EVOLUTION OF MOLECULAR FUNCTION IN MAMMALIAN NEURONS

Chantal Francis

A DISSERTATION

in

Biology

Presented to the Faculties of the University of Pennsylvania

in

Partial Fulfillment of the Requirements for the Degree of Doctor of Philosophy

2011

Supervisor of Dissertation

Signature _____

Junhyong Kim, Ph.D.

Co-Supervisor of Dissertation

Signature _____

Jim Eberwine, Ph.D.

Graduate Group Chairperson

*Signature*_____

Doris Wagner, Ph.D.

Dissertation Committee

Nancy Bonini, Ph.D., Professor of Biology, (Dissertation Committee chair)

Maja Bucan, Ph.D., Professor of Genetics

Dustin Brisson, Ph.D., Assistant Professor of Biology

Marc Schmidt, Ph.D., Associate Professor of Biology

EVOLUTION OF MOLECULAR FUNCTION IN MAMMALIAN NEURONS

Chantal Francis

2011

DEDICATION

I would like to dedicate this dissertation to my family for all their love and support.

Especially to my beloved husband Peter.

ACKNOWLEDGEMENTS

There are no shortcuts for an achievement. None of this work would have been possible without the intellectual and mental support of many, many people to whom I am deeply grateful.

I would like to acknowledge the members of my thesis committee: Dr. Nancy Bonini, Dr. Maja Bucan, Dr. Dustin Brisson, and Dr. Marc Schmidt for great advice and constructive criticism throughout these years. Their advice and suggestions helped me to refine and achieve my thesis goals.

I would also like to acknowledge all past and current members of the Kim lab: Dan, Dexter, Fan, Hannah, Hoa, Kathryn, Miler, Mugdha, Sheng, Saurav, Shreedhar, and Stephen. Their different backgrounds and amazing computational skills helped me incredibly and allowed me to move forward in my data analysis. Despite the diversity of their work, all Kim lab members were united by their friendliness. This is so important in making your graduate student life more enjoyable! Dan and Miler were always so friendly and available to help me with my basic programming skills. They both are very organized and such talented teachers. They were so patient with me! From my collaboration with Shreedhar, I also learned many aspects related to computational biology as well as life. Stephen's contribution, creating an online database, was extremely valuable to my *in situ* project. His extensive computer skills and wise advice helped me avoid lots of computer crashes and data loss. I am grateful to Kathleen (an "unofficial" Kim lab member) for her great advice on surviving the pressures of dissertation writing and for making sure I had something to eat for my late nights in the lab. The hub of talent and kindness from the Kim lab is but a small reflection of who Junhyong is. My first interaction with Junhyong was as a technician and throughout the years the more I interacted with him the deeper my admiration became. It is sincerely impossible for me to describe in few words all my gratitude for such an empathic boss, knowledgeable teacher and advisor, and considerate human being.

One of the greatest things that happened to me as a graduate student was to perform my thesis work in a partnership with Dr. Jim Eberwine. This collaboration quickly became the keystone of my thesis thanks to Jim's generosity and welcome in his lab. Jim was an incredible co-mentor from whom I have learned to be more confident in my thoughts and to think more as a scientist. Beside his extreme talent and passion for

science, Jim was also very understanding and supportive of my family obligations. For this, I am sincerely thankful. Working in Jim's lab not only allowed me to learn many new technical skills but also allowed me the unique opportunity to meet and interact with his lab members. I would like to acknowledge Ditte, Jackie, Jai, Jae-Hee, Jenn, Jeanine, Kevin, Pete, Sean, Terri, Tiina and Thom. All of them helped in some way ranging from teaching me single cell dissection, *in situ* techniques and the beauty of microscopy to even providing advice on my writing. Beside their technical assistance, they have been great co-workers reviving the bench work area with fun discussions about food, American sports games, which was a big novelty for me that I am still learning a lot about, or Nobel prizes.

I have to acknowledge that none of my experiments would have been feasible without the thousands of primary neuron cultures that Margie prepared for me. It was always a pleasure going to Margie's lab to pick up my samples and to see her bright smile that cheered up my day. I am also grateful to the very talented students Irmina, Emily and John who helped me on the immense and tedious tasks of the *in situ* project during their internship in Jim's lab.

I must also thank Dr. Brian Gregory who opened his lab to me, and walked me through all of the steps in the library preparation of next generation sequencing.

Being a TA during the first two years of Graduate school was tough but such a great learning experience. It made me realize how challenging it is to convey your thoughts and understanding to others and that teaching a subject to others is far different from understanding it yourself. I am grateful to Dr. Linda Robinson for training me and for giving me great tips about speaking in public and dealing with delicate situation with students.

The underlying force behind my choice to enter graduate school comes from Paul Sniegowski's lab. I came to the U.S. with a master's degree and the thought that I was done as far as my education. Joining Paul's lab revived my passion for science and pushed my ambition further. Paul was an amazing teacher and his former graduate student's, Heidi and Helen, enthusiasm and persistence in their work was a key motivation for me.

My graduate student years were enormously enriched in basic science, technical skills and wonderful discussions with very bright people…it was also full of ups and downs. Junhyong once said: "Graduate school is like a marathon". It is all about

endurance, learning from the down side and gaining even more force from it to pursue your goals to the end, even though science never ends.

If I was able to put this dissertation together and stay physically and mentally sane, it is only because of the support of my friends, family and beloved son and husband.

To Helen, Maurits, Vero, Paul, Sophie, and Hélène, meeting you during grad school was such a breath of fresh air! To Danielle & Amr, our friendship and your advice were so valuable to me throughout the toughest moments of my grad school. To Milay & David, meeting with you and discussing everything but science felt so good. Our interactions, your Cuban food and parties always reminded me that life is fun!

My lovely mom, Evelyne, taught me that anyone could reach his or her goals with patience and determination. She always reminded me that smiling is free and can make ones life so much brighter. Thanks to her and my three cherished brothers Tony, Richard & Alain I became the person I am today, always thriving to learn more and believe in and pursue my passion.

Having Téo, my son, during my third year of graduate school was a big challenge but by far the most beautiful thing that ever happened to me. No matter how busy I was or if lab experiments were depressingly not working, seeing Téo at the end of the day and playing with him like a kid magically recharged all my batteries and gave me more energy and efficiency upon returning to work.

Last and by far not the least, my love, my best friend, my everything: Peter. He always believed in me, much more than I did in myself. Without him I wouldn't have had the time to fulfill my mommy tasks and enjoy quality time with my family. Without him, I would be eating canned and frozen food every single day. Without him dragging me to bed late at night, I would have become a sleep-deprived zombie. His love, patience and understanding are the secret of my achievement. Thanking him is never enough but all I can say is that this dissertation would not have been possible without his support.

ABSTRACT

EVOLUTION OF MOLECULAR FUNCTION IN MAMMALIAN NEURONS

Chantal Francis

Junhyong Kim, Ph.D., Jim Eberwine, Ph.D.

A common question in neuroscience is what forms the neurological basis of the variety of behaviors in mammals. While many studies have compared mammalian anatomy and evolution, few have investigated the physiologic functions of individual neurons in a comparative manner. Neurons' ability to modify their features in response to stimuli, known as synaptic plasticity, is fundamental for learning and memory. A key feature of synaptic plasticity involves delivering mRNA to distinct domains where they are locally translated. Regulatory coordination of these events is critical for synaptogenesis and synaptic plasticity as defects in these processes can lead to neurological diseases. In this work, we combine computational and experimental biology to investigate subcellular localization of mRNAs in dendrites of mouse and rat neurons. Differential subcellular localization of specific gene products may highlight differential synaptic function and allow the uncovering of evolutionary conserved as well as divergent molecular functions in neurons.

First, we performed a comparative analysis of the dendritic transcriptome in mouse and rat via microarrays. We found that their dendritic transcriptome are significantly more divergent than other homologous tissues and that these evolutionarily changes could be associated with transposon activity. Second, we

comprehensively determined subcellular expression patterns for neuronal genes in mice and rats via a systematic *in situ* survey on a curated list of dendritic mRNA from our previous microarray study. This survey highlighted that dendritic localization of specific transcripts occurs in a species-specific fashion. We uncovered species-specific *cis* and *trans*-elements with possible implications in transcript localization and gene expression regulation in neurons. The interactions between these elements might play a major role in the proper development and evolution of complex nervous systems. Our data will be publically available in a database (http://kim.bio.upenn.edu/insitu/public/), which could guide future investigations. Finally, we investigated single cell variability by combining microarrays, RNAseq and *in situ* experiments. Our preliminary results underlined the extent of this variability and its contributions in establishing cell's unique identity.

In conclusion, our study suggests the existence of species-specific mRNA localization mechanisms and supports the idea that evolution of phenotypes might be linked to the evolution of subcellular localization of transcripts.

Contents

List of Tables

List of Figures

Chapter 4: SINGLE-CELL VARIABILITY: FICTION OR REALITY?

Chapter 1

INTRODUCTION

1.1 INTRODUCTION TO BRAINS, NEURONS AND EVOLUTION

Mammalian behavior is extremely heterogeneous: spanning from complex language, social interactions and behaviors to very simple stimulus responses. What is the neurological basis for such a variety of behavioral traits in mammals? A classical answer would relate the complexity of an organism's behavior to the complexity of its nervous system. Comparative neuroanatomy studies have led to an understanding of brain differences and evolution between species [1, 2]. As we move from invertebrates to vertebrates, brain evolution is illustrated by a dramatic increase in its size and complexity [3]. Particularly, the increased structural complexity of the cerebral cortex is one of the hallmarks of the brain's evolution. Increased number of convolutions in the brain is a characteristic of species with more advanced brains and more complex behavior. In early vertebrates' brains, the neocortices were composed of only a few cortical areas, and with more evolved lineages, such as primates, the neocortex expanded dramatically [3]. Large numbers of studies have compared mammalian brain anatomy, architecture, and evolution, but few have investigated the physiological functions of individual neurons in a comparative manner.

Neurons, the basic building blocks of the nervous system, are highly polarized cells comprised by three main parts: the dendrites, the cell body and the axon [4]. The dendrites and axon are designed to receive and transmit information respectively (See Figure 1.1).

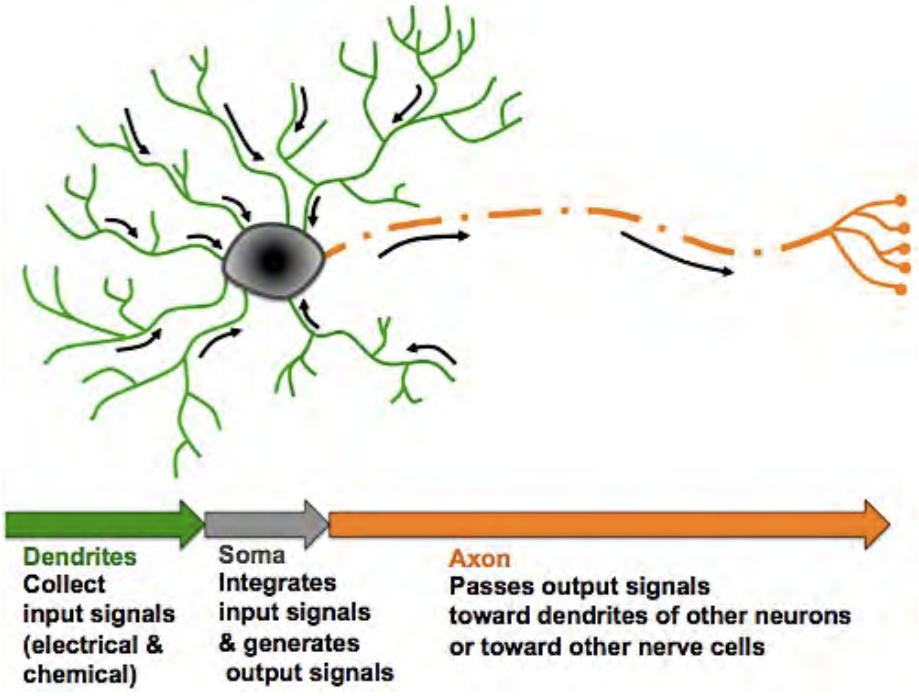


Figure 1.1: Flow of information in a neuron

A remarkable characteristic seen in all nerve cells is the similarity in signaling properties. In general, neuron-to-neuron communication is based upon electric signals that travel through the axon of the presynaptic cell leading to neurotransmitter release, which interacts with the postsynaptic dendrites. Intracellular and extracellular stimuli induce a series of highly coordinated cellular and molecular interactions within neurons that allow them to adjust their

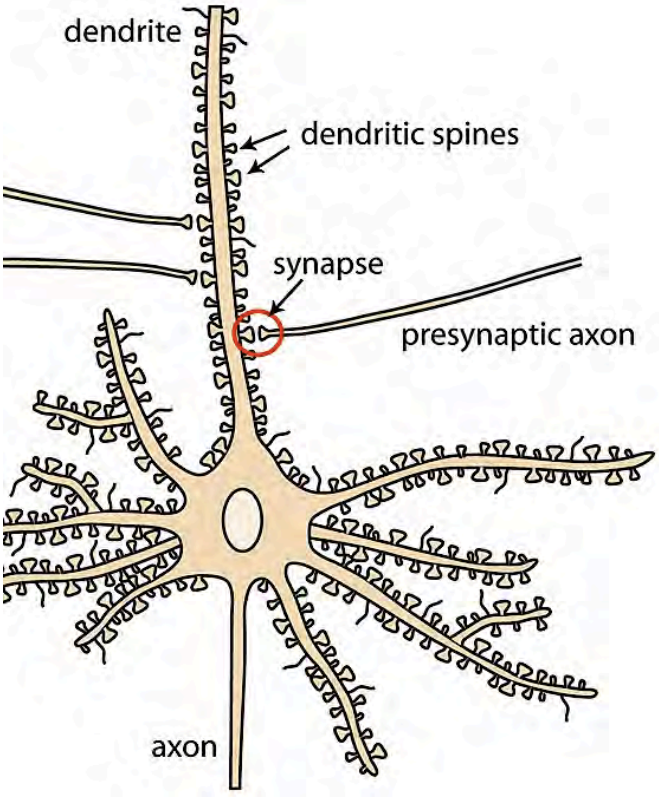
response accordingly [5, 6]. Neurons are interconnected by synapses, which are the major mediators of communication in nervous systems. Their abundance, up to a thousand per single-cell, allows wide connections between neurons that enable the formation of a large information-processing network. Far from being a simple track passively transmitting signals from a presynaptic neuron to a postsynaptic neuron, a synapse constantly changes its transmission efficiency based on its recent history. This remarkable ability, called synaptic plasticity, most likely generates the cellular basis of learning and memory [7, 8].

A key feature of synaptic plasticity involves gene expression and the delivery of mRNAs to distinct subcellular domains where their localized translation is thought to play a role in affecting synaptic efficacy [9]. In the past two decades, research on neuronal gene expression has provided strong evidence on the importance of mRNAs localization, post-transcriptional modifications, and translation in dendrites.

Our main objective is to elucidate how the molecular programs controlling and regulating gene expression are manifested in rodent neurons, particularly in the dendrites, and to test the hypothesis that organism-specific events regulate synaptic plasticity mechanisms and gene expression and localization in dendrites. Before focusing on our goal, it is worth reviewing the main characteristics of dendrite structure and evolution, as well as the key findings on the relevance of mRNA localization in cells and particularly in neurons.

1.2 DENDRITE STRUCTURE

Dendrites were first described by Santiago Ramòn y Cajal (1852-1934), the founder of modern neuroscience, as the receptive surfaces of neurons. Dendrites receive synaptic inputs from other presynaptic neurons' axons, and then process and conduct this activity to their own corresponding axon. The received inputs occur either directly on the shaft of the dendrites or on the spines, which are specialized protrusions (~2µm) ending in a bulbous head (See Figure 1.2)



Smrt and Zhao, Frontiers Biology 2010

(Permission to reproduce this image given by Richard Smrt, Ph.D.)

Figure 1.2: Dendritic Spines' distribution on a neuronal cell

Dendrites exhibit a wide variety of forms. They are characterized by extensive branching, known as arborizations, which spread into specific domains. Unlike axons, dendritic connections are relatively local extending between 15 µm to 2000 µm. Their shape and composition are continuously influenced by the environment [10]. Several studies, mainly based on electron microscopy, have highlighted their broad variety and revealed their inclusion of main organelles such as the Golgi apparatus, mitochondria, the rough and smooth endoplasmic reticulum, ribosomes and polyribosomes. Moreover, based on their branching patterns, neurons have been classified as unipolar, bipolar, multipolar, conical or pyramidal, to name a few.

During his extensive study on the nervous system, Ramòn y Cajal noticed that the complexity of dendrites reflects the complexity and number of connections its neuron has. The wide diversity seen in the structure, composition and plasticity of dendrites suggests broad functional contributions of these structures to brain health. Indeed, dendritic morphology has proven important in the context of normal brain function. A tight correlation was shown to exist between several neurological diseases and dendrite pathology. Reductions in dendritic arborization have been reported in several pathological conditions, such as mental retardation, Alzheimer’s disease or even aging. Appearance of dendritic varicosities like protein aggregations or vacuolar dystrophy, have been

reported in Parkinson's and Creutzfeldt-Jakob disease respectively. Moreover, changes in the dendritic spines density, either by increasing or decreasing, have been reported in several pathologies such as ADHD, autism, mental retardation, and fragile X syndrome.

Understanding the evolution of dendrites and the relationship between dendrites morphology and function is essential for understanding their contribution to brain function.

1.3 EVOLUTION OF DENDRITES

Dendrites from various neurons share similar basic features but they also exhibit a large divergence in their shape, molecular composition and ability to process and conduct action potentials. The large diversity reported in dendrites both within and between species implies their evolutionary adaptation to a wide range of functionalities. Most of the evolutionary changes in dendrites that have been reported are mainly based on comparative studies in anatomy, biophysics and biochemistry.

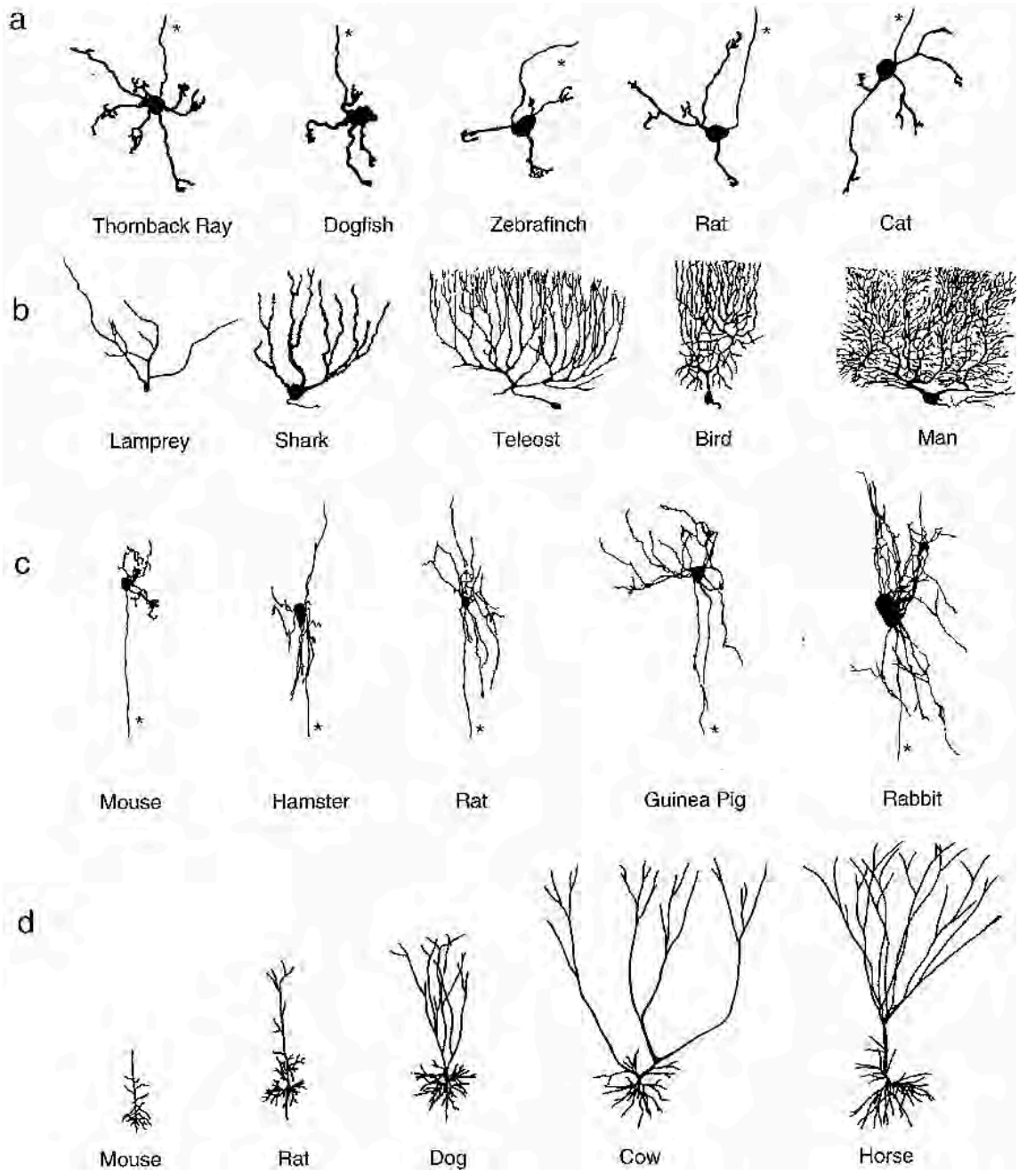
1.3.1 Comparative anatomy: scaling of dendritic arbors with brain size

More than a hundred years ago, Santiago Ramòn y Cajal stated the fundamental observation that neurons increase in complexity as the animal size and brain size increase. Throughout the years, more comparative studies refined

Cajal's original observations and highlighted great structural complexity in neurons from different brain sizes.

In general, cell types that have continuous ability to divide have roughly the same size across species but their number in a given organ increase as the organs scale with the body size [4, 11]. Many cellular functions in these "invariant" cell types are expected to scale up with the cell number; for instance, doubling the red blood cells numbers would double the amount of oxygen carried. As for non-dividing cells, like neurons, they tend to become more complex and larger as the overall brain size or body size increases (except Cerebellar granule cells). It is known now that the information-processing capacities of the brain depend not only on the number of neurons but on the neuronal interconnected network and its size.

Even if their general branching patterns is preserved, variation of dendritic complexity in different neuronal types across species has been reported by several investigations. Figure 1.3 illustrates some of the species diversity in the architecture of dendrites. The divergence seen in dendrites branching patterns among neuron classes suggests that these different cell types have evolved either to perform different computations or under different sets of constraints.



(a) Cerebellar granule cells (b) Cerebellar Purkinje neurons (c) Sympathetic neurons (d) Neocortical pyramidal neurons (*) axons

Figure 1.3: Comparison of vertebrate dendrites across phylogeny

1.3.2 Comparative electrophysiology

Dendrites vary not only in their shape but also in their diameter along their length. As they become larger and more branched, voltage signals travel a longer distance before reaching the cell body. The greater the distance the more filtering properties of the dendritic membrane are involved. A tight relationship exists between the diameter and the signaling properties of the dendrites. For instance, in wider dendrites, the passive and active signals propagations are faster [12, 13]. Across species, some types of neurons keep a constant diameter of their dendrites, such as the reticula of the substancia nigra [14], while other neurons, such as the CA1 pyramidal neurons, have a dendritic diameter and thickness that scales with the size of the cell [15]. In general, independently of the size or pattern of the dendrites, the relative efficacies of the synapses on different locations are preserved.

1.3.3 Comparative molecular physiology

The processing of input signals in dendrites is based on the composition and spatial distribution of the passive and active ions channels and depends on many neurotransmitters, neuromodulators, postsynaptic receptors and signal transduction molecules. These main molecular building blocks are conserved across the vertebrates; however, their level of expression could vary extensively

which might have a great effect on the firing properties of neurons and potentially could impact synaptic plasticity [16].

Most of the recent systematic approaches to study neuron evolution are based on comparative proteomics and genomics together with molecular phylogenetic approaches. These comparative studies, done in majority in the *Aplysia genus*, have reinforced the understanding of the relationships between species and the mechanisms of molecular evolution. In 1997, Gilly and colleagues [17] showed that gating parameters may have adapted between several Mollusks species as animals with a faster rate of movement have also a faster gating of sodium channels in neurons. More recently, a comparative transcriptomic analysis on neurons from two *Aplysia* species highlighted that genes involved in signal transduction have a higher rate of evolution of their proteins than the housekeeping genes. This study also provided a list of new candidate genes that might be involved in synaptic plasticity, learning and memory [18].

Even with all the advances in the “omics”, multi-species investigations on the evolution of neuronal functions are still very sparse. Very few studies have investigated the molecular functions of individual neurons in a comparative manner across species.

1.4 IMPLICATION OF DENDRITES IN SYNAPTIC PLASTICITY

A major function of dendrites is their involvement in synaptic plasticity. Based on recent activity, a synapse modifies its efficiency to either transmit a strong long lasting in signal, in a process known as Long Term Potentiation (LTP), or a reduced one via Long Term Depression (LTD). These two opposing mechanisms are widely believed to be the molecular basis of learning and memory. A key feature in both LTP and LTD involves delivery of mRNAs to distinct sub-cellular domains where their localized translation is thought to play a role in affecting synaptic efficacy [19]. In fact mRNAs localization has been shown to be a critical event in gene expression in a wide variety of organism [20]. I will first give a brief review of the main findings on the relevance of transcripts' sub-cellular localization in various systems then I will narrow down to investigations done more specifically on transcripts in neurons.

1.5 IMPORTANCE OF MRNA LOCALIZATION IN CELLS

Cells are organized into compartments, membranes, and organelles that perform specialized biological and regulatory functions. All cells are more or less polarized and this polarity is essential for their viability.

Until two decades ago, protein specific localization was thought to occur only after translation. Nowadays, many studies have shown that local protein synthesis can also be regulated at the mRNA level and that post-transcriptional processing such as splicing, editing and translation are critical steps in gene expression regulation [21-23].

The phenomenon of mRNA subcellular localization has been conserved throughout evolution and used by organisms to control protein synthesis to specific regions of the cell. RNA localization has been described in a large variety of species ranging from yeast to mammals. It occurs in both dividing cells such as fibroblasts as well as post-mitotic cells such as neurons [21]. Recent studies, particularly in *Drosophila* embryos, showed that the majority of mRNAs were asymmetrically localized highlighting the fact mRNA localization is very widespread [24].

Indeed, several potential advantages support RNA localization compared to protein localization. First, it is very cost-effective for the cell to perform several rounds of protein synthesis from a localized mRNA instead of moving each protein molecule individually [25]. That is, the movement of a small number of RNA molecules and local translation can have the same effect as the movement of very large number of protein molecules. Second, localized translation allows a fast and efficient spatio-temporal control of cellular response to external stimuli, which is particularly relevant for neurons [26, 27]. Third, mRNA localization allows the restriction of translation to specialized organelles or cellular domains [28-32]. Particularly, this process helps prevent protein presence in ectopic areas that might have a harmful outcome on the cell. For example, the mislocalization of *Oskar* mRNA, which is normally localized to the posterior half of the *Drosophila* oocyte, leads to the formation of a second abdomen [33, 34]. In general, targeted mRNAs were shown to be involved in various biological functions such as oocyte differentiation with the establishment

of morphogenesis gradients [33, 35, 36], early stage development of embryos [37-43], or even the establishment of synaptic plasticity [44].

1.6 IMPORTANCE OF MRNA LOCALIZATION IN DENDRITES

mRNA localization is particularly important in highly polarized cells such as neurons where the final destination of a protein could be spatially very far from its original site of transcription. Indeed, given its structure, the neuron must overcome several molecular challenges when executing even basic cellular functions like protein synthesis. Long physical distances exist between the soma, the distal axonal, and dendritic compartments, which makes it difficult for neurons to generate fast responses to environmental changes or synaptic input. Moreover, during development, complex regulatory mechanisms are needed to confer a spatial patterning to neurons [45, 46]. Therefore, it is favorable for the cell to rely on localized pools of RNAs that can be rapidly translated when needed instead of being solely dependent upon protein synthesis in the cell body and their trafficking toward the dendritic region of interest.

The idea of mRNA transport and local protein synthesis in neurons was first proposed when ribosomes were revealed via electron microscopy in dendrites almost fifty years ago [47]. It has since been demonstrated that a large population of transcripts are confined in dendrites [48-50] and that these transcripts not only are locally translated [9, 51-56] but also locally edited [57].

This localized control of gene expression in dendrites allows much faster and accurate spatio-temporal control of cellular response to environmental changes or synaptic input. Having a population of incompletely processed RNA present and available at a synapse would allow a deeper level of response to external stimuli and a potentially large diversity of protein to be synthesized from the same original gene.

The well characterized dendritic transcripts fall into a wide variety of categories from structural proteins (MAP2), enzymes (CamK2a), growth factors (BDNF), ligand or voltage gated ions channels (Glutamate and GABA receptor subunits; Calcium channels) and even transcription factors (CREB) to name a few [55, 58-61].

1.7 MECHANISMS OF RNA TRANSPORT TO DENDRITES

How are certain populations of transcripts selected for transport into dendrites and how does this targeting mechanism occur? Many evidences have ruled out the hypothesis of simple passive diffusion as only a subpopulation of mRNAs is found in dendrites. The most robust explanation relies on mRNA association with specific proteins facilitating their transport. These transport granules, referred as messenger ribonucleoproteins (mRNPs), are macromolecular structures (200-600µm of diameter) made by a combination of RNAs, ribosomes and/or proteins. These structures are not thought to be transcriptionally competent but rather a storage place that would release mRNAs

upon receiving specific signals [62, 63]. A variety of mRNAs have been identified within the mRNPs such as the calcium/calmodulin-dependent protein kinase 2-alpha (Camk2a), the activity-regulated cytoskeletal-associated (Arc), beta-actin, and the noncoding BC1. Numerous biochemical and proteomic studies suggest a heterogeneous composition of the mRNPs. These could contain a combination of several different RNA binding proteins (RBPs) such as Staufen, Zipcode-binding protein1 (ZBP1), heteronuclear RNP-A2 (hnRNP-A2), Pur-alpha and the fragile X mental retardation protein (FMRP) as well as multiple cytoskeletal motors such as the kinesin (Kif5) and dynein which could facilitate the transport along dendrites' microtubules [64-68].

The RBPs are highly involved in selectively targeting certain mRNAs to dendrites by recognizing specific sequences (i.e. Cytoplasmic Polyadenylation Element, CPE, in the 3'UTR) or secondary structures within the mRNAs [69-71]. The identification of these targeting elements has been very challenging. Table1.1 lists examples of the ones that have been reported in neurons.

Table 1.1: List of mRNAs whose localization elements have been characterized in neurons.

Transcript	Organism	Subcellular Location	Targeting region	Minimum length	Binding protein	Ref.
Arc	Rat	Dendrites	3'UTR	350nt, 370nt		[72]
βactin	Rat	Dendritic filopodia	3'UTR	ACACCC within 54nt	ZBP-1	[73]
BC1	Rat	Dendritic filopodia	5'UTR	62 nt stem loop	ZBP-1	[62, [74]
BNDF	Mouse, Rat	Dendrites	3'UTR	Unknown		[75]
CamK2a	Rat	Dendrites	3'UTR	CPE (UUUUUUAUU X2 separated by 82nt)	CPEB	[76]
CamK2a	Rat	Dendrites	3'UTR	1200nt, 30nt	CPEB	[77]
Dendrin	Rat	Dendrites	3'UTR	1000nt		[78]
IP3RI	Mouse	Dendrites	3'UTR	Unknown	Hfz	[79]
MBP	Rat	Myelin sheath	3'UTR	GCCAAGGAGUC	hnRNP-A2	[80]
MBP	Rat	Oligodendrocyte processes	A2RE (11nt)	GCCAAGGAGCC	hnRNP-A2	[81]
MAP2	Rat	Dendrites	3'UTR	640nt	MARTA1/2	[82]70]
Nanos	*Drosophila*	Dendrites	3'UTR	Unknown		[83]
Neurogranin	Rat	Dendrites	3'UTR	30nt		[76]
NMDA_NR1	Rat	Dendrites	5'UTR	24nt	60-70KDa proteins	[84]
PKMz	Rat	Dendrites	3'UTR	84nt		[85]
RhoA	Rat	Axons	3'UTR	Unknown		[85, 86]
Shank1	Rat	Dendrites	3'UTR	200nt		[87]
Syntaxin	*Aplysia*	Axon hillock	3'UTR	CPE	CPEB	[88]
Tau	Rat	Axons	3'UTR	240nt	HuD	[89]
Tau	Rat	Axons	3'UTR	91nt	Ilf3; NF90	[90] 79]
Vasopressin	Rat	Dendrites	ORF+3'UTR	Unknown		[91]81]

BDNF = brain-derived nerve factor mRNA
CamK2a = calcium/calmodulin-dependent protein kinase II alpha mRNA
hnRNP-A2 = heteronuclear ribonucleoprotein-A2
IP3RI = type1 inositol 1,4,5-trisphosphate receptor mRNA
MAP2 = microtubule associated protein 2 mRNA
MARTA1/2 = MAP2 RNA trans-acting protein 1 and 2
CPE = Cytoplasmic polyadenylation element binding protein
PABP = poly(A) binding protein
CPEB = CPE binding protein
UTR = Untranslated regions
ORF = Open Reading Frame
NMDA = N-Methyl-D-aspartate
ZBP-1 = zipcode-binding protein 1
nt = nucleotides

Additionally, recent investigations have shown that, in response to extracellular stimuli, the 3'UTR region of a large number of mRNAs is not only involved in coordinating their targeting but also in regulating their expression at

specific locations [75, 92-96]

Non-coding regions have also gained increased attention as introns' retention in some transcripts was shown to be critical for their subcellular localization and for dendritic neuro-physiology [97]. Particularly, in a novel study, rodent ID elements, members of the interspersed nuclear elements (SINEs) family, were shown to positively correlate with genes expressed in the brain [98, 99] and their presence within cytoplasmic intron sequence-retaining transcripts (CIRTs) was shown to mediate dendritic localization of several transcripts [100].

Even though the number of known localized transcripts has grown in the past few decades [21-23, 67], and that recent findings allowed a better understanding of the process of mRNAs targeting, many aspects of these events are still unclear like their prevalence, variety or evolutionary conservation.

1.8 SPECIFIC AIMS

In order to get a better understanding of localized neuronal regulatory mechanisms, and to assess whether dendritic molecular physiology is conserved, I investigated subcellular localization of mRNAs in dendrites of both model organisms mouse and rat. The differences in subcellular localization of specific transcripts in these two rodents may highlight differential synaptic function and allow uncovering of evolutionary conserved as well as divergent molecular functions in neurons. In this thesis work, I combined both computational and different experimental techniques and technologies to

address this question and characterize systematically rat and mouse dendritic transcriptome.

First, I focused my investigation on amplifying dendritic mRNAs from single neurons to best assess this population and to overcome the issue of cellular heterogeneity in brain tissue. Using microarrays, I performed a comparative analysis of the dendritic transcriptome in mouse and rat that resulted in a comprehensive characterization of their dendritic transcripts (Chapter 2). Second, I defined the dendritic transcriptome and its spatial patterns through a large-scale *in situ* hybridization (Chapter 3). This novel systematic *in situ* survey is expected to be a new standard for mRNAs subcellular localization in rodent neurons. Third, in addition to comparative functional genomics work in dendrites of neurons, I also collaborated on a research to delineate the single-cell RNA variability of rat and mouse neurons. In order to establish a reference for single cell analysis and the usage of antisense RNA (aRNA) amplification procedure, I carried out series dilution experiment through a combination of microarrays and RNAseq experiments (Chapter 4). Finally, I performed *in situ* hybridizations on candidate transcripts with different 3'UTR isoforms in the effort of investigating their potential role in single cell variability (Chapter 4). The flow chart in Figure 1.4 displays my experimental outline.

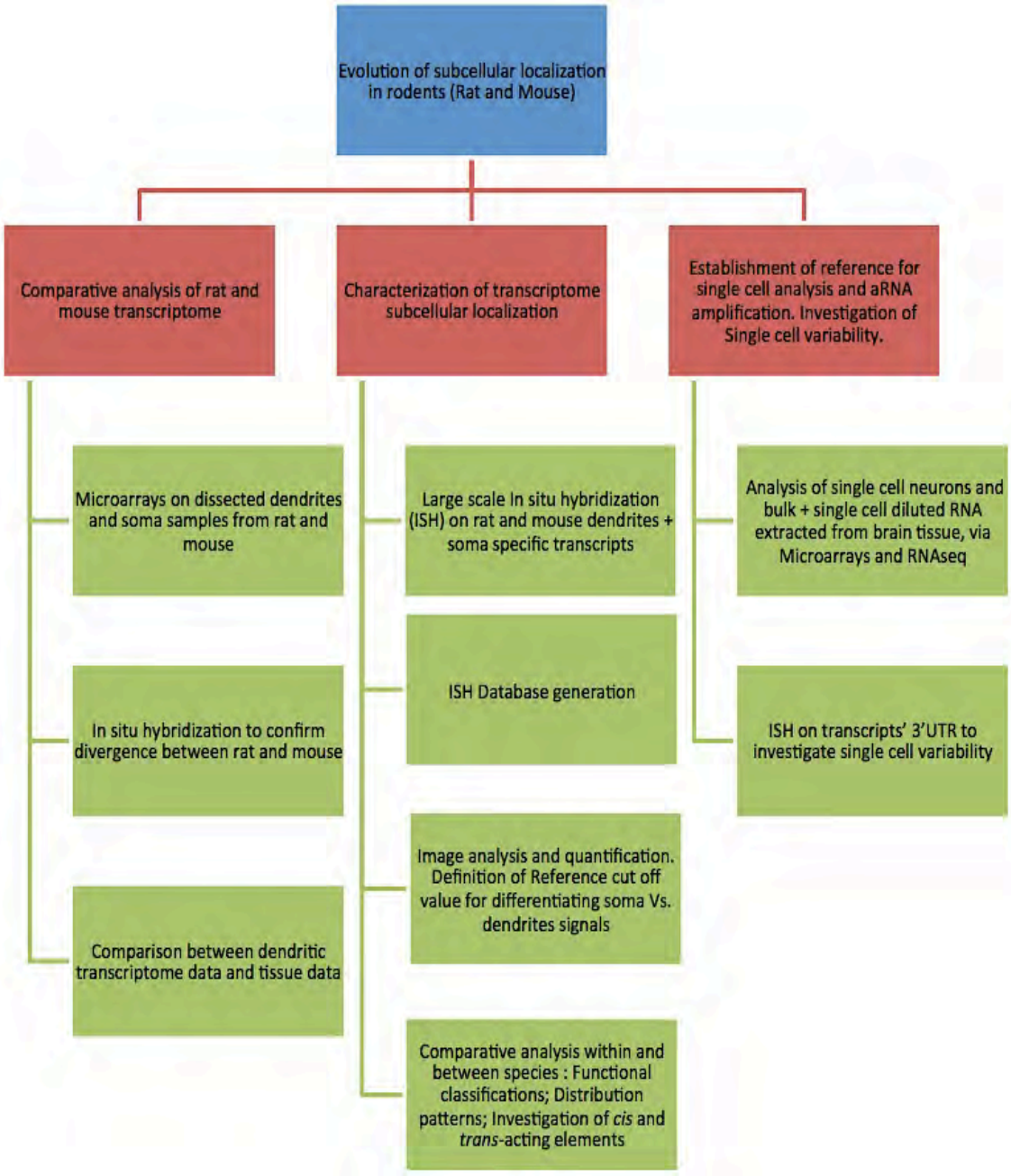


Figure 1.4: Experimental Outline

REFERENCES

1. Carroll, S.B., *Genetics and the making of Homo sapiens.* Nature, 2003. **422**(6934): p. 849-57.
2. Crick, F. and C. Koch, *A framework for consciousness.* Nat Neurosci, 2003. **6**(2): p. 119-26.
3. Karlen, S.J. and L. Krubitzer, *The evolution of the neocortex in mammals: intrinsic and extrinsic contributions to the cortical phenotype.* Novartis Found Symp, 2006. **270**: p. 146-59; discussion 159-69.
4. Huxley, J., *problems of relative growth*, ed. Methuen1932, London.
5. Loftus, S.K. and W.J. Pavan, *The use of expression profiling to study pigment cell biology and dysfunction.* Pigment Cell Res, 2000. **13**(3): p. 141-6.
6. Moggs, J.G. and G. Orphanides, *The role of chromatin in molecular mechanisms of toxicity.* Toxicol Sci, 2004. **80**(2): p. 218-24.
7. Bailey, C.H., D. Bartsch, and E.R. Kandel, *Toward a molecular definition of long-term memory storage.* Proc Natl Acad Sci U S A, 1996. **93**(24): p. 13445-52.
8. Bliss, T.V. and G.L. Collingridge, *A synaptic model of memory: long-term potentiation in the hippocampus.* Nature, 1993. **361**(6407): p. 31-9.
9. Steward, O. and E.M. Schuman, *Compartmentalized synthesis and degradation of proteins in neurons.* Neuron, 2003. **40**(2): p. 347-59.

10. Holtmaat, A. and K. Svoboda, *Experience-dependent structural synaptic plasticity in the mammalian brain.* Nature reviews. Neuroscience, 2009. **10**(9): p. 647-58.

11. Thompson, D.W., *On growth and form*, ed. C.U. Press1942, Cambridge.

12. Purves, D. and J.W. Lichtman, *Geometrical differences among homologous neurons in mammals.* Science, 1985. **228**(4697): p. 298-302.

13. Olsen, O., et al., *Uniform growth and neuronal integration.* Journal of neurophysiology, 1996. **76**(3): p. 1850-7.

14. Yelnik, J., et al., *Golgi study of the primate substantia nigra. I. Quantitative morphology and typology of nigral neurons.* The Journal of comparative neurology, 1987. **265**(4): p. 455-72.

15. Bekkers, J.M. and C.F. Stevens, *Two different ways evolution makes neurons larger.* Progress in brain research, 1990. **83**: p. 37-45.

16. Raymond, J.L., S.G. Lisberger, and M.D. Mauk, *The cerebellum: a neuronal learning machine?* Science, 1996. **272**(5265): p. 1126-31.

17. Gilly, W.F., R. Gillette, and M. McFarlane, *Fast and slow activation kinetics of voltage-gated sodium channels in molluscan neurons.* Journal of neurophysiology, 1997. **77**(5): p. 2373-84.

18. Choi, S.L., et al., *Differential evolutionary rates of neuronal transcriptome in Aplysia kurodai and Aplysia californica as a tool for gene mining.* Journal of neurogenetics, 2010. **24**(2): p. 75-82.

19. Goelet, P., et al., *The long and the short of long-term memory--a molecular framework.* Nature, 1986. **322**(6078): p. 419-22.

20. Sylvestre, J., et al., *The role of the 3' untranslated region in mRNA sorting to the vicinity of mitochondria is conserved from yeast to human cells.* Mol Biol Cell, 2003. **14**(9): p. 3848-56.

21. Bashirullah, A., R.L. Cooperstock, and H.D. Lipshitz, *RNA localization in development.* Annual review of biochemistry, 1998. **67**: p. 335-94.

22. Kloc, M., N.R. Zearfoss, and L.D. Etkin, *Mechanisms of subcellular mRNA localization.* Cell, 2002. **108**(4): p. 533-44.

23. Czaplinski, K. and R.H. Singer, *Pathways for mRNA localization in the cytoplasm.* Trends Biochem Sci, 2006. **31**(12): p. 687-93.

24. Martin, K.C. and A. Ephrussi, *mRNA localization: gene expression in the spatial dimension.* Cell, 2009. **136**(4): p. 719-30.

25. Jansen, R.P., *mRNA localization: message on the move.* Nature reviews. Molecular cell biology, 2001. **2**(4): p. 247-56.

26. Cohen, S. and M.E. Greenberg, *Communication between the synapse and the nucleus in neuronal development, plasticity, and disease.* Annual review of cell and developmental biology, 2008. **24**: p. 183-209.

27. Flavell, S.W. and M.E. Greenberg, *Signaling mechanisms linking neuronal activity to gene expression and plasticity of the nervous system.* Annual review of neuroscience, 2008. **31**: p. 563-90.

28. Lawrence, J.B. and R.H. Singer, *Intracellular localization of messenger RNAs for cytoskeletal proteins.* Cell, 1986. **45**(3): p. 407-15.

29. Mingle, L.A., et al., *Localization of all seven messenger RNAs for the actin-polymerization nucleator Arp2/3 complex in the protrusions of fibroblasts.* Journal of cell science, 2005. **118**(Pt 11): p. 2425-33.

30. Zhang, H.L., et al., *Neurotrophin-induced transport of a beta-actin mRNP complex increases beta-actin levels and stimulates growth cone motility.* Neuron, 2001. **31**(2): p. 261-75.

31. Lambert, J.D. and L.M. Nagy, *Asymmetric inheritance of centrosomally localized mRNAs during embryonic cleavages.* Nature, 2002. **420**(6916): p. 682-6.

32. Adereth, Y., et al., *RNA-dependent integrin alpha3 protein localization regulated by the Muscleblind-like protein MLP1.* Nature cell biology, 2005. **7**(12): p. 1240-7.

33. Ephrussi, A., L.K. Dickinson, and R. Lehmann, *Oskar organizes the germ plasm and directs localization of the posterior determinant nanos.* Cell, 1991. **66**(1): p. 37-50.

34. Ephrussi, A. and R. Lehmann, *Induction of germ cell formation by oskar.* Nature, 1992. **358**(6385): p. 387-92.

35. Driever, W. and C. Nusslein-Volhard, *A gradient of bicoid protein in Drosophila embryos.* Cell, 1988. **54**(1): p. 83-93.

36. Gavis, E.R. and R. Lehmann, *Localization of nanos RNA controls embryonic polarity.* Cell, 1992. **71**(2): p. 301-13.

37. Broadus, J., S. Fuerstenberg, and C.Q. Doe, *Staufen-dependent localization of prospero mRNA contributes to neuroblast daughter-cell fate.* Nature, 1998. **391**(6669): p. 792-5.

38. Hughes, J.R., S.L. Bullock, and D. Ish-Horowicz, *Inscuteable mRNA localization is dynein-dependent and regulates apicobasal polarity and spindle length in Drosophila neuroblasts.* Current biology : CB, 2004. **14**(21): p. 1950-6.

39. Li, P., et al., *Inscuteable and Staufen mediate asymmetric localization and segregation of prospero RNA during Drosophila neuroblast cell divisions.* Cell, 1997. **90**(3): p. 437-47.

40. Neuman-Silberberg, F.S. and T. Schupbach, *The Drosophila dorsoventral patterning gene gurken produces a dorsally localized RNA and encodes a TGF alpha-like protein.* Cell, 1993. **75**(1): p. 165-74.

41. Simmonds, A.J., et al., *Apical localization of wingless transcripts is required for wingless signaling.* Cell, 2001. **105**(2): p. 197-207.

42. Takizawa, P.A., et al., *Actin-dependent localization of an RNA encoding a cell-fate determinant in yeast.* Nature, 1997. **389**(6646): p. 90-3.

43. Zhang, J., et al., *The role of maternal VegT in establishing the primary germ layers in Xenopus embryos.* Cell, 1998. **94**(4): p. 515-24.

44. Du, J., et al., *Modulation of synaptic plasticity by antimanic agents: the role of AMPA glutamate receptor subunit 1 synaptic expression.* The Journal of

neuroscience : the official journal of the Society for Neuroscience, 2004. **24**(29): p. 6578-89.

45. Arimura, N. and K. Kaibuchi, *Neuronal polarity: from extracellular signals to intracellular mechanisms.* Nature reviews. Neuroscience, 2007. **8**(3): p. 194-205.

46. Barnes, A.P. and F. Polleux, *Establishment of axon-dendrite polarity in developing neurons.* Annual review of neuroscience, 2009. **32**: p. 347-81.

47. Bodian, D., *A Suggestive Relationship of Nerve Cell Rna with Specific Synaptic Sites.* Proc Natl Acad Sci U S A, 1965. **53**: p. 418-25.

48. Kacharmina, J.E., et al., *Stimulation of glutamate receptor protein synthesis and membrane insertion within isolated neuronal dendrites.* Proc Natl Acad Sci U S A, 2000. **97**(21): p. 11545-50.

49. Paradies, M.A. and O. Steward, *Multiple subcellular mRNA distribution patterns in neurons: a nonisotopic in situ hybridization analysis.* J Neurobiol, 1997. **33**(4): p. 473-93.

50. Poon, M.M., et al., *Identification of process-localized mRNAs from cultured rodent hippocampal neurons.* J Neurosci, 2006. **26**(51): p. 13390-9.

51. Aakalu, G., et al., *Dynamic visualization of local protein synthesis in hippocampal neurons.* Neuron, 2001. **30**(2): p. 489-502.

52. Crino, J.E.a.P., *Analysis of mRNA Populations from Single Live and Fixed Cells of the Central Nervous System.* Current Protocols in Neuroscience, 2001. **5**(5.3).

53. Eberwine, J., *Single-cell molecular biology.* Nature Neuroscience, 2001. **4**: p. 1155-1156.

54. Job, C. and J. Eberwine, *Localization and translation of mRNA in dendrites and axons.* Nat Rev Neurosci, 2001. **2**(12): p. 889-98.

55. Crino, P.B. and J. Eberwine, *Molecular characterization of the dendritic growth cone: regulated mRNA transport and local protein synthesis.* Neuron, 1996. **17**(6): p. 1173-87.

56. Schuman, E.M., J.L. Dynes, and O. Steward, *Synaptic regulation of translation of dendritic mRNAs.* The Journal of neuroscience : the official journal of the Society for Neuroscience, 2006. **26**(27): p. 7143-6.

57. Glanzer, J., et al., *RNA splicing capability of live neuronal dendrites.* Proc Natl Acad Sci U S A, 2005. **102**(46): p. 16859-64.

58. Crino, P., et al., *Presence and phosphorylation of transcription factors in developing dendrites.* Proc Natl Acad Sci U S A, 1998. **95**(5): p. 2313-8.

59. Miyashiro, K., M. Dichter, and J. Eberwine, *On the nature and differential distribution of mRNAs in hippocampal neurites: implications for neuronal functioning.* Proc Natl Acad Sci U S A, 1994. **91**(23): p. 10800-4.

60. Burgin, K.E., et al., *In situ hybridization histochemistry of Ca2+/calmodulin-dependent protein kinase in developing rat brain.* The Journal of neuroscience : the official journal of the Society for Neuroscience, 1990. **10**(6): p. 1788-98.

61. Racca, C., A. Gardiol, and A. Triller, *Dendritic and postsynaptic localizations of glycine receptor alpha subunit mRNAs.* The Journal of neuroscience : the official journal of the Society for Neuroscience, 1997. **17**(5): p. 1691-700.

62. Krichevsky, A.M. and K.S. Kosik, *Neuronal RNA granules: a link between RNA localization and stimulation-dependent translation.* Neuron, 2001. **32**(4): p. 683-96.

63. Rook, M.S., M. Lu, and K.S. Kosik, *CaMKIIalpha 3' untranslated region-directed mRNA translocation in living neurons: visualization by GFP linkage.* The Journal of neuroscience : the official journal of the Society for Neuroscience, 2000. **20**(17): p. 6385-93.

64. Ule, J. and R.B. Darnell, *RNA binding proteins and the regulation of neuronal synaptic plasticity.* Curr Opin Neurobiol, 2006. **16**(1): p. 102-10.

65. Wells, D.G., *RNA-binding proteins: a lesson in repression.* The Journal of neuroscience : the official journal of the Society for Neuroscience, 2006. **26**(27): p. 7135-8.

66. Kiebler, M.A. and G.J. Bassell, *Neuronal RNA granules: movers and makers.* Neuron, 2006. **51**(6): p. 685-90.

67. St Johnston, D., *Moving messages: the intracellular localization of mRNAs.* Nature reviews. Molecular cell biology, 2005. **6**(5): p. 363-75.

68. Kosik, K.S. and A.M. Krichevsky, *The message and the messenger: delivering RNA in neurons.* Science's STKE : signal transduction knowledge environment, 2002. **2002**(126): p. pe16.

69. Huang, Y.S., et al., *Facilitation of dendritic mRNA transport by CPEB.* Genes & development, 2003. **17**(5): p. 638-53.

70. Steward, O. and P.F. Worley, *Selective targeting of newly synthesized Arc mRNA to active synapses requires NMDA receptor activation.* Neuron, 2001. **30**(1): p. 227-40.

71. Tiruchinapalli, D.M., et al., *Activity-dependent trafficking and dynamic localization of zipcode binding protein 1 and beta-actin mRNA in dendrites and spines of hippocampal neurons.* The Journal of neuroscience : the official journal of the Society for Neuroscience, 2003. **23**(8): p. 3251-61.

72. Kobayashi, H., et al., *Identification of a cis-acting element required for dendritic targeting of activity-regulated cytoskeleton-associated protein mRNA.* The European journal of neuroscience, 2005. **22**(12): p. 2977-84.

73. Eom, T., et al., *Localization of a beta-actin messenger ribonucleoprotein complex with zipcode-binding protein modulates the density of dendritic filopodia and filopodial synapses.* The Journal of neuroscience : the official journal of the Society for Neuroscience, 2003. **23**(32): p. 10433-44.

74. Muslimov, I.A., et al., *RNA transport in dendrites: a cis-acting targeting element is contained within neuronal BC1 RNA.* The Journal of neuroscience : the official journal of the Society for Neuroscience, 1997. **17**(12): p. 4722-33.

75. An, J.J., et al., *Distinct role of long 3' UTR BDNF mRNA in spine morphology and synaptic plasticity in hippocampal neurons.* Cell, 2008. **134**(1): p. 175-87.

76. Mori, Y., et al., *Two cis-acting elements in the 3' untranslated region of alpha-CaMKII regulate its dendritic targeting.* Nature Neuroscience, 2000. **3**(11): p. 1079-84.

77. Blichenberg, A., et al., *Identification of a cis-acting dendritic targeting element in the mRNA encoding the alpha subunit of Ca2+/calmodulin-dependent protein kinase II.* The European journal of neuroscience, 2001. **13**(10): p. 1881-8.

78. Kremerskothen, J., et al., *Postsynaptic recruitment of Dendrin depends on both dendritic mRNA transport and synaptic anchoring.* Journal of neurochemistry, 2006. **96**(6): p. 1659-66.

79. Iijima, T., et al., *Hzf protein regulates dendritic localization and BDNF-induced translation of type 1 inositol 1,4,5-trisphosphate receptor mRNA.* Proc Natl Acad Sci U S A, 2005. **102**(47): p. 17190-5.

80. Ainger, K., et al., *Transport and localization elements in myelin basic protein mRNA.* The Journal of cell biology, 1997. **138**(5): p. 1077-87.

81. Hoek, K.S., et al., *hnRNP A2 selectively binds the cytoplasmic transport sequence of myelin basic protein mRNA.* Biochemistry, 1998. **37**(19): p. 7021-9.

82. Blichenberg, A., et al., *Identification of a cis-acting dendritic targeting element in MAP2 mRNAs.* The Journal of neuroscience : the official journal of the Society for Neuroscience, 1999. **19**(20): p. 8818-29.

83. Brechbiel, J.L. and E.R. Gavis, *Spatial regulation of nanos is required for its function in dendrite morphogenesis.* Current biology : CB, 2008. **18**(10): p. 745-50.

84. Pal, R., et al., *Selective dendrite-targeting of mRNAs of NR1 splice variants without exon 5: identification of a cis-acting sequence and isolation of sequence-binding proteins.* Brain Res Mol Brain Res, 2003. **994**(1): p. 1-18.

85. Muslimov, I.A., et al., *Dendritic transport and localization of protein kinase Mzeta mRNA: implications for molecular memory consolidation.* The Journal of biological chemistry, 2004. **279**(50): p. 52613-22.

86. Wu, K.Y., et al., *Local translation of RhoA regulates growth cone collapse.* Nature, 2005. **436**(7053): p. 1020-4.

87. Bockers, T.M., et al., *Differential expression and dendritic transcript localization of Shank family members: identification of a dendritic targeting element in the 3' untranslated region of Shank1 mRNA.* Molecular and cellular neurosciences, 2004. **26**(1): p. 182-90.

88. Liu, J., et al., *Two mRNA-binding proteins regulate the distribution of syntaxin mRNA in Aplysia sensory neurons.* The Journal of neuroscience : the official journal of the Society for Neuroscience, 2006. **26**(19): p. 5204-14.

89. Aranda-Abreu, G.E., et al., *Embryonic lethal abnormal vision-like RNA-binding proteins regulate neurite outgrowth and tau expression in PC12 cells.* The Journal of neuroscience : the official journal of the Society for Neuroscience, 1999. **19**(16): p. 6907-17.

90. Behar, L., et al., *cis-acting signals and trans-acting proteins are involved in tau mRNA targeting into neurites of differentiating neuronal cells.* International journal of developmental neuroscience : the official journal of the International Society for Developmental Neuroscience, 1995. **13**(2): p. 113-27.

91. Prakash, N., et al., *Dendritic localization of rat vasopressin mRNA: ultrastructural analysis and mapping of targeting elements.* The European journal of neuroscience, 1997. **9**(3): p. 523-32.

92. Yudin, D., et al., *Localized regulation of axonal RanGTPase controls retrograde injury signaling in peripheral nerve.* Neuron, 2008. **59**(2): p. 241-52.

93. de Moor, C.H., H. Meijer, and S. Lissenden, *Mechanisms of translational control by the 3' UTR in development and differentiation.* Seminars in cell & developmental biology, 2005. **16**(1): p. 49-58.

94. Sandberg, R., et al., *Proliferating cells express mRNAs with shortened 3' untranslated regions and fewer microRNA target sites.* Science, 2008. **320**(5883): p. 1643-7.

95. Irier, H.A., et al., *Control of glutamate receptor 2 (GluR2) translational initiation by its alternative 3' untranslated regions.* Molecular pharmacology, 2009. **76**(6): p. 1145-9.

96. Irier, H.A., et al., *Translational regulation of GluR2 mRNAs in rat hippocampus by alternative 3' untranslated regions.* Journal of neurochemistry, 2009. **109**(2): p. 584-94.

97. Bell, T.J., et al., *Cytoplasmic BK(Ca) channel intron-containing mRNAs contribute to the intrinsic excitability of hippocampal neurons.* Proc Natl Acad Sci U S A, 2008. **105**(6): p. 1901-6.

98. Milner, R.J. and J.G. Sutcliffe, *Gene expression in rat brain.* Nucleic Acids Res, 1983. **11**(16): p. 5497-520.

99. Milner, R.J., et al., *Brain-specific genes have identifier sequences in their introns.* Proc Natl Acad Sci U S A, 1984. **81**(3): p. 713-7.

100. Buckley, P.T., et al., *Cytoplasmic intron sequence-retaining transcripts can be dendritically targeted via ID element retrotransposons.* Neuron, 2011. **69**(5): p. 877-84.

Chapter 2

DENDRITIC TRANSCRIPTOMES OF RATS AND MICE EXHIBIT ELEVATED LEVEL OF DIVERGENCE

2.1 ABSTRACT

Using mechanically isolated dendrites from primary cultures of hippocampal neurons of two mouse strains (C57BL/6 and Balb/c) and one rat strain (Sprague-Dawley), we assayed the evolutionary differences in subcellular

dendritic localization of mRNAs. We found significantly greater evolutionary diversification of RNA localization in the dendritic transcriptomes (81% difference among highly expressed genes) compared to the transcriptomes of 11 different CNS and non-CNS tissues (average of 44%). Differentially localized genes include many genes involved in CNS function. These results suggest the existence of species-specific RNA localization mechanisms, which is consistent with our recent finding demonstrating a rapidly evolving retroviral element functions in rat-specific RNA localization. We speculate that the differences in the localized RNA may mediate activity-dependent functional differences in neurons leading to rapid diversification of brain function through modulation of individual cell function.

2.2 INTRODUCTION

Brain evolution is characterized by changes in size, structural complexity and connectivity of the central nervous system (CNS), commonly referred to as mosaic evolution [1-3]. Recently, with the affluence of genomic studies, gene expression was shown to be tightly connected to the evolution of phenotypes [4-6]. Particularly, changes in gene expression in the human lineage, have been identified in brain evolution [7, 8]. In fact any transcriptional dis-regulation would affect both development and physiology of the CNS [9, 10]. Neurons are highly polarized cells interconnected by key functional units known as synapses therefore, in addition to changes in neuro-anatomy, brain evolution should also involve modulation of the synaptic compartments [11]. In this context, studies of the last past two decades revealed the importance of subcellular mRNA localization and local translation in mediating synaptic function [12, 13]. Perturbation of these events can have serious effects at the cellular and organismal level, leading to neurological diseases such as Fragile X Syndrome and Spinal Muscular Atrophy [14]. Subcellular localization of mRNA is mediated by posttranscriptional regulatory factors, and is generally important for all cell function, but little is known about evolution of subcellular localization.

In a recent study from our lab, Buckley *et al.* [15] identified a novel mechanism for dendritic localization of mRNA mediated by a family of SINE retrotransposons called ID elements, contained within cytoplasmic intron

sequence-retaining transcripts (CIRTs). ID elements are hypothesized to be derived from the BC1 RNA gene [16] and expanded in the rodent genome. The BC1 RNA gene contains a 5' stem-loop structure that has been identified as containing localization signals [17]. Many of the dendritic genes with ID elements within CIRTs preserve the spatial localization signals of the BC1 gene and functional studies showed that these ID elements play a role in dendritic localization of rat neurons [15].

The genomic copy numbers of ID elements are highly variable in the rodent lineage, with estimates ranging from ~300 copies in guinea pigs to over 150,000 copies in rats [18]. As these copy-numbers were mainly based on gel electrophoresis studies, their exactitude may not be very reliable but the general order of magnitude in ID elements difference between rodents is undeniable. We re-annoted ID element numbers in both mouse and rat genome based on computational sequence analysis detailed in the Materials and Methods section. And found that mice seem to carry at least 50% less ID elements than the ones estimated in rats. It is estimated that the mouse-rat evolutionary split originated ~8-10 million years ago [19], suggesting a relatively rapid evolution of ID element insertions in these two species. Given the role of the ID elements in dendritic localization in rats [15] and the much lower number of these elements in mouse compared to rat, we hypothesized that the mouse dendritic transcriptome may be more divergent from the rat dendritic transcriptome than previously thought.

Here, we characterize the dendritic transcriptome of mice and rats' hippocampal neurons to assess whether a significant difference in subcellular

localization in the neurons of these two closely related species exists. We chose neurons from the hippocampus, as this brain region is known to be involved in learning and memory. Using microarrays on a collection of mechanically isolated dendrites from mouse and rat neurons, we find a high degree of evolutionary divergence in the dendritic transcriptome of these two species (81% difference among highly expressed genes), significantly greater than the divergence seen in other organs or whole brain tissues (average of 44%). Also, a functional category analysis revealed that many genes, previously described to have functional roles in synaptic plasticity and neurodegenerative threat showed significantly different level of expression in the dendrites of rats and mice.

We propose that the neuronal architectures of even closely related mammalian species might show substantial evolutionary diversification at the subcellular level. This suggests that learning abilities and behavioral differences between species may involve not only neuro-anatomical differences, but also subcellular differences in synaptic compartments and molecular functions of individual neurons.

2.3 RESULTS

2.3.1 Microarray analysis reveals a high degree of divergence between mouse and rat dendritic transcriptomes.

To assess neuronal expression divergence between mice and rats, we used the Affymetrix array platform to assay the transcriptomes of mechanically dissected individual dendrites of hippocampal neurons in dispersed primary cell cultures from Sprague-Dawley rat (9 biological replicates), C57/BL6 mouse (14 biological replicates), and Balb/c mouse (5 biological replicates). The detailed procedure of samples preparation is provided in the material and methods section and the different steps in the collection of dendrites are illustrated in Figure 2.1

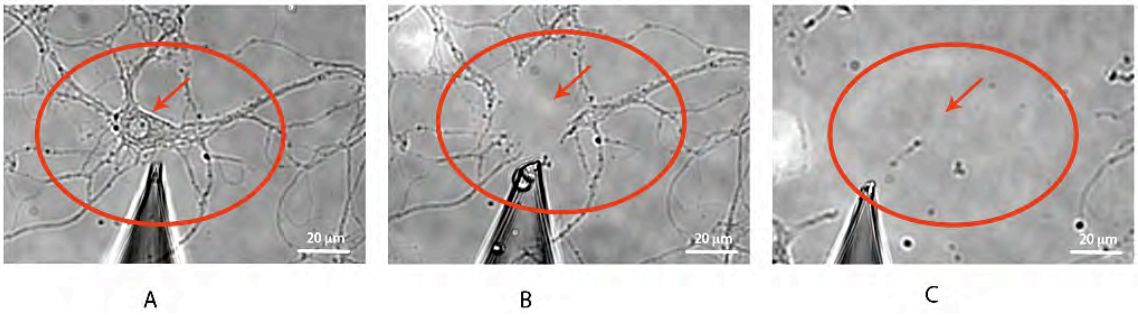


Figure 2.1: Mechanical severing of dendrites from neurons.

Rat hippocampal neuron with its soma (red arrow) and dendrites (red circle) before (A) and after aspiration by a glass micropipette of the soma (B) and dendrites (C).

All replicate samples from the two mouse strains and the rat strain show good concordance with average pair-wise Pearson's correlation of 0.80. In

addition, we also validated our RNA amplification protocol with a series of synthetic dilution and replicate amplification experiments and obtained an average correlation of 0.74 across all of our amplification controls (see Materials and Methods).

To compare expressed genes between rat and mouse, we constructed a stringent BLAST reciprocal-best-hit homology map (see Materials and Methods), yielding 10,833 conservatively mapped mouse-rat orthologs (Table 2.S1 in the appendices, provided as a separate document due to its large size). Using a t-test on normalized array expression values with a stringent FDR of 0.1%, we found 4713 out of 10,833 genes with significantly different dendritic localization between rat and the C57BL/6 mouse. In contrast, a within-species comparison between the C57BL/6 and Balb/c mouse strains yielded only 54 significantly different genes (FDR 0.1%). At the soma level, none of the genes exhibited significant difference in its expression neither between rat and mouse nor within the two mouse strains (FDR 0.1%). We should note that for the soma samples, the lack of statistically significant difference is due to large single cell variability within each species rather than a lack of difference between the two species. In Chapter 4, I discuss our control experiments for single cell variation that suggests the observed biological variability is inherent to individual cells. Therefore, in the results below we do not compare the expressed levels of RNA for the soma units, as the single cell variability precludes consistent estimates of between species divergence.

In order to compare the two dendritic transcriptomes more conservatively using highly expressed genes, we computed the median rank of the expression levels across the biological replicates of the ortholog-mappable genes for each species, and then assessed the overlap in gene identity of the top 5% of the highly expressed set. At this broad level, a surprisingly small fraction, ~19% (105) genes, are shared between the top 5% expressed genes in mouse and rat. The same comparison between C57BL/6 and Balb/c mice yields an overlap of 58% (312 genes), showing that the expression divergence is a function of evolutionary distance of the strains and species.

2.3.2 *In situ* experiments show inter-species differences in dendritic localization

We selected three sets of genes that show differing dendritic localization in rats and mice and carried out *in situ* assays of the spatial expression patterns on cultured rat and C57BL/6 mouse neurons, as shown in Figure 2.2. Each set of genes included three genes, which brings for each species, a total of 9 different *in situ* hybridizations. The identity of the genes investigated and their corresponding probe's sequence are detailed in the materials and methods section. The images and their *in situ* signals quantification via manual tracing using custom imaging software (Figure 2.3) verified subcellular localization differences in these three sets (See Materials and Methods). For example, *SFRS16* showed high signal in the cell soma of both the mouse and the rat

neurons, but high dendritic signal only in the mouse neuron (Figures 2.2*A and 2.3A*). In contrast, *ZFP410* showed high dendritic signal in rat but not mouse (Figures 2.2*B and 2.3B*), and *OLFM1* showed consistently high dendritic signal in both mouse and rat (Figures 2.2C *and 2.3C*).

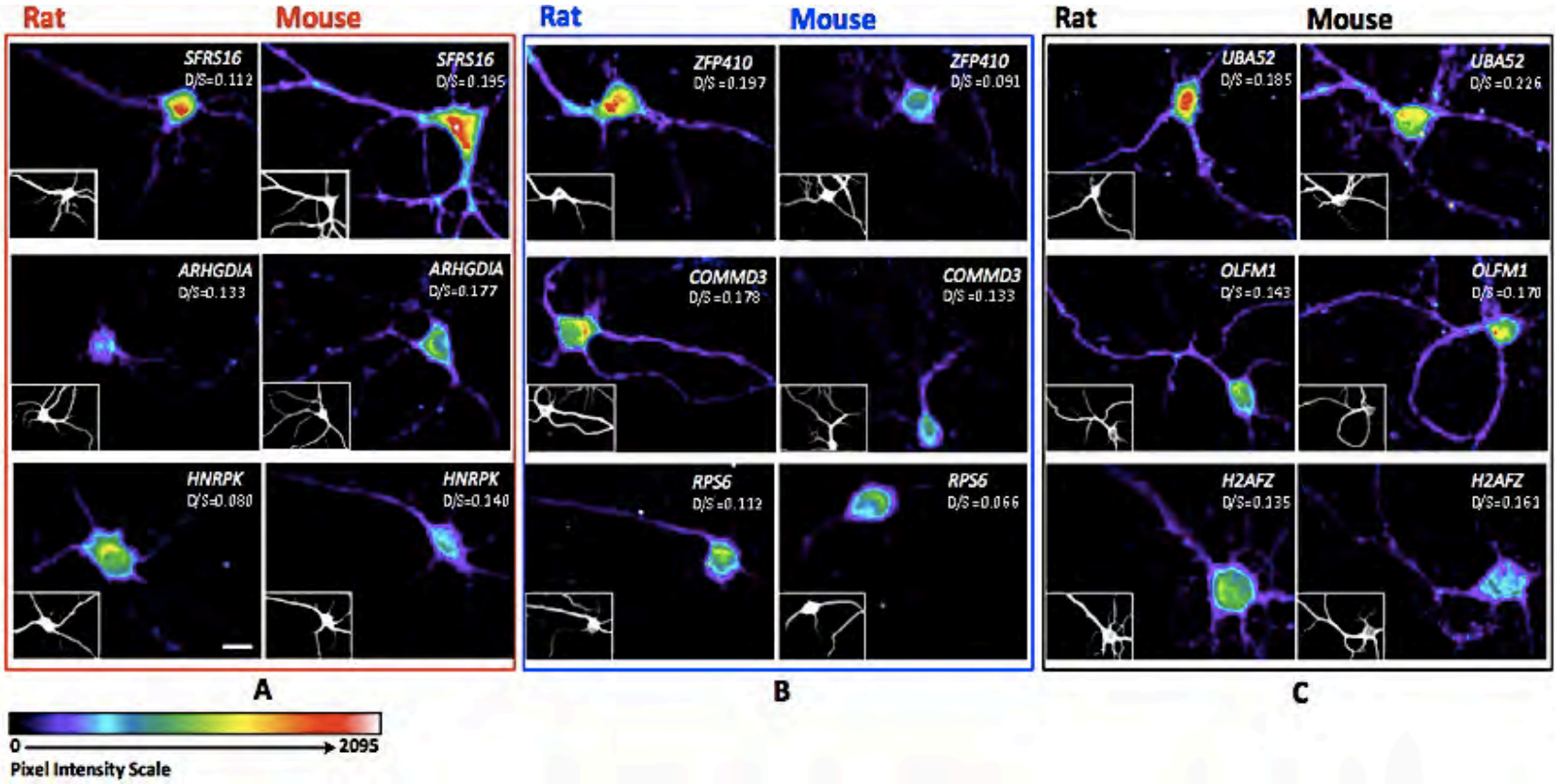


Figure 2.2: *In situ* hybridization reveals species-specific patterns of localization in neuronal dendrites.

Fluorescent Microscopy evaluation of biotin-conjugated oligoprobes on paraformaldehyde fixed 14-day cultured rat and mouse cortical neurons hybridized with 9 biotin-conjugated oligoprobes detected with streptadivin-Alexa568. For each probe images set, the small bottom left corner panels represent MAP2 immuno-staining. Scale bar = 20µm.
(A), Probes against *SFRS16, ARHGDIA* and *HNRPK* transcripts show higher dendritic localization in mouse neurons than in rat neurons (Red box).
(B), Probes against *ZFP410, COMMD3* and *RSP6* transcripts show higher dendritic localization in rat neurons than in mouse neurons (Blue box).
(C), Probes against *UBA52, OLFM1* and *H2AFZ* transcripts show high dendritic localization in both rat and mouse neurons (Black box).
D/S = Ratio of average pixel intensity in dendrites versus soma (computed via IGOR Pro 6.04 image software)

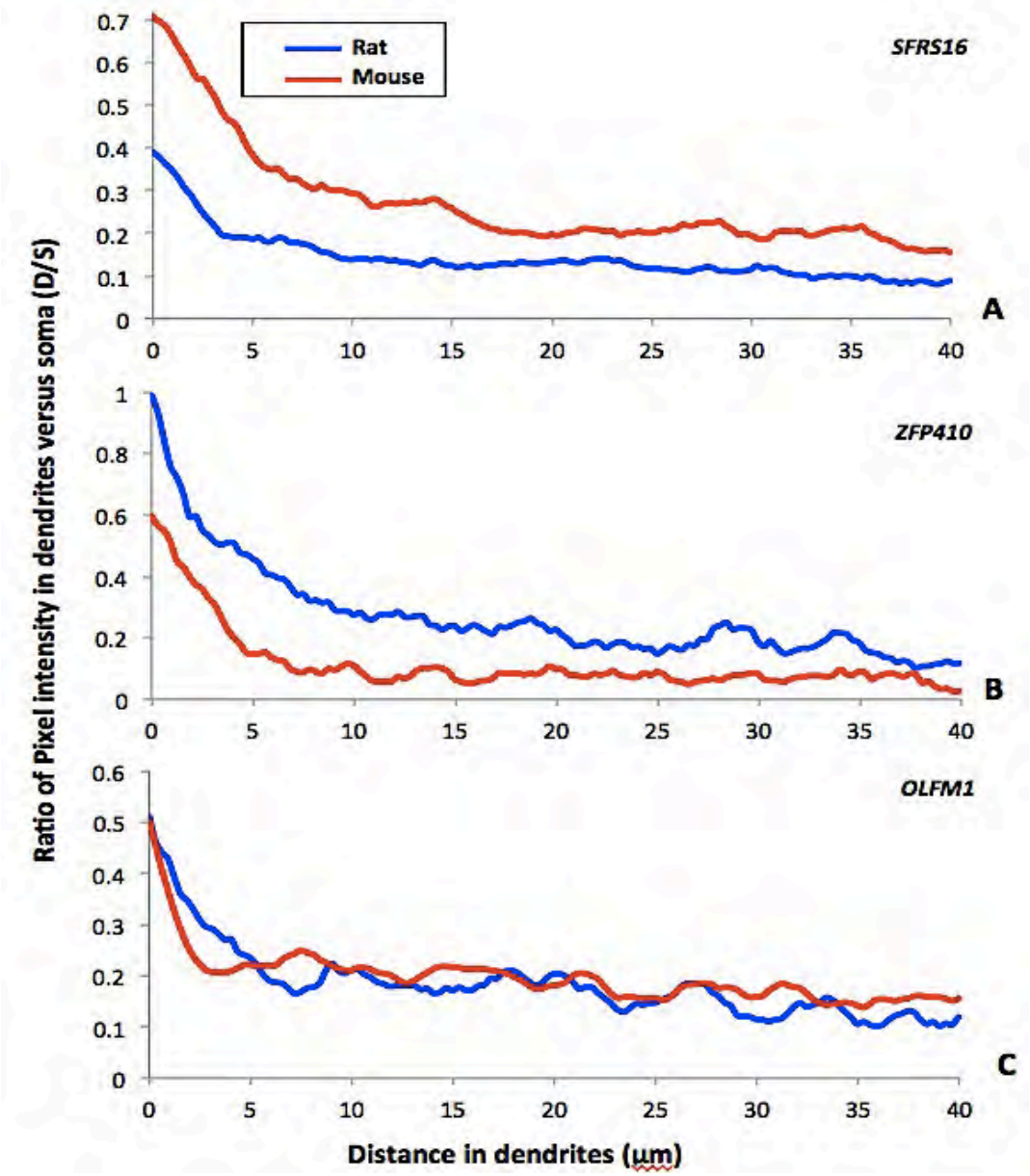


Figure 2.3: Quantified *In situ* hybridization signal shows species-specific localization in dendrites.

Graphs represent the ratio of *in situ* signal in dendrites versus soma (D/S) against the distance (from soma toward the dendrites).
(A), The probe against *SFRS16* transcript shows higher dendritic localization in mouse neurons than in rat neurons.
(B), The probe against *ZFP410* transcript shows higher dendritic localization in rat neurons than in mouse neurons.
(C), The Probe against *OLFM1* transcript shows similar level of dendritic localization in both rat and mouse neurons.

2.3.3 Dendritic transcriptomes are more divergent than other tissues.

To place the above dendritic transcriptome comparisons in context, we analyzed the expression data for 11 different organs/tissues of the Sprague-Dawley rat and C57BL/6 mouse available from the Genomics Institute of the Novartis Research Foundation (GNF) [20]. We computed the overlap percentage of the top 5% expressed genes for each of the 11 different tissue arrays between the two species (the total number of ortholog-mappable genes here is 3839 due to array version difference). Figure 2.4 shows a heatmap of the overlap percentages within each species across the tissues (Figs. 2.4A and 2.4B) as well as between species across the tissues (Fig. 2.4C). The last row and column of each of the heatmap shows the overlap percentage of the dendritic transcriptome. The diagonal elements in Figure 2.4C show the overlap percentages of homologous tissues across the two species and the off-diagonal elements show the overlap percentages of the non-homologous tissues.

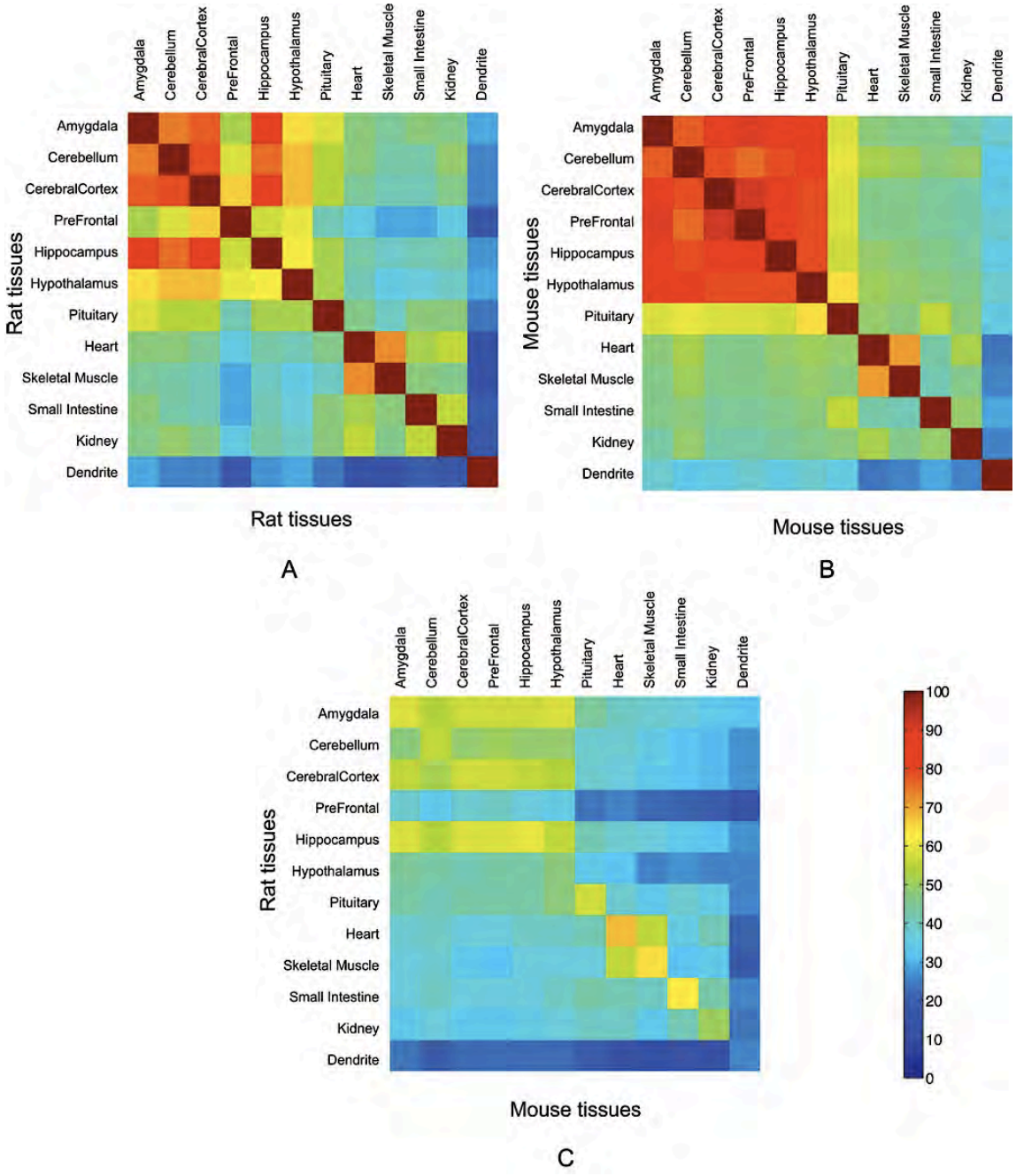


Figure 2.4: Heatmap of overlap percentages for the top 5% expressed genes.

The diagonal elements show the overlap percentages of homologous tissues and the off-diagonal elements show the overlap percentages of the non-homologous tissues. The last row and column of each of the heatmap shows the overlap percentage of the dendritic transcriptome with each tissue and transcriptome.
A) Overlap between tissues for rat.
B) Overlap between tissues for mouse.
C) Overlap between tissues across species.

As expected, tissues from brain anatomical regions show more similar gene expression compared to tissues from other organs both within and across the species. However, the dendritic transcriptome shows greater divergence in both rat and mouse than any other tissue -- even greater than non-homologous tissue comparisons. The fraction of overlap between the top 5% expressed ortholog-mappable genes of the mouse and rat dendritic transcriptomes was significantly different from the fraction of overlap of the homologous tissues (arcsine transformed t-test, $p < 10^{-7}$). Of special note in Figure 2.4 are the patterns related to the rat prefrontal cortex and hypothalamus. Both of these tissues showed higher divergence within rat non-homologous tissue comparisons as well as between homologous rat-mouse comparisons. The GNF dataset indicated that the prefrontal cortex was from a 20-week old rat while the hypothalamus was from a 16-week old rat. All other tissues in both rat and mouse were collected from 10-week old animals. Despite the developmental timing disparities between these samples, the overlap percentages were 39.6% and 47.4%, respectively for the prefrontal cortex and the hypothalamus, which was still significantly higher than for the overlap between dendritic transcriptomes. Our dendritic samples were extracted from developmentally matched time points (see Materials and Methods). Homologous developmental points can be difficult to define but our dendritic comparisons showed significantly greater divergence than the tissue data, which ranges between 10-20 week animals; thus, the dendritic divergence cannot be explained by mismatched developmental samples.

While mechanisms of transcriptional regulation can account for differences in gene expression between the two species, we believe that posttranscriptional regulatory mechanisms contribute to the dendritic divergence. This is supported by the fact that almost 85% of the genes in the top 5% overlap fraction of the rat and mouse hippocampus tissue, were not present in the top 5% overlap fraction for the dendritic transcriptomes.

To compare the overlap in highly expressed genes at other ranks than Top 5%, we also computed the number of common genes at each k rank (for k > 30) for all homologous tissue and dendritic transcriptomes for the rat and C57BL/6 mouse. Figure 2.5 shows the percent overlap in gene identity between the two species as a function of rat expression rank k up to 500 for the average of all GNF tissues and our dendritic transcriptome. For each curve we also computed 95% Bonferroni corrected confidence interval as well as the min and max of the tissue overlap percentages. The figure shows that the dendritic transcriptome was significantly different in overlap percentages compared to tissue comparisons at all ranks (until the curves converge to random at very large ranks, not shown in this figure).

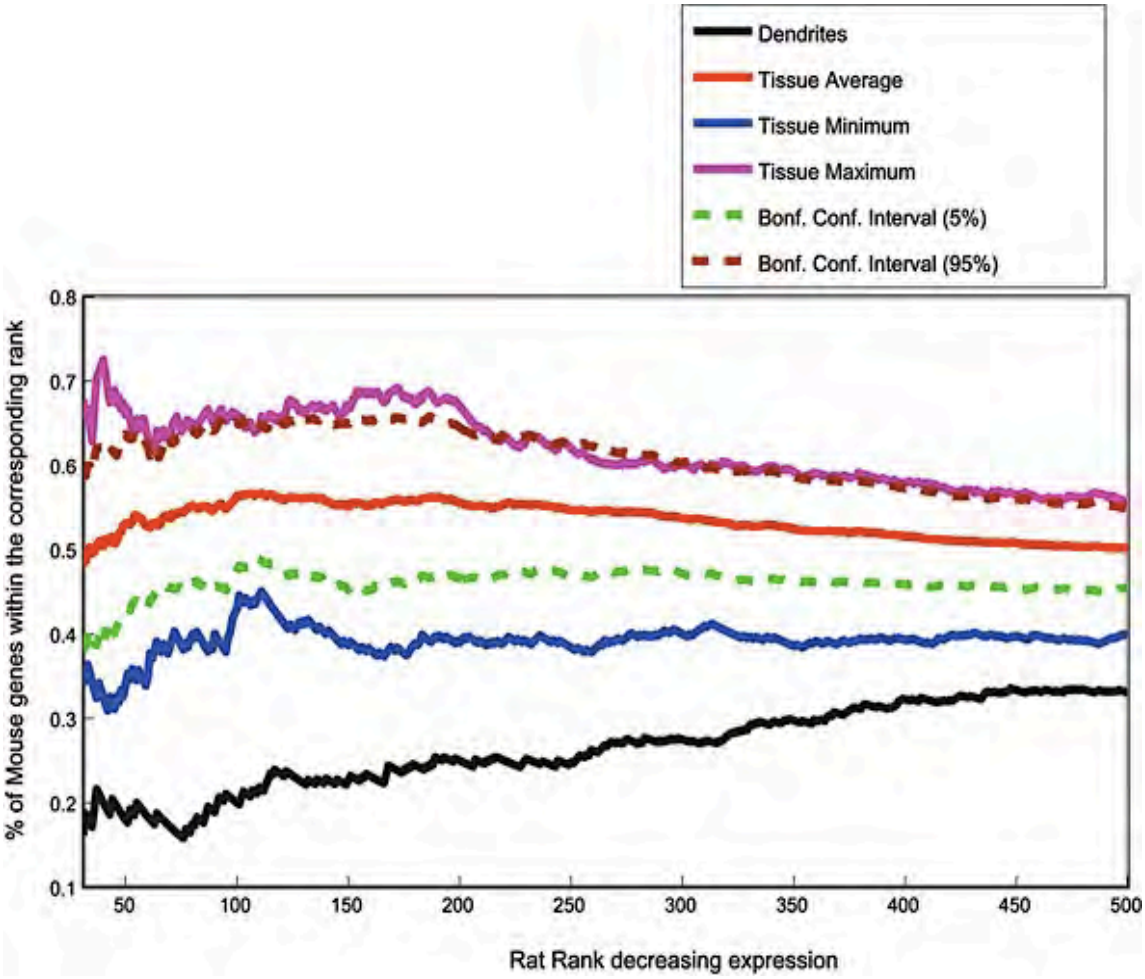


Figure 2.5: Rank concordance map between rat and mouse for dendrites and tissues gene expression.

Curves show fraction of mouse genes (y-axis) that have ranks lesser than or equal to rank (for rat) represented on the x-axis. The black thick curve shows trend for the dendrites. The red line shows average trend across 11 tissues. The green and brown dotted lines show the lower and upper Bonferroni corrected confidence intervals for the average trend (red line). The blue and the pink lines show the trends for the minimum and maximum values (across all tissues) for the tissue trend data. The rank on the x-axis ranges from rank 31 to rank 500 for the rat expression data.

2.3.4 Functional annotation of divergently localized dendritic genes.

The analysis above showed that the dendritic transcriptome of rats and mice are significantly more divergent than other homologous tissues. As described earlier, dendritic transcripts and their localized translation are critical determinants of synaptic plasticity, which is necessary for brain functions such as learning and memory. Understanding the evolution of higher brain function in terms of dendritic physiology will not only help understand the neurochemical basis of behavior but also identify therapeutic targets for neurological diseases linked to loss of memory and motor function (e.g. Alzheimer and Parkinson).

As a first step toward understanding species-specific effectors of dendritic physiology, we questioned, for the top 2000 expressed dendritic genes within each species, which broad functional categories are significantly represented. A GO analysis of these genes, performed using DAVID [21] and GOEAST [22] (see Materials and Methods for more details), highlighted categories such as *localization, neurogenesis, and ribosomal components* as enriched in both rat and mouse (Table 2.1A-B and Figure 2.6). However, even within shared GO categories between these two species, there is expression divergence within gene families, suggesting species-specific functionalization. For instance within the RAB family, which is involved in vesicular trafficking and neurotransmitter release, RAB3 and RAB10 are mainly present in the top 2000 mouse dendritic transcripts while RAB1, 8, 15 and 21 are mainly present in the top2000 rat dendritic transcripts.

Table 2.1A: GO analysis for the top 2000 mouse dendritic expressed genes

GO Term	GO Description	FDR
GO:0009987	Cellular process	4.5E-26
GO:0044237	Cellular metabolic process	4.8E-10
GO:0006412	Translation	1.1E-06
GO:0051179	Localization	5.3E-06
GO:0006091	Generation of precursor metabolites and energy	1.9E-05
GO:0007399	Nervous system development	1.7E-04
GO:0022900	Electron transport chain	8.5E-04
GO:0006810	Transport	1.1E-03
GO:0019538	Protein metabolic process	2.1E-02
GO:0006119	Oxidative phosphorylation	3.7E-02
GO:0022008	Neurogenesis	4.0E-02
GO:0005840	Ribosome	5.3E-09
GO:0030529	Ribonucleoprotein complex	1.8E-07
GO:0043227	Membrane-bounded organelle	1.8E-07
GO:0043228	Non-membrane-bounded organelle	2.5E-07
GO:0043232	Intracellular non-membrane-bounded organelle	2.5E-07
GO:0070469	Respiratory chain	7.7E-07
GO:0031966	Mitochondrial membrane	8.3E-07
GO:0005740	Mitochondrial envelope	4.4E-06
GO:0019866	Organelle inner membrane	6.2E-06
GO:0005743	Mitochondrial inner membrane	6.3E-06
GO:0031090	Organelle membrane	2.6E-05
GO:0031967	Organelle envelope	2.6E-05
GO:0044429	Mitochondrial part	3.1E-05
GO:0005856	Cytoskeleton	4.3E-04
GO:0005198	Structural molecule activity	1.5E-07
GO:0005515	Protein binding	3.8E-06
GO:0015077	Monovalent inorganic cation transmembrane transporter activity	5.2E-02
GO:0003954	NADH dehydrogenase activity	7.4E-02
GO:0008137	NADH dehydrogenase (ubiquinone) activity	7.4E-02
GO:0050136	NADH dehydrogenase (quinone) activity	7.4E-02
GO:0015078	Hydrogen ion transmembrane transporter activity	7.8E-02

Table 2.1B: GO analysis for the top 2000 rat dendritic expressed genes

GO Term	GO Description	FDR
GO:0009987	Cellular process	5.84E-31
GO:0006414	Translational elongation	1.13E-04
GO:0007399	Nervous system development	1.76E-04
GO:0034621	Cellular macromolecular complex subunit organization	2.50E-04
GO:0022008	Neurogenesis	1.22E-03
GO:0034622	Cellular macromolecular complex assembly	3.06E-03
GO:0051179	Localization	1.12E-08
GO:0006810	Transport	2.09E-07
GO:0001568	Blood vessel development	3.56E-03
GO:0048514	Blood vessel morphogenesis	6.73E-03
GO:0043234	Protein complex	4.38E-06
GO:0031090	Organelle membrane	1.59E-05
GO:0043005	Neuron projection	8.96E-05
GO:0005740	Mitochondrial envelope	1.19E-04
GO:0031966	Mitochondrial membrane	1.28E-04
GO:0044445	Cytosolic part	1.83E-04
GO:0022626	Cytosolic ribosome	4.25E-04
GO:0022627	Cytosolic small ribosomal subunit	8.11E-04
GO:0044429	Mitochondrial part	2.95E-03
GO:0015935	Small ribosomal subunit	5.33E-03
GO:0019866	Organelle inner membrane	1.68E-02
GO:0070469	Respiratory chain	2.09E-02
GO:0005743	Mitochondrial inner membrane	3.53E-02
GO:0043228	Non-membrane-bounded organelle	4.89E-02
GO:0043232	Intracellular non-membrane-bounded organelle	4.89E-02
GO:0005515	Protein binding	9.71E-25
GO:0005516	Calmodulin binding	2.09E-02
GO:0015078	Hydrogen ion transmembrane transporter activity	3.35E-02
GO:0015077	Monovalent inorganic cation transmembrane transporter activity	7.52E-02

Table 2.1A-B: GO analysis for the top 2000 mouse and rat dendritic expressed genes.

GO analysis for the top2000 ranked dendritic genes in mouse (Table 2.2A) and rat (Table 2.2B) with FDR < 0.1 used as threshold value.
"Blue color" = Biological Process GO Category
"Gray color" = Cellular Component GO Category
"Purple color" = Molecular Function GO Category
"**Bold**"= GO Category common in rat and mouse

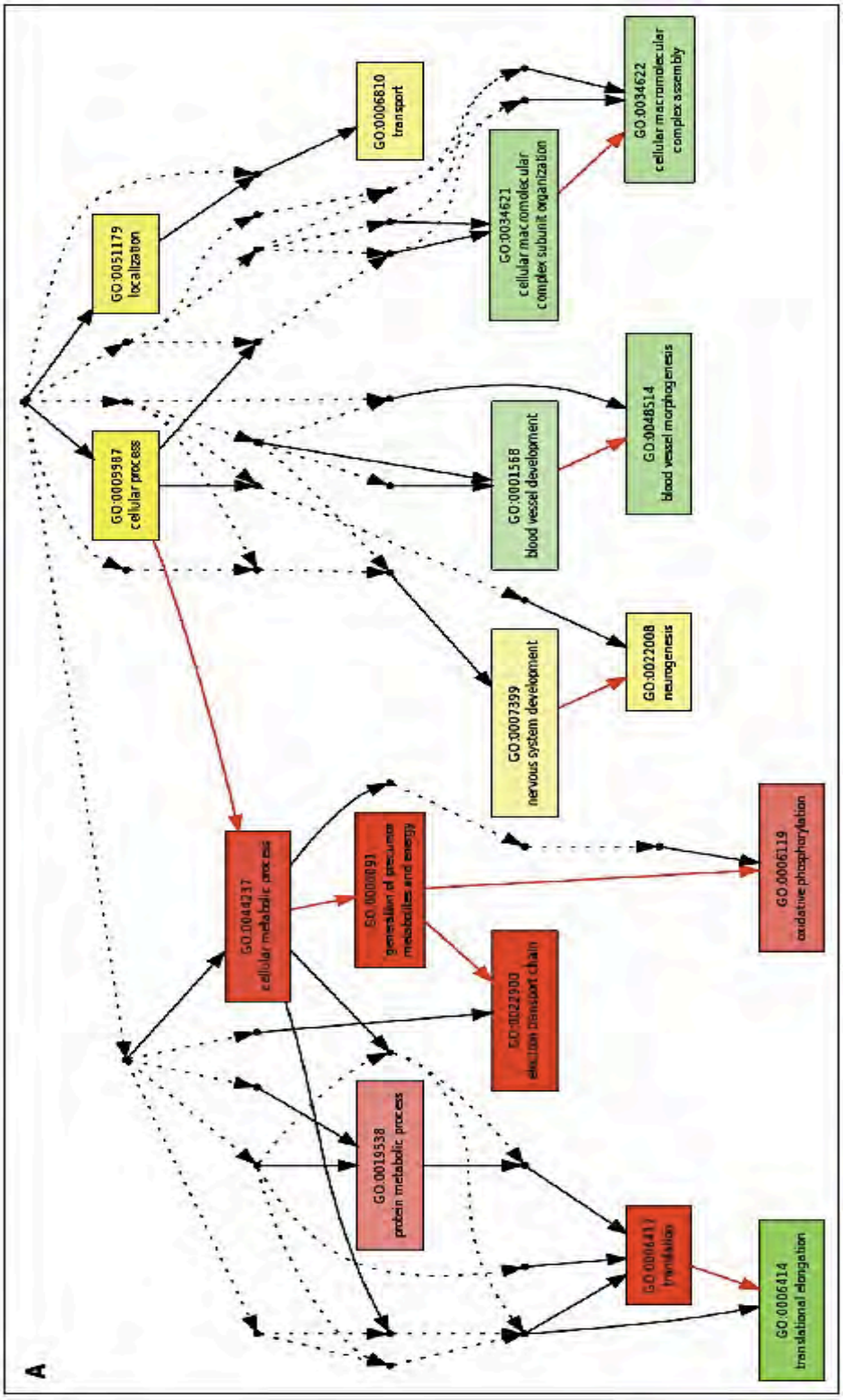


Figure 2.6A: "Biological Process" as GO category

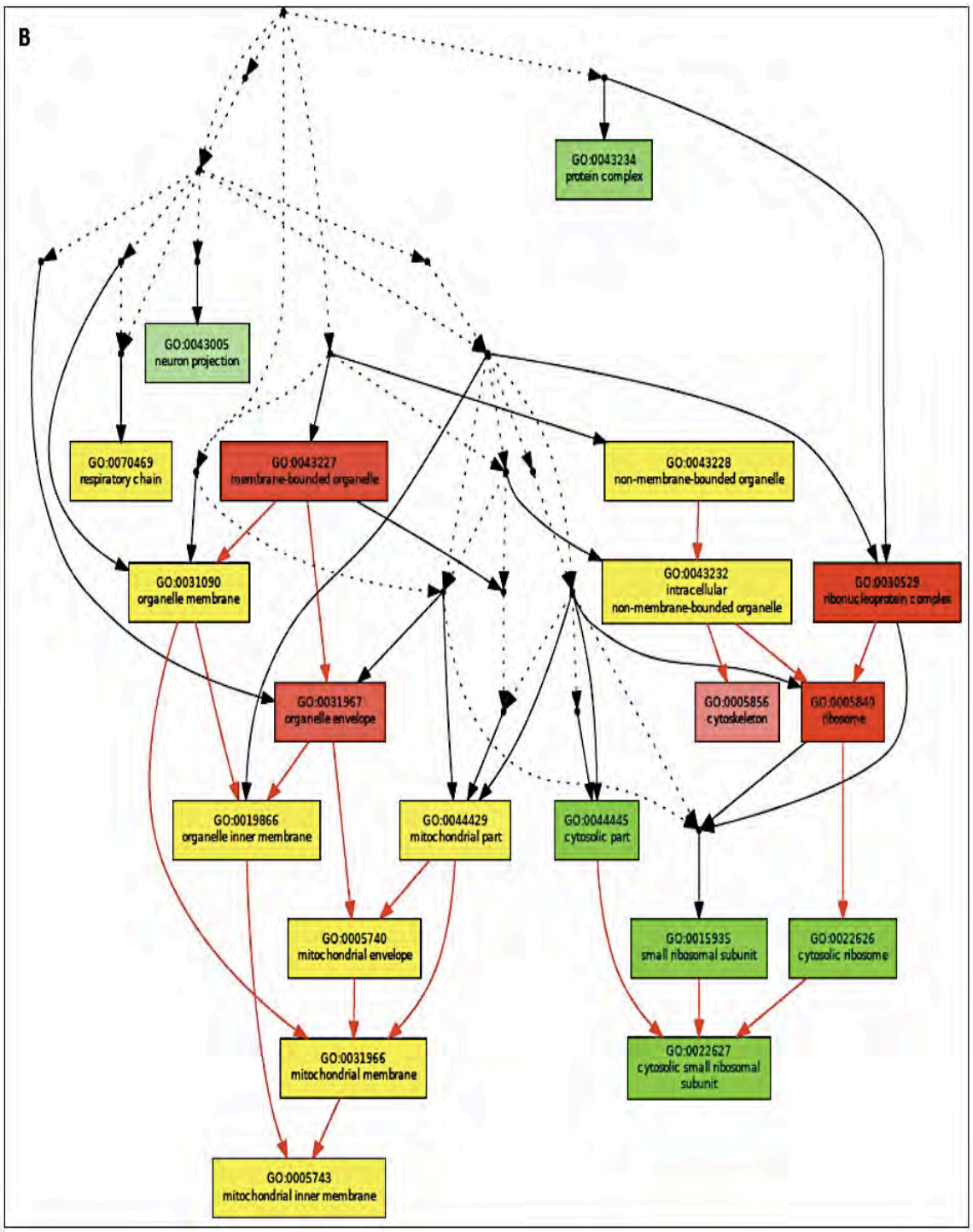


Figure 2.6B: "Cellular Component" as GO category

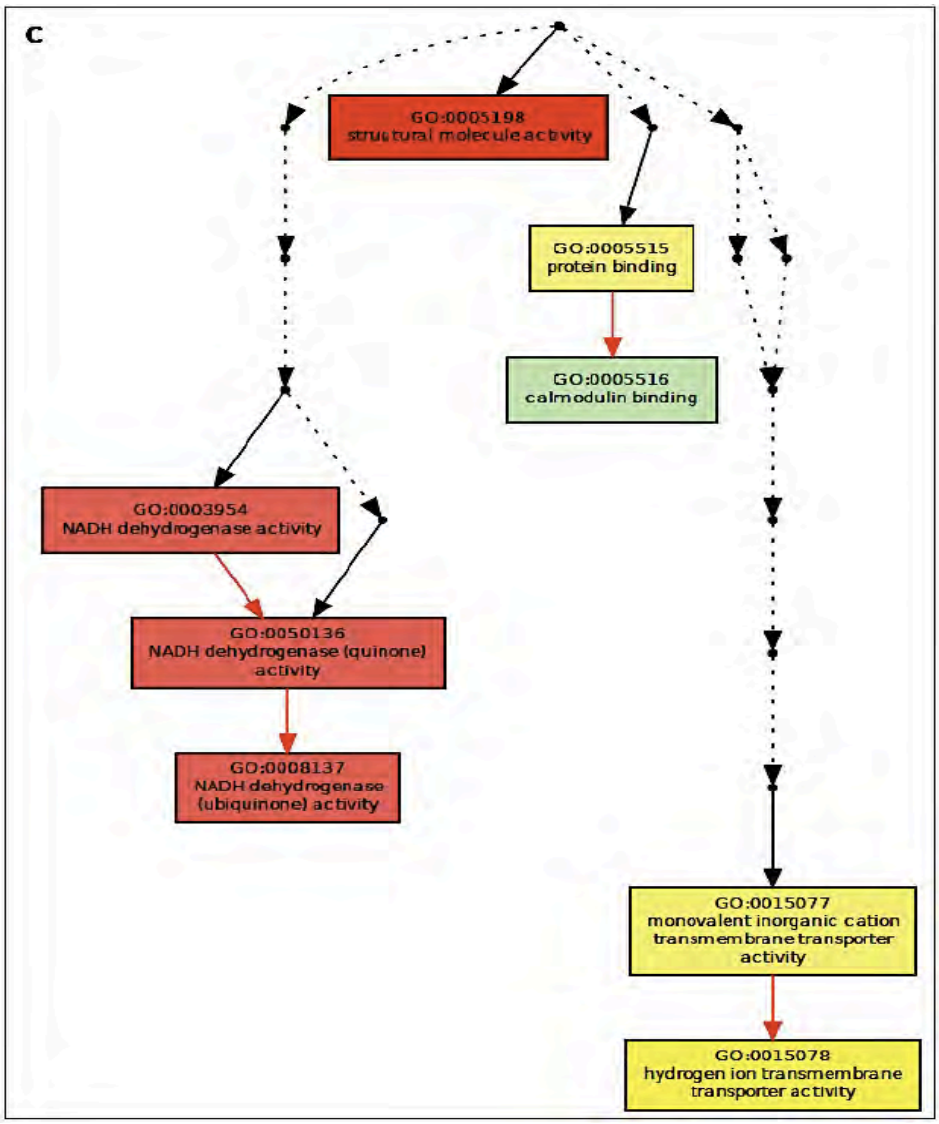


Figure 2.6C: "Molecular Function" as GO category

Figure 2.6 A-C: Graphs for the results of the GO analysis done on the top2000 ranked dendritic genes in rat and mouse.

These graphs display enriched GO IDs and their hierarchical relationships in "biological process" (A), "cellular component" (B) or "molecular function" (C) GO categories. Significantly enriched GO terms are marked in green, red or yellow if represented in rat, mouse, or both species respectively. The degree of color saturation of each node is positively correlated with the significance of enrichment of the corresponding GO term. Non-significant GO terms within the hierarchical tree are drawn as points. Branches of the GO hierarchical tree without significant enriched GO terms are not shown. Edges stand for connections between different GO terms. Red edges stand for relationship between two enriched GO terms, black solid edges stand for relationship between enriched and un-enriched terms, black dashed edges stand for relationship between two un-enriched GO terms (Performed via GOEAST, [22]).

We examined other genes and categories known to be important in synaptic function. Of particular note is the GO category *Mitochondria*, which is significantly enriched in the dendritic transcriptome of both species, highlighting its relevance in neuronal function. Mitochondria is involved in calcium sequestration, thus contributing to modulation of local calcium concentration - a factor important in Long-term potentiation (LTP) and long-term depression (LTD). We also see species-specific mechanisms for calcium regulation. For example, several calcium-sensitive genes show gene expression divergence (t-test FDR <0.0001, Table 2.S1 in the appendices, provided as a separate document due to its large size), including synaptotagmins and *CAMK2,* which can modulate calcium microenvironment in the dendrite [23]. Among the potassium channel genes (*KCNX*) genes relevant for neuronal excitability, almost half of them show significant difference (t-test FDR <0.0001, Table 2.S1 in the appendices, provided as a separate document due to its large size) between the two species. In addition, potassium channel auxiliary subunits, Beta1 and Beta2, that regulate potassium channels using different mechanisms [24] are also differentially expressed in dendrites, with Beta1 being higher in rat and Beta2 being higher in mouse. It has been proposed that the beta subunits can function as oxidoreductases that can link the redox state of the dendrites to the electrical activity of the cell [24]. The *Netrin* receptor *DCC*, which has been implicated in spatial control of translation [25] and in modulation of synaptic plasticity, also shows significant difference in dendritic expression. Other genes that show divergent expression include a G-coupled protein receptor (*GPR37*) and the ion

channel *KCNN2*. *GPR37* has been identified to interact with *Parkin*, a protein associated with Parkinson's disease, [26, 27], thus GPR37 could be of therapeutic relevance. *KCNN2* encodes the SK2 subunit of small-conductance calcium activated potassium channels and mutations in its 5' end has been associated with abnormal after hyperpolarization (AHP) activity and Parkinson's like phenotype in mice [28]. A more detailed listing of functionally important genes is provided in Table 2.2. This list consists of genes that are either in the top 5% for all Sprague-Dawley rat, C57BL/6 mouse and Balb/c mouse or are highly variable across the three samples (in the top 5% for at least one and in the bottom 50% for another). Table 2.S2 (in the appendices provided as separate document due to its large size) shows the complete list of receptors and synaptic genes categorized by a finer scale of rank expression.

Table 2.2: Synaptic plasticity genes show divergent level of expression in rats and mice dendrites.

Family	Class	SynapticFunction	GeneSymbol	Rat	Balb/c	C57BL/6
Channel	Voltage-gated	--	*Trpm1*	+	-	-
Channel	Voltage-gated	ARGs, LTP	*Cnga2~**	+	-	-
Channel	Voltage-gated	--	*Cacna1g*	+	-	-
Channel	Ligand-gated	--	*Chrna1*	+	+	+
Channel	Voltage-gated	LTP	*Hcn1**	+	+	+
Channel	Voltage-gated	--	*Kcnn2*	-	+	+
Gprotein	Gprotein	--	*Gng11*	+	+	+
Receptor	GPCR_A	LTP	*Htr1f**	+	-	-
Receptor	GPCR_A	--	*Ghsr*	+	-	-
Receptor	GPCR_A	--	*P2ry6*	+	-	-
Receptor	GPCR	--	*Gpr108*	+	-	-
Receptor	GPCR_A	--	*Npffr2*	+	+	+
Receptor	GPCR_C	LTP	*Grm8**	-	+	+
Receptor	GPCR_A	--	*Mchr1*	-	+	+
Receptor	GPCR_A	--	*Gpr61*	-	+	+
Receptor	GPCR	--	*Gprc5a*	-	+	+
Receptor	Receptor	--	*Ssr2*	+	-	-
Receptor	Receptor	--	*Ifngr1*	-	+	+
Receptor	Receptor	--	*Grb2*	-	+	+
Receptor	Receptor	--	*Agtr1a*	-	+	+
Receptor	Receptor	--	*Adipor2*	-	+	+
Other	--	LTP	*Art5**	+	-	-
Other	--	ARGs, LTP	*Calm3~**	+	-	-
Other	--	ARGs	*Crybb2~*	+	-	-
Other	--	ARGs	*Prx~*	+	-	-
Other	--	ARGs	*Fuca1~*	+	-	-
Other	--	ARGs	*Cx3cl1~*	+	-	-
Other	--	ARGs	*Rt1.aa~*	+	-	-
Other	--	ARGs	*Anxa8~*	+	-	-
Other	--	ARGs	*Sgcg~*	+	-	-
Other	--	ARGs	*Ggnbp1~*	+	-	-
Other	--	ARGs	*Pax8~*	+	-	-
Other	--	ARGs	*Nfya~*	+	-	-
Other	--	LTP	*Sod1**	+	-	+
Other	--	LTP, LTD	*Mapk3*#*	+	+	+
Other	--	LTP	*Stmn4**	+	+	+
Other	--	ARGs, LTP, LTD	*Nrgn~*#*	+	+	+
Other	--	ARGs	*Hyal2~*	+	+	+
Other	--	ARGs	*Anxa1~*	+	+	+
Other	--	ARGs	*Atp1b1~*	+	+	+
Other	--	ARGs	*Tapbp~*	+	+	+
Other	--	ARGs	*Rps29~*	+	+	+
Other	--	ARGs	*Ttc35~*	-	-	+
Other	--	ARGs	*Naca~*	-	-	+
Other	--	LTP	*Inhbc**	-	+	+
Other	--	LTP	*Ppp1r2**	-	+	+
Other	--	ARGs	*Aldh3a2~*	-	+	+
Other	--	ARGs	*Foxa2~*	-	+	+
Other	--	ARGs	*Tat~*	-	+	+
Other	--	ARGs	*H2afy~*	-	+	+
Other	--	ARGs	*Arhgdib~*	-	+	+
Other	--	ARGs	*Gdpd5~*	-	+	+
Other	--	ARGs	*Wdsub1~*	-	+	+
Other	--	ARGs	*Ect2~*	-	+	+
Other	--	ARGs	*Cfb~*	-	+	+
Other	--	ARGs	*Meox1~*	-	+	+
Other	--	ARGs	*Tppp3~*	-	+	+
Other	--	ARGs	*Vhl~*	-	+	+
Other	--	ARGs	*Spg7~*	-	+	+
Other	--	ARGs	*Dusp11~*	-	+	+

*Long Term Potentiation Genes (LTP); #Long Term Depression Genes (LTD); ~LTP activity regulated genes (ARGs); G protein Coupled Receptor (GPCR); GPCR group A (GPCR_A); GPCR group C (GPCR_C)
"+" Gene Expression ≥ Top5%; "-" Gene Expression ≤Top50%

These results demonstrate that many genes, previously described to have functional roles in synaptic plasticity and neurodegenerative risk showed significantly divergent localization in the dendrites of rats and mice. To obtain an overall view of the impact and extent of these differentially localized transcripts on the CNS, we examined their occurrence in three key pathways (critical for synaptic plasticity): Long-term potentiation (LTP), long-term depression (LTD) and Calcium signaling. Pathway information was downloaded from the KEGG database [29, 30]. Figures 2.7A, B and C illustrate that a large proportion of genes in these pathways display divergent expression between the two species supporting our hypothesis that neuronal function and its regulation involve common as well as species-specific mechanisms. The species-specific difference related to dendritic physiology can occur at multiple levels; from regulation of translational controlling mechanisms such as *DCC* receptors and ribosomal proteins, to fine-tuned controls of electrical characteristics through calcium sensing mechanisms and channel regulatory proteins. From this data we hypothesize that there is a functional divergence in the synaptic compartment of mouse and rat induced by RNA localization divergence.

Figure 2.7 A-C: Pathways highlighting genes with dendritic gene expression difference between rat and mouse

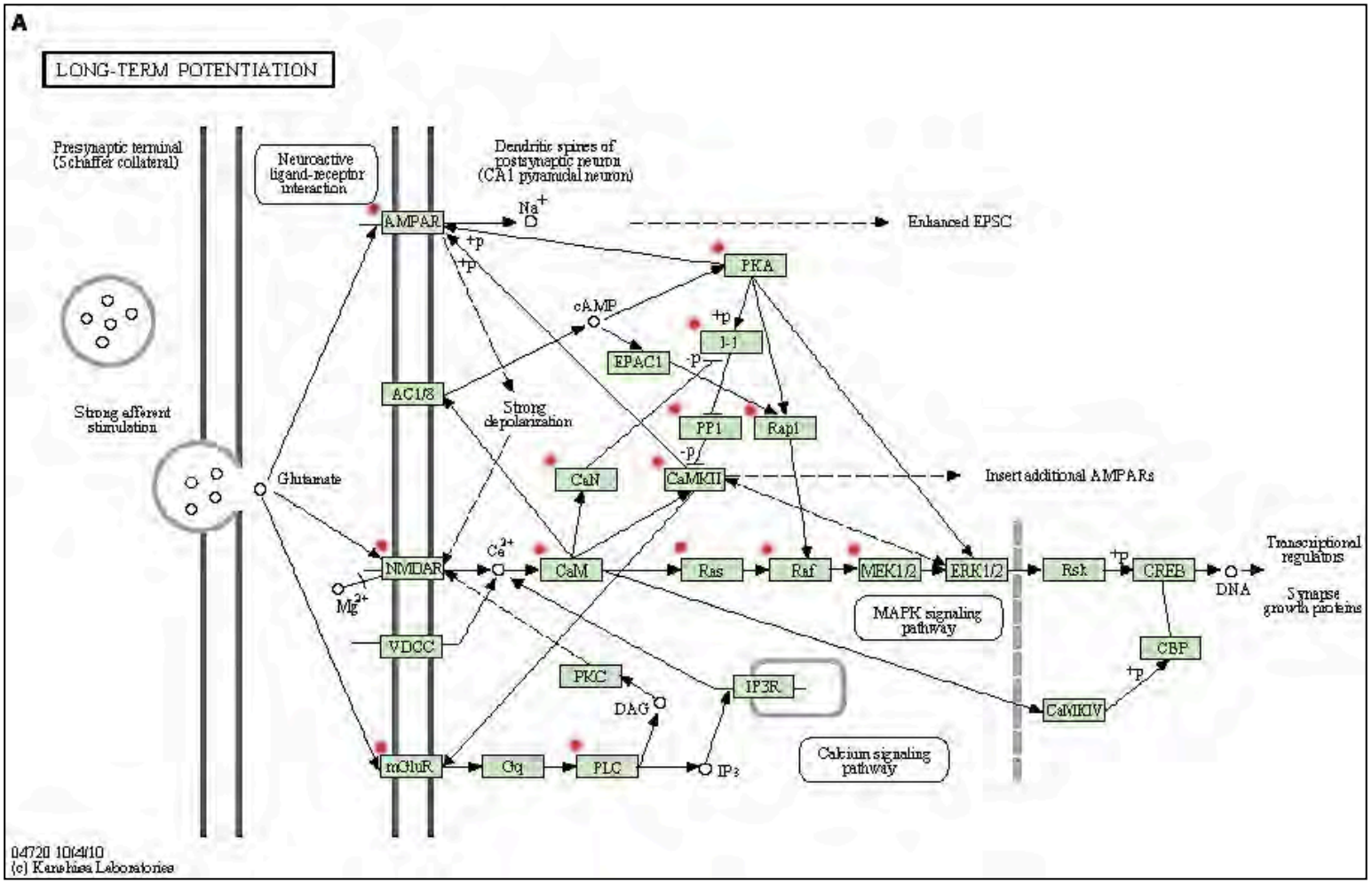


Figure 2.7A: Long-term potentiation pathway. The dendritically divergent genes (T-test significant, FDR <0.001) are highlighted in red

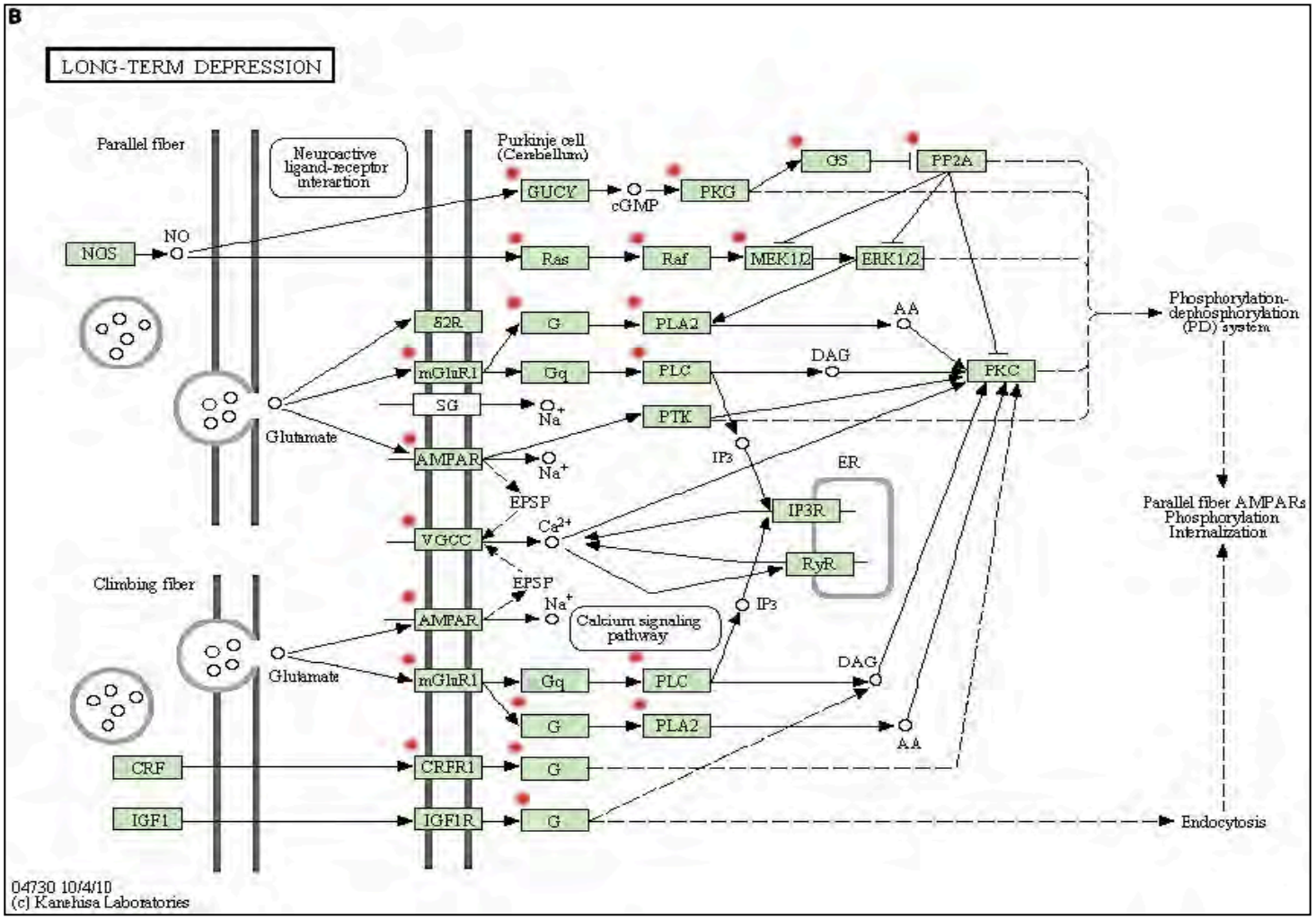


Figure 2.7B: Long-term depression pathway highlighting genes with dendritic gene expression difference between rat and mouse. The dendritically divergent genes (T-test significant, FDR <0.001) are highlighted in red

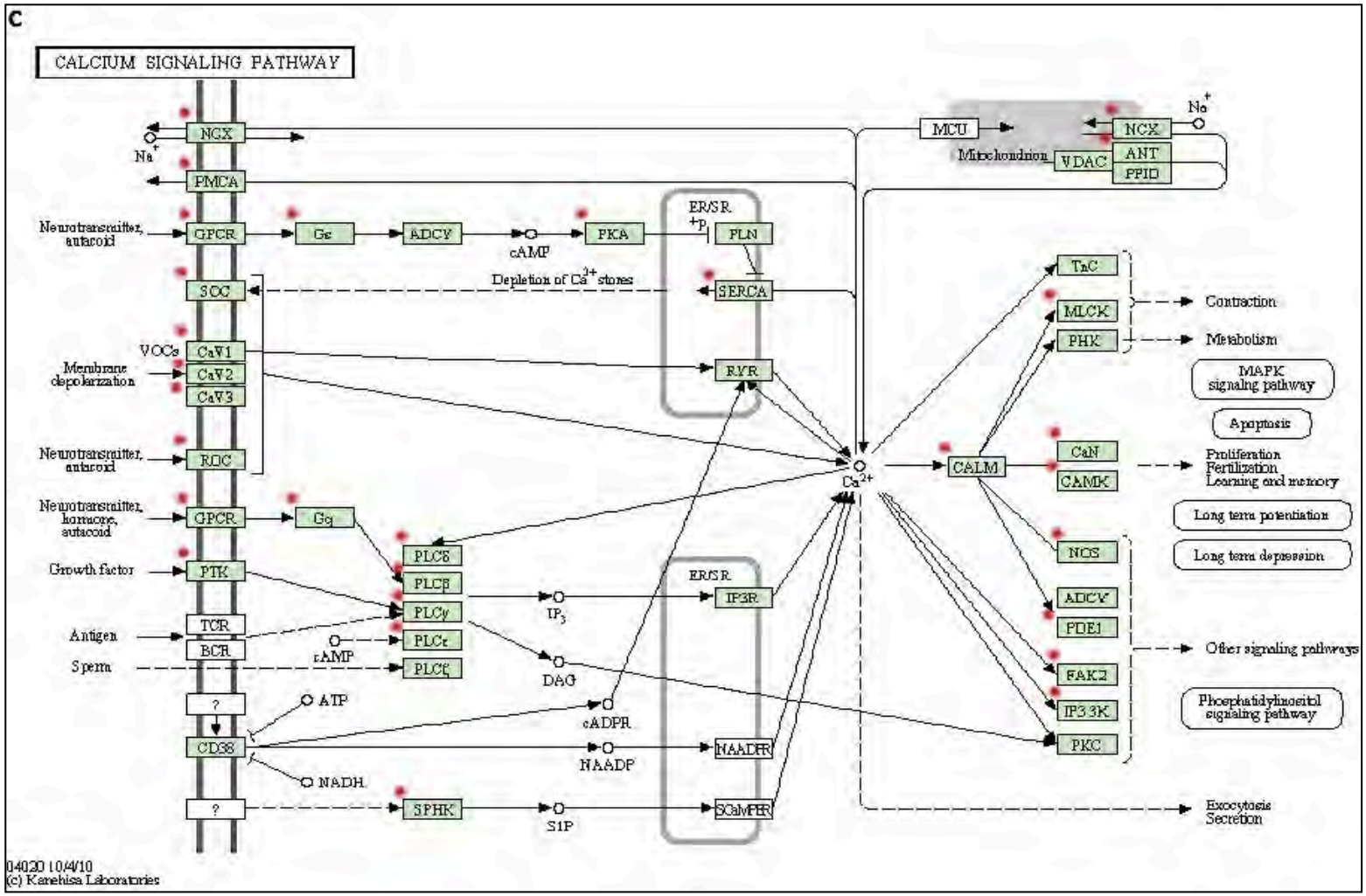


Figure 2.7C: Calcium signaling pathway highlighting genes with dendritic gene expression difference between rat and mouse. The dendritically divergent genes (T-test significant, FDR <0.001) are highlighted in red

2.4 DISCUSSION

In this study, we used mechanically isolated individual dendrite preparations to assay the dendritically localized transcriptomes from mouse and rat hippocampal neurons in a comparative genomics analysis. Our results show that the dendritic transcriptome is significantly more diverged in these two species than for other tissue- and organ-level transcriptomes. The level of divergence is not explained by experimental variability, as shown by our in vitro dilution and amplification control studies (see Materials and Methods), as well as comparisons between two mouse strains (C57BL/6J and Balb/c). Shadt *et al.* [31] compared the liver transcriptomes of two mouse strains (C57BL/6J and DBA/2J) using 111 F2 lines from the two strains and found 33.3% of the genes differentially expressed at a p-value of 0.05. When we used the same gene-wise p-value level instead of the family-wise FDR correction, we find 32.8% of the genes significantly differ between the dendritic transcriptome of C57BL/6J and Balb/c mouse strains, nearly identical to the Schadt *et al.* results [31]. Fernandes *et al.* [32] studied the hippocampal gene expression of eight different inbred strains of mice and found 252 genes significantly different out of 12,888 probes at a Bonferroni corrected p-value of 0.05. At the same Bonferroni corrected level, we find 45 genes significantly different for our 10833 genes between the dendritic transcriptome of Balb/c and C57BL/6 mice. This smaller number is consistent with the smaller degrees of freedom represented in our two-strain comparison versus the eight-strain comparison in the Fernandes *et al.* study. In conclusion,

these results suggest that our dendritic assays using micro-dissection and linear RNA amplification are consistent with whole tissue methods in terms of quantification accuracy. Furthermore, the rate of within-species divergence in dendritically localized expression that we see is consistent with tissue level assays in other studies [31, 32].

We reported that 43.5% of the orthologous genes between the rat and mouse dendritic transcriptomes showed significantly different expression at an FDR correction of 0.1%, and that the two transcriptomes shared only 19% identity among the top 5% expressed genes in each species. The lack of replicates in the GNF tissue level data precluded a t-test comparison, but the fraction of overlap among the top expressed mouse-rat orthologs was significantly lower for the dendritic transcriptomes compared to any other tissue including various brain tissues. Humans and orangutans are putatively separated by 13 million years of evolution, which is similar to the rat-mouse split. Hsieh *et al.* [33] reanalyzed the transcriptome data from primate species for brain and liver tissues and found that 52.3% and 0.8% of genes differ between brain expression at p-value of 0.05 and Bonferroni corrected (Bonf.) p-value of 0.05, respectively; they also report a difference of 31.3% (p = 0.05) and 1.3% (Bonf. p = 0.05) of the genes for liver expression. When we assessed our rat-mouse dendritic transcriptome at these p-values, we obtained 70.5% (p = 0.05) and 25.8% (Bonf. p = 0.05) of genes that were significantly different. Previous evidences suggested that human brain expression experienced rapid evolutionary diversification compared to other tissues [8, 34]. Although the mouse-rat pair was similar in

divergence to the human-orangutan pair, the dendritic transcriptome in the mouse-rat pair showed significantly greater divergence, especially in the class of large deviation genes (25.8% vs. 1.3%, respectively at Bonf. p = 0.05).

The molecular basis of divergent brain function has been previously studied at the level of individual genes. Previous reports on strain or species variation in molecular brain function include neuropeptides and their receptor structure and distribution [35] as well as the protein levels of *CAMK2*, *MAPK*, *CREB*, and *BDNF* [29, 36] and other genes involved in development [10]. Our results supported the hypothesis that functional characteristics of the postsynaptic compartment may have rapidly diversified in rats and mice through changes in the processes that regulate subcellular localization. In our study, nearly twice the numbers of highly expressed genes showed significant divergence in subcellular localization but not in tissue level expression. Subcellular localization is mediated by post-transcriptional regulation involving both *cis*-factors within the transcript and *trans*-factors interaction. We examined the 5' and 3' UTRs of the mRNAs for signatures of greater sequence divergence in the set of genes with divergent dendritic localization. At the level of whole UTR regions, we did not detect any significant differences in sequence evolution for the two classes of genes. This may be due to the localization signal being a much smaller subset of UTR regions or because the localization signal is embedded in other sequence features as we showed in Buckley *et al.* [15]. In fact, upon examination of the most divergent dendritic genes, we find that well-formed ID elements (i.e. having a secondary structure similar to BC1 [16, 17], see

Materials and Methods for more details) occur in significantly higher rates among the top 5% rat-specific dendritic transcripts (2.36 ID elements per 10,000 nts) as compared to transcripts that are dendritically highly expressed in both mouse and rat (1.82 per 10,000 nts, $p < 0.001$ by binomial test). Furthermore, amongst the 33 genes studied in Buckley *et al.* [15], 17 genes show dendritic localization in mouse and 9 out of these 17 genes have ID elements in non-homologous positions in their introns. Since the functional experiments performed by Buckley *et al.* [15] demonstrated the implication of ID elements in transcripts targeting to dendrites, these mouse genes with introns-retained ID elements are most likely to be associated with dendritic localization. Functional experiments in mouse are currently being undertaken in our lab to validate this observation. Further work is needed to determine the *cis*-elements involved in species-specific targeting for individual transcripts, but our results suggested the existence of species-specific mechanisms including those associated with transposon activity.

Although the functional significance of these large-scale dendritic transcriptome differences is not completely clear at this stage, our results point at the evolution of system level molecular physiology of dendrites rather than restricted to a small number of receptors or developmental differences in tissue organization. Rapid evolution of subcellular localization may be mediated by the fact that post-transcriptional regulation does not have epistatic effects with transcriptional regulation. Finally, our results also highlight that the choice of an animal model might require detailed species-specific knowledge of neuronal function.

2.5 MATERIALS AND METHODS

Sample collection for transcriptome analysis

Hippocampi primary cultures from mouse E18 (C57BL/6, Charles River Laboratories, Inc.) and rat E19 (Sprague-Dawley Charles River Laboratories, Inc.) were plated at 100,000 per ml in neurobasal medium (Invitrogen) with B-27 supplement (Sigma) on 12-mm round German Spiegelglas coverslips (Bellco Glass) and grown for 14 days [37]. Mouse and rat embryonic samples used for primary cultures were developmentally matched based on the protocol provided by Charles River Laboratories (http://www.criver.com/SiteCollectionDocuments/rm_rm_d_pregnant_rodent.pdf) These primary cultures allowed single-cell harvesting where a pool of 100-400 dendrites was mechanically isolated. At least 9 biological “dendrites-pool” replicates were collected in each species.

RNA Isolation and Microarrays

All samples were assessed through standard aRNA amplification methods, as described previously [12, 38]. After 2 rounds of amplification, a final aRNA amplification was performed with the Ambion Illumina TotalPrep RNA Amplification kit with an incubation time of 14 h. The integrity of these amplified aRNAs was evaluated with an Agilent Technologies 2100 Bioanalyzer and RNA Nano LabChip. 5µg of each aRNA was used for Affymetrix Rat 230 2.0 and Mouse 430.2 analysis.

In situ hybridization and imaging

Species-specific 25 DNA-oligomer biotin-labeled were custom-made (Sigma-Genosys®). 14 day-old primary rat and mouse cortical neurons were fixed for 15 minutes in 4% paraformaldehyde, washed in 1X PBS and permeabilized with 0.2% TritonX-100 for 10min at room temperature (RT). Cells were prehybridized at 36°C with 50% formamide, 1X Denhardt's solution, 4X SSC, 10mM DTT, 0.1% CHAPS, 0.1% Tween-20, 500µg/ml yeast tRNA, 500µg/ml salmon sperm DNA. *In situ* hybridization was performed for 16h at 36°C with 15ng/µl probe in prehybridization buffer. After probe hybridization, Rabbit anti-MAP2 (Microtubule Associated Protein 2) primary antibody (1:1000) was added to cells for 1h at RT followed by addition of secondary antibodies Alexa 488 goat anti-rabbit antibody (1:750) and Alexa 568 streptavidin conjugated (1:750) for 1h at RT. The co-staining for MAP2 was performed for two main reasons: First, MAP2 is known to be a marker for dendrites, second MAP2 is conserved in mammals and its expression is known to coincide with the maturation of neuronal morphology, and thus could be used as reference baseline for the maturity of both rat and mouse neurons fixed after 14 days in culture [39-41]. DAPI staining was performed before mounting the slides. The samples were visualized by fluorescent microscopy (Axiovert 200M Inverted Fluorescent Microscope – Zeiss Inc., 20x Objective). The collected images were processed in Metamorph® image analysis software. IGOR Pro 6.04 software (WaveMatrics, Inc.) was used to extract the pixel intensity information for the regions of interest and the extracted pixel

intensity values were plotted in Excel. For each transcript and in each species, a manual tracing was done on an average of 3 cell somas and 9 dendrites (The dendritic path origin started at the end of soma and went out along the dendritic process). The quantification paths were manually drawn based on MAP2 immuno-staining and automatically generated for the corresponding *in situ* hybridization channel. The ratio of the average pixel intensity along the paths in the dendrites (D) versus the soma (S) were computed and plotted against the distance from the dendrites path origin (40μm). Figure 2.3 illustrates 3 examples of the outcome plots.

Probe's Symbol	**Sequence (Biotin5'-3')**
ZFP410	GGACTGGGAATTCATAGACACCAGG
COMMD3	CGTCTGGTTTTCCTCTAGGCTCCTG
RPS6	TGCGCTTCCTCTCTCCAGTTCTCCT
SFRS16	AGAAACCCAGCAGCATAACAGCCCC
ARHGDIA	CGTGAACTTGGTCCCACGTTTGTCC
HNRPK	TCCACAGCATCAGATTCGAGCGGGA
UBA52	CGATGGAAGGGGACTTTATTTGGTC
RPL6	GCGATGACAAACTTCTGGTGTGTCC
H2AFZ	GTCCACTGGAATCACCAACACTGGA

Control Experiment

Mouse adult female brain's cortex (C57BL/6, Charles River Laboratories, Inc.) was isolated and stored immediately at -80°C. Subsequently, the mRNA (15μg) was isolated using TRIzol Reagent and MicroFastTrack 2.0 Kit (Invitrogen). A Sample of 5μg was assessed on Affymetrix Mouse 430.2 array. Aliquots from the

leftovers of the same cortical mRNA were diluted to single-cell RNA levels (0.1, 1, and 10 pg) and independently amplified, as described above, for a total of 2 and 4 rounds and assessed on Affymetrix Mouse 430.2 arrays

Computational Analysis of Single-Cell Transcriptome

Array quantification

The expression intensities of the probes were summarized using the upper decile statistic by using Affymetrix RMA 2.0 methods [42, 43]. All the arrays were median centered and scaled by the range of expression values between the 10^{th} and the 90^{th} percentile in each array.

Rat-Mouse Ortholog Map

Orthologs were identified using reciprocal-best-hits from a blast nucleotide (blastn) analysis between RefSeq version 37 for mouse and rat with an e-value threshold of 1e-5. We also carried out a blastn search of the Affymetrix probes against the respective sequence set for each species for probe sets for which the Affy mapping was not available or was outdated. We used the top hit from the probeset-mRNA blast search with the constraint of at least 24/25 base matches. The mapping was further restricted by stipulating that at least 9/11 probes, in each probe-set, should map to the same mRNA. By combining all these mapped relationships we constructed a mapping between the rat and the mouse probe-sets. This map includes both unique matches as well as many-to-many matches.

In order to resolve the many-to-many matches, we identified all connected components within and across both species (excluding the unique matches). The connected component was called a metagene and a unique identifier was assigned to it, and median values for each species connected component were used for the metagene. Figure 2.8 illustrates the workflow for creating the rat-mouse probe-set map.

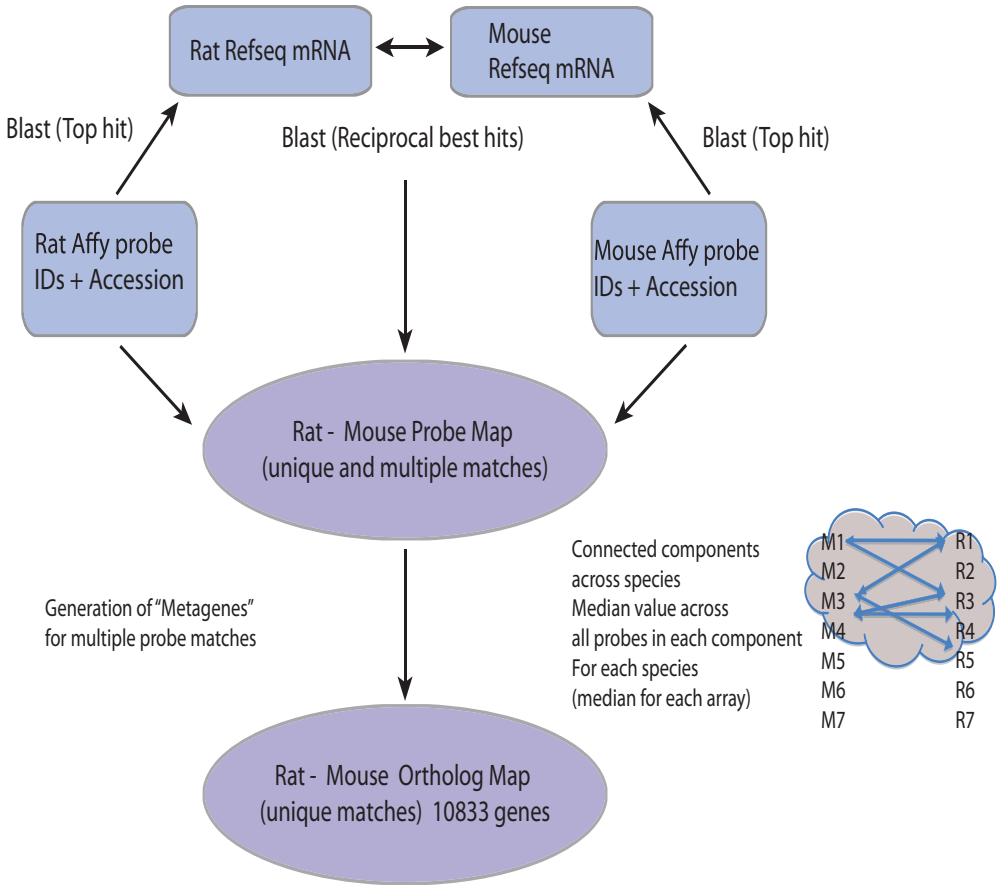


Figure 2.8: Workflow showing the construction of the Rat-Mouse Ortholog map.

The Blast results from the Probe-mRNA match are used only for cases where the Affymetrix accession numbers corresponding to probes were missing from the mRNA dataset.

Statistical tests

All statistical tests including Benjamini-Hotchberg FDR correction were carried out using custom programs and the Statistics Toolbox from MATLAB [44]. The p-value for the difference in the overlap of top 5% for dendrites vs. tissues were computed using an Arcsin transformation of percentages and an one-sample t-test.

GNF tissues data

Raw expression data for 11 tissues in mouse and rat was downloaded from the GNF BioGPS system [20]. The data was processed in the same manner as our dendritic data using the RMA algorithm and median centering and percentile range scaling. Since the Rat GNF expression analysis was carried out on a different platform, we used the best match probe mapping provided by GNF between their platform and the Affymetrix Platform resulting in a total of 3839 probe sets that mapped between rat and mouse.

Rankmap

We computed ranks with ties for rat and mouse expression data and each ortholog pair were sorted with respect to increasing rat rank (decreasing expression). The rankmap was made by computing for each rat rank k, the fraction of mouse genes that were equal to below rank k. We refer to this as the concordance level for the mouse. Note that the rank ordering of genes is specific to each tissue and dendrite. Thus, the concordance levels correspond to different

subsets of genes in each case. The rankmap confidence intervals were computed using the binomial distribution and applying a Bonferroni correction by a factor of 500 (for the total ranks compared).

Gene Ontology and Pathway analysis

GO analysis for the top2000 ranked dendritic genes in rat and mouse was carried out using the online resource – DAVID (Database for Annotation, Visualization and Integrated Discovery) [21]. A False Discovery Rate (FDR) < 0.1 was used as threshold value. The summary results of the above GO analysis were graphically displayed via GOEAST [22]. The graphs display enriched GO IDs and their hierarchical relationships in "biological process", "cellular component" or "molecular function" GO categories (Figure 2.6A-C)

Neuronal Functions Table

We combined four different resources to construct a table that highlights genes involved in synaptic plasticity, ion channels and receptors (Table 2.2 and Table 2.S2). We first extracted, from the Affymetrix (Rat 230 2.0 and Mouse 430.2) annotation files, all the genes that are described as “channels”, “G protein coupled receptors” or other “receptors”. Next, we extended the annotations by including available gene description from Park *et al*. [45], KEGG pathways [29, 30] and IUPHAR database [36].

ID elements mapping in mouse and rat (in collaboration with Mugdha Khaladkar, Ph.D.)

The RepeatMasker [46] annotations were downloaded from UCSC Genome Browser for mouse (mm9) and rat (rn4) genome assemblies [47]. The genomic coordinates for all SINE ID elements were extracted and analyzed for overlap with the Refseq annotated mRNAs in both sense (S) as well as anti-sense (AS) orientation. The presence of ID elements in the intronic, exonic or UTR regions was noted. The 3'UTR region for this purpose was extended 1000 bp downstream of the annotated 3'UTR end in order to account for longer un-annotated 3'UTRs. Overall, there were 33,406 ID elements in mouse (S: 15414, AS: 17992) and 61,311 ID elements in rat (S: 27396, AS: 33915) that mapped to the Refseq mRNAs.

ID elements enrichment analysis (in collaboration with Miler Lee, Ph.D.)

Well-formed ID elements were identified using the characteristics defined in Buckley *et al*. [15]. Genes occurring in the top 5% of rat dendrite expression were partitioned into two sets: genes whose mouse orthologs also occurred in the top 5% of mouse dendrite expression (n=103), and genes whose mouse orthologs occur in the bottom 50% of mouse dendrite expression (i.e., rat-specific dendritic genes) (n=148). The occurrence rate of well-formed ID elements per nucleotide of intronic sequence between these two sets was compared using a two-tailed binomial proportion test. A similar comparison between mouse-

specific dendritic genes and common rat/mouse dendritic genes did not yield a significant difference.

2.6 PUBLICATION NOTES

Data from this chapter appeared in part in: “Rapid divergence of RNA localization between rat and mouse neurons reveals the potential for rapid brain evolution”. In Preparation for submission. Chantal Francis*, Shreedhar Natarajan*, Miler T. Lee, Peter T. Buckley, Jai-Yoon Sul, James Eberwine[#] and Junhyong Kim[#].

*, [#] denotes joint authorship

REFERENCES

1. Redies, C. and L. Puelles, *Modularity in vertebrate brain development and evolution.* Bioessays, 2001. **23**(12): p. 1100-11.
2. Crick, F. and C. Koch, *A framework for consciousness.* Nat Neurosci, 2003. **6**(2): p. 119-26.
3. Karlen, S.J. and L. Krubitzer, *The evolution of the neocortex in mammals: intrinsic and extrinsic contributions to the cortical phenotype.* Novartis Found Symp, 2006. **270**: p. 146-59; discussion 159-69.
4. Wray, G.A., et al., *The evolution of transcriptional regulation in eukaryotes.* Mol Biol Evol, 2003. **20**(9): p. 1377-419.
5. Stern, D.L., *A role of Ultrabithorax in morphological differences between Drosophila species.* Nature, 1998. **396**(6710): p. 463-6.
6. Brunetti, C.R., et al., *The generation and diversification of butterfly eyespot color patterns.* Curr Biol, 2001. **11**(20): p. 1578-85.
7. Khaitovich, P., et al., *Evolution of primate gene expression.* Nature reviews. Genetics, 2006. **7**(9): p. 693-702.
8. Pfefferle, A.D., et al., *Comparative expression analysis of the phosphocreatine circuit in extant primates: Implications for human brain evolution.* Journal of human evolution, 2011. **60**(2): p. 205-12.
9. Gilbert, S.L., W.B. Dobyns, and B.T. Lahn, *Genetic links between brain development and brain evolution.* Nature reviews. Genetics, 2005. **6**(7): p. 581-90.
10. Rakic, P., *Evolution of the neocortex: a perspective from developmental biology.* Nature reviews. Neuroscience, 2009. **10**(10): p. 724-35.

11. Ryan, T.J. and S.G. Grant, *The origin and evolution of synapses.* Nature reviews. Neuroscience, 2009. **10**(10): p. 701-12.

12. Crino, J.E.a.P., *Analysis of mRNA Populations from Single Live and Fixed Cells of the Central Nervous System.* Current Protocols in Neuroscience, 2001. **5**(5.3).

13. Glanzer, J., et al., *RNA splicing capability of live neuronal dendrites.* Proc Natl Acad Sci U S A, 2005. **102**(46): p. 16859-64.

14. Bassell, G.J. and S. Kelic, *Binding proteins for mRNA localization and local translation, and their dysfunction in genetic neurological disease.* Curr Opin Neurobiol, 2004. **14**(5): p. 574-81.

15. Buckley, P.T., et al., *Cytoplasmic Intron Sequence-Retaining Transcripts Can Be Dendritically Targeted via ID Element Retrotransposons.* Neuron, 2011. **69**(5): p. 877-84.

16. Kim, J., et al., *Rodent BC1 RNA gene as a master gene for ID element amplification.* Proc Natl Acad Sci U S A, 1994. **91**(9): p. 3607-11.

17. Muslimov, I.A., et al., *Spatial codes in dendritic BC1 RNA.* J Cell Biol, 2006. **175**(3): p. 427-39.

18. Ono, T., et al., *Genomic organization and chromosomal distribution of rat ID elements.* Genes Genet Syst, 2001. **76**(4): p. 213-20.

19. Steppan, S., R. Adkins, and J. Anderson, *Phylogeny and divergence-date estimates of rapid radiations in muroid rodents based on multiple nuclear genes.* Systematic biology, 2004. **53**(4): p. 533-53.

20. Orozco, C., et al., *BioGPS: an extensible and customizable portal for querying and organizing gene annotation resources.* Genome Biology\, 2009. **10\(**11\): p. R130\.

21. Dennis, G., Jr., et al., *DAVID: Database for Annotation, Visualization, and Integrated Discovery.* Genome Biol, 2003. **4**(5): p. P3.

22. Zheng, Q. and X.J. Wang, *GOEAST: a web-based software toolkit for Gene Ontology enrichment analysis.* Nucleic Acids Res, 2008. **36**(Web Server issue): p. W358-63.

23. Pozzo-Miller, L.D., J.A. Connor, and S.B. Andrews, *Microheterogeneity of calcium signalling in dendrites.* The Journal of Physiology, 2000. **525**(1): p. 53-61.

24. Pongs, O. and J.R. Schwarz, *Ancillary subunits associated with voltage-dependent K+ channels.* Physiol Rev, 2010. **90**(2): p. 755-96.

25. Tcherkezian, J., et al., *Transmembrane receptor DCC associates with protein synthesis machinery and regulates translation.* Cell, 2010. **141**(4): p. 632-44.

26. Imai, Y., et al., *A product of the human gene adjacent to parkin is a component of Lewy bodies and suppresses Pael receptor-induced cell death.* J Biol Chem, 2003. **278**(51): p. 51901-10.

27. Rezgaoui, M., et al., *The neuropeptide head activator is a high-affinity ligand for the orphan G-protein-coupled receptor GPR37.* J Cell Sci, 2006. **119**(Pt 3): p. 542-9.

28. Faber, E.S. and P. Sah, *Functions of SK channels in central neurons.* Clin Exp Pharmacol Physiol, 2007. **34**(10): p. 1077-83.

29. Kanehisa, M. and S. Goto, *KEGG: kyoto encyclopedia of genes and genomes.* Nucleic Acids Res, 2000. **28**(1): p. 27-30.

30. Kanehisa, M., et al., *KEGG for representation and analysis of molecular networks involving diseases and drugs.* Nucleic acids research, 2010. **38**(Database issue): p. D355-60.

31. Schadt, E.E., et al., *An integrative genomics approach to infer causal associations between gene expression and disease.* Nat Genet, 2005. **37**(7): p. 710-7.

32. Fernandes, C., et al., *Hippocampal gene expression profiling across eight mouse inbred strains: towards understanding the molecular basis for behaviour.* Eur J Neurosci, 2004. **19**(9): p. 2576-82.

33. Hsieh, W.P., et al., *Mixed-model reanalysis of primate data suggests tissue and species biases in oligonucleotide-based gene expression profiles.* Genetics, 2003. **165**(2): p. 747-57.

34. Khaitovich, P., et al., *Parallel patterns of evolution in the genomes and transcriptomes of humans and chimpanzees.* Science, 2005. **309**(5742): p. 1850-4.

35. Insel, T.R. and L.J. Young, *Neuropeptides and the evolution of social behavior.* Curr Opin Neurobiol, 2000. **10**(6): p. 784-9.

36. Harmar, A.J., et al., *IUPHAR-DB: the IUPHAR database of G protein-coupled receptors and ion channels.* Nucleic Acids Res, 2009. **37**(Database issue): p. D680-5.

37. Buchhalter, J.R. and M.A. Dichter, *Electrophysiological comparison of pyramidal and stellate nonpyramidal neurons in dissociated cell culture of rat hippocampus.* Brain Res Bull, 1991. **26**(3): p. 333-338.

38. Eberwine, J., *Single-cell molecular biology.* Nature Neuroscience, 2001. **4**: p. 1155-1156.

39. Viereck, C., et al., *Phylogenetic conservation of brain microtubule-associated proteins MAP2 and tau.* Neuroscience, 1988. **26**(3): p. 893-904.

40. Tucker, R.P., L.I. Binder, and A.I. Matus, *Neuronal microtubule-associated proteins in the embryonic avian spinal cord.* The Journal of comparative neurology, 1988. **271**(1): p. 44-55.

41. Tucker, R.P., *The roles of microtubule-associated proteins in brain morphogenesis: a review.* Brain research. Brain research reviews, 1990. **15**(2): p. 101-20.

42. Irizarry, R.A., et al., *Exploration, normalization, and summaries of high density oligonucleotide array probe level data.* Biostatistics, 2003. **4**(2): p. 249-64.

43. Sul, J.Y., et al., *Transcriptome transfer produces a predictable cellular phenotype.* Proc Natl Acad Sci U S A, 2009. **106**(18): p. 7624-9.

44. Mathworks, *MATLAB.*

45. Park, C.S., et al., *Molecular network and chromosomal clustering of genes involved in synaptic plasticity in the hippocampus.* J Biol Chem, 2006. **281**(40): p. 30195-211.

46. Smit, A., Hubley, R and Green, P. *RepeatMasker Open-3.0.* 1996-2010; Available from: http://www.repeatmasker.org.

47. Fujita, P.A., et al., *The UCSC Genome Browser database: update 2011.* Nucleic acids research, 2011. **39**(Database issue): p. D876-82.

Chapter 3

CHARACTERIZATION AND COMPARATIVE ANALYSIS OF MRNAS POPULATION IN RAT AND MOUSE DENDRITES VIA LARGE-SCALE *IN SITU* HYBRIDIZATION

3.1 ABSTRACT

It is well recognized by now that certain populations of mRNAs are specifically targeted to neuronal projections, which allows for a localized control of translation and protein distributions. Nevertheless, the extent of this event and its potential phenotypic and evolutionary impacts remain unclear. To clarify the relationship between spatial patterns of dendritic transcriptome localization and genetic differences between species, I conducted a large-scale *in situ* hybridization. This was done on a curated list of mRNAs shown in our previous transcriptome study to be highly expressed in rat and mouse dendrites (Chapter 2).

I first optimized the imaging and data analysis procedure to allow a between and within species consistency throughout this study. After employing a very stringent filtering procedure on the signal intensity detected between the dendrites and soma, we concluded with high confidence that out of the 200 transcripts investigated commonly in rat and mouse, 84.5% showed localization in dendrites. Specifically 58% displayed a dendritic localization in both species; 29% and 13% displayed mouse and rat exclusive dendritic localization,

respectively. Overall, these three clusters contained a wide variety of transcripts involved in many neuronal functions that were either shared or species-specific. A sequence motif search highlighted the implication of introns-retained ID elements in RNA localization in dendrites, both in rat, as previously described in our lab, and in mouse. This study also provided a connection to diverse RNA binding proteins and microRNAs that could potentially be involved in mRNAs targeting and post-transcriptional regulation via either common or species-specific mechanisms.

These results complement previous studies examining transcripts localization in dendrites. It also reinforces the evolutionary diversification of this mechanism, as described in Chapter 2. The similarities and differences between mRNAs localization and pattern of distribution in rat and mouse dendrites underscore potential implications of this mechanism in the fine-tuned regulation of neuronal functions and the establishment of species-specific cognitive features.

3.2 INTRODUCTION

Phenotypic and behavioral diversity seen within and between species is partly due to genetic variation. Numerous and complex mechanisms can generate phenotypic variability [1-3]. The most well-studied of these involve sequence modifications that alter amino acid coding and ultimately affecting protein function. Differences in gene dosage effects as well as copy number variation have also been associated with phenotypic variation [4-8]. However, very few investigations have explored the relationship between differential spatial distribution of transcripts within tissues or cells and its potential impact on phenotypic differences.

In the brain, several studies have established that dendritically localized mRNAs are numerous and have a large functional repertoire. More specifically, our previous study using microarrays on isolated dendrites (detailed in Chapter 2) provided a comprehensive description of the dendritic RNA population in two closely related rodent species, rat and mouse. Our results showed that these species exhibit evolutionary divergence in their dendritic transcriptome with 81% difference among the highest expressed genes. The differential sub-cellular localization of specific genes suggests differences in neuronal architecture that could play an important role in determining species-specific cognitive features. Therefore, a greater understanding of the evolution of transcripts regulation and localization may provide new insights into key genes that are likely to be relevant in the evolution and diversification of phenotypes.

A major gap common to most dendritic transcriptome investigations is the lack of comprehensive assays at the single-cell level that would verify dendritic localization and provide detailed spatial information. Visualization via *in situ* hybridization can yield quantification and a refined resolution of differential spatial distribution of transcripts within and between species. In order to investigate the spatial pattern of RNA dendritic localization along with species differences, I conducted a large scale *in situ* hybridization using 400 most representative dendrite-specific genes in rat and mouse taken from our previous microarrays analysis (detailed in Chapter 2).

I carried out *in situ* hybridization on 14-day old primary cultures of cortical neurons hybridized with biotin-conjugated oligo-probes detected with a streptadivin-conjugated fluorescent dye. In addition, I imaged the dendrites processes using MAP2 immuno-staining and the nucleus using DAPI staining (see Materials and Methods). A custom-designed image analysis program allowed to extract spatial features and optimize the replicability of the data and verified the consistency between and within species. After a stringent filtering of all the data using the ratio of distal dendrites to soma of control transcripts, I recovered more than 165 significantly localized mRNAs in both species' dendrites.

Using the comprehensive image set of 400 probes, I explored several hypotheses. The functional classification of the dendritic transcripts highlighted their wide implications in common as well as species-specific cellular functions. I searched for sequence motifs within the dendrites-specific transcripts of both rat

and mouse. This outcome underlined a high incidence and a potential implication in transport or gene expression regulation, of non-coding elements such as the intron-retained ID-elements or the 3'UTR AU-rich elements. I established a relationship between the dendritically localized transcripts and potential trans-acting factors such as RNA binding proteins and microRNAs. All the recovered *cis* or *trans*-elements candidates are most likely acting together or in conjunction with other factors to regulate mRNAs trafficking and post-transcriptional events. The putative regulatory factors governing dendritic RNA localization show both common and species-specific patterns, the latter of which may underlie their role in the evolution of species-specific brain functions and diversification of phenotypes. All data from this study has been compiled in a database, http://kim.bio.upenn.edu/insitu/public/, which could become a potential public resource and guideline for a broad variety of future investigations.

3.3 RESULTS

3.3.1 Method Development, Screening, And Database Implementation

Identifying dendritically localized transcripts is made difficult by the fact that the mRNA concentration in dendrites is lower than the one in the soma by several orders of magnitude, which can produce RNA presence due to non-active diffusive flow from the soma. Therefore, it is essential to be able to define the relative level of a given transcript in dendrites compared to the soma. In order to do that, I carefully selected soma-specific transcripts in both rat and mouse (see Materials and Methods). These reference set of putatively soma-specific transcripts were used to establish a soma-specific reference value designated as "control cut off". Briefly, for each reference soma-specific gene, I computed average dendrite/soma intensity ratio and then set two standard deviations above the mean ratio as the threshold for significant dendritic localization (see Materials and Methods). To differentiate mRNAs present in dendrites due to random diffusion compared to the ones that are actively transported, I chose to estimate the functional localization of RNAs by geometrically dividing the neuronal cell into somatic, proximal and distal dendritic regions as detailed in the Materials and Methods section. I only designated transcripts to be dendritically localized if they exhibited a localization signal in the distal dendrites and the dendrites/soma (D/S) transcript ratio was higher than the reference cut-off value. In order to correct for various noise associated with *in situ* hybridization techniques, in collaboration with Jai-Yoon Sul, Ph.D. and Stephen Fisher, Ph.D., we developed

an image analysis procedure with optimal imaging resolution, consistency and high throughput data extraction (see Materials and Methods for more details). Finally, we created a neuronal mRNA localization database resource (http://kim.bio.upenn.edu/insitu/public/), containing annotation terms and representative images of all transcripts investigated in rat and mouse neurons.

3.3.2 Most of the transcripts investigated in mouse neurons are localized to dendrites.

Out of 242 transcripts investigated in mouse neurons, the majority (67.4%) displayed dendritic localization with a high confidence, as their D/S ratio was above the mouse control cut-off value. If the stringency in defining this cut-off value was decreased to only 1 standard deviation (SD) above the soma average instead of 2 SD, as described in the materials and methods section, 3.72% additional transcripts could be considered dendritically localized. Figure 3.1 illustrates the ranked D/S ratios detected for all transcripts examined via *in situ* hybridization on mouse neurons. In general, the D/S ratio trend increases linearly but the highest 10% transcripts exhibit a steeper rise. This top 10% cluster includes transcripts involved in signal transduction, transport and RNA and protein metabolisms. In contrast the bottom 10% cluster, corresponding to the transcripts with a D/S ratio just above the cut off value threshold (starting at 1SD above cut off which is the yellow bar in the Figure 3.1), includes vesicle trafficking functions, protein phosphorylation and acetylation and regulation of neurons

processes. Table 3.1 (provided at the end of this chapter) displays the list of all transcripts identified as dendritically localized in mouse.

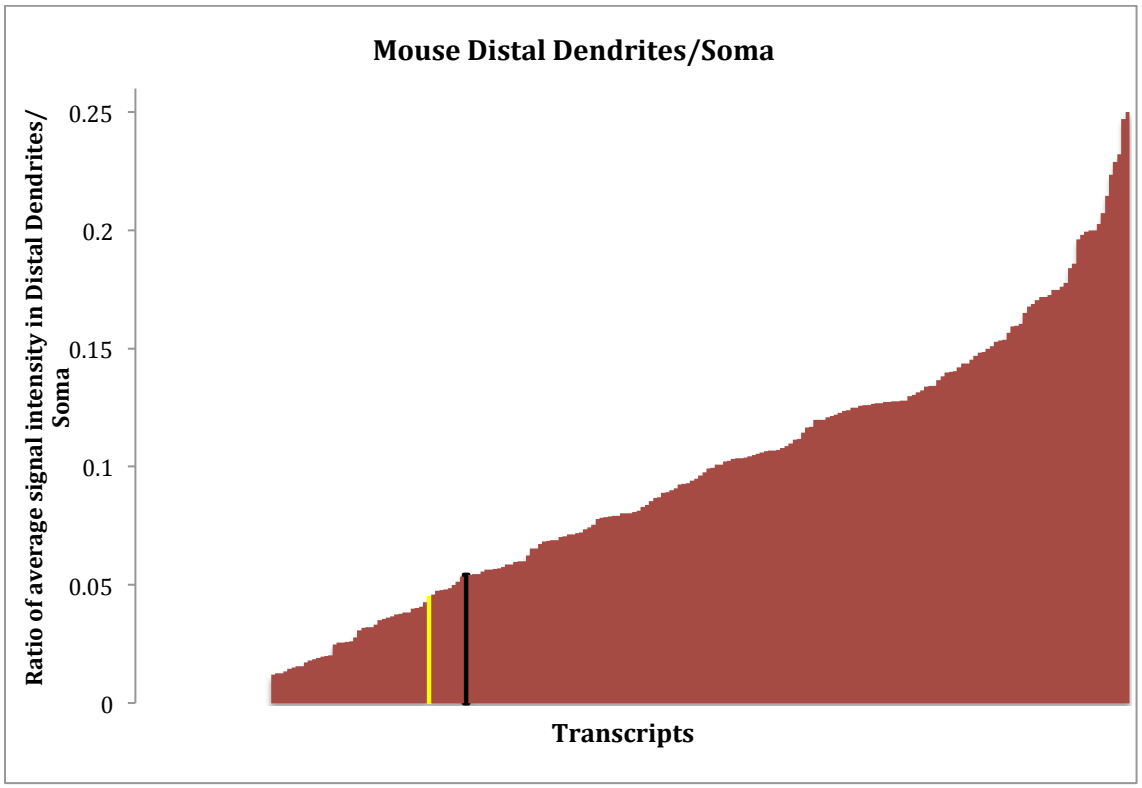


Figure 3.1: List of 242 transcripts examined via *in situ* hybridization on mouse neurons.

The Black bar represents the boundary above which transcripts are considered dendritically localized with high confidence and the yellow bar represents the boundary above which transcripts are considered dendritically localized with a lower 1 SD confidence.

3.3.3 Most of the transcripts investigated in rat neurons are localized to dendrites.

Similar to mouse, the majority (58.62%) of 290 transcripts investigated in rat neurons displayed dendritic localization with a high confidence, where their D/S ratio was above the rat control cut-off value. An additional 4.14% transcripts could be considered dendritically localized after decreasing the cut-off value's stringency to 1 standard deviation (SD) above the soma average instead of 2 SD, as described in the materials and methods section. Figure 3.2 illustrates the D/S ratio detected for all transcripts examined via *in situ* hybridization on rat neurons. Similarly to what was reported previously in the mouse, the general D/S ratio trend increases linearly until approximately the highest 10% transcripts where the increase of D/S ratio becomes much steeper. This top 10% rat cluster does not share any common transcripts with the Top10% cluster of the mouse however the biological functions of these rat transcripts correspond to similar functions as in the ones reported in the Top 10% mouse such as signal transduction, transport and metabolisms. The bottom 10% rat cluster, corresponding also to the transcripts with a D/S ratio just above the cut off value threshold (starting at 1SD above cut off which is the yellow bar in the Figure 3.2), includes protein translation, phosphorylation and acetylation, ion transport functions and regulation of neurotransmitter release. Table 3.2 (provided at the end of this chapter) displays the list of all transcripts considered dendritically localized in rat.

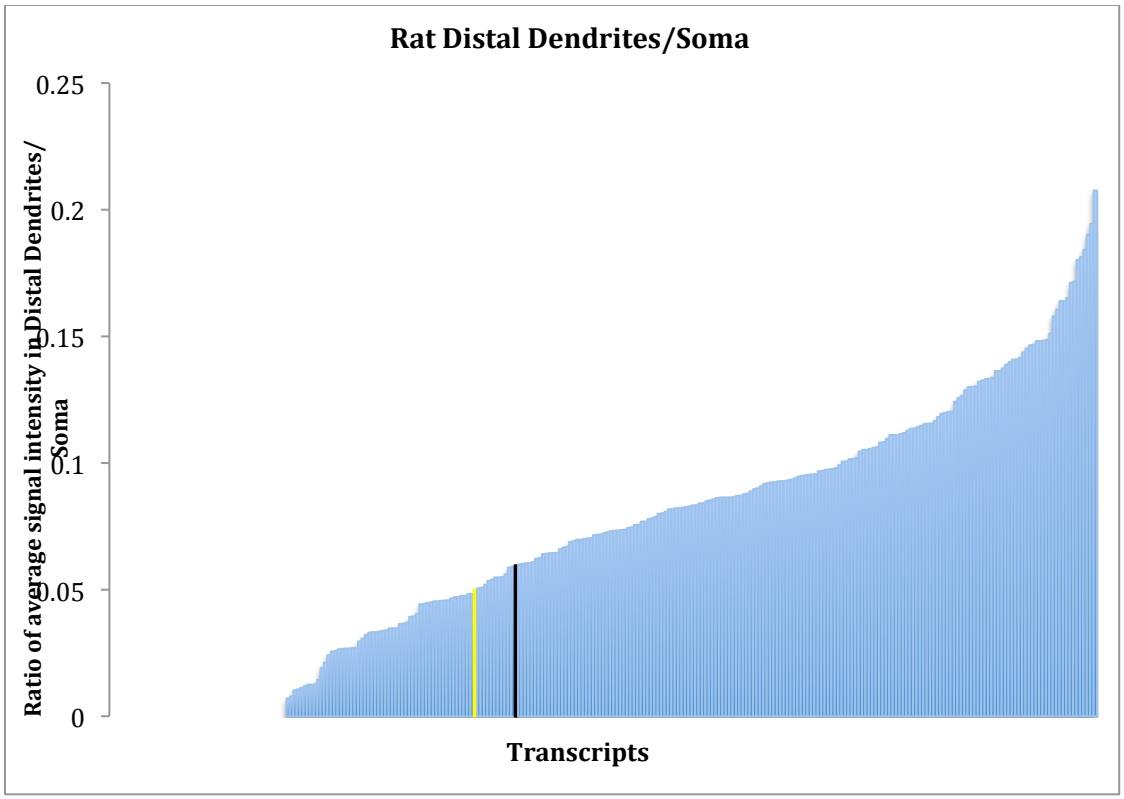


Figure 3.2: List of 290 transcripts examined via *in situ* hybridization on rat neurons.

The Black bar represents the boundary above which transcripts are considered dendritically localized with high confidence, and the yellow bar represents the boundary above which transcripts are considered dendritically localized with a lower 1SD confidence.

In general, and for both species, the main functionalities of the dendritically localized transcripts highlight their implication in local protein synthesis and in neuronal functions. Within these transcripts some that were reported previously by others such as MAP2, Glutamate and GABA receptor subunits and Calcium and Potassium channels [9-13].

Note that, regardless of the species in question, non-significant detection of mRNAs in dendrites via *in situ* hybridization does not necessarily mean lack of

localization. The absence or the weak labeling might be due to their low abundance or to a particular configuration that makes them inaccessible to the probe during the hybridization process. This lack of signal might also be due to physical masking of these mRNAs in dendrites, which interferes with the hybridization between the transcripts and their complementary probes. Another possible interpretation might be that these mRNAs, detected via microarrays, have in reality a relatively low concentration in dendrites compared to soma but are nevertheless present in the dendrite as we previously detected their presence in our microarray analysis of pooled dissected dendrites [14].

3.3.4 Analysis of the subcellular localization of mRNA transcripts within rat and mouse via *in situ* hybridization

Functional annotation of dendritic transcripts

I performed a Gene Ontology (GO) analysis, using the online tool for functional classification available in DAVID [15], to functionally classify the transcripts identified as dendritically localized with high confidence. I also repeated this analysis on the whole transcriptome in order to use this classification as a reference baseline. Not surprisingly, some categories in the dendritic transcripts, such as metabolism and transport, stood out in comparison to the reference transcriptome which implies that some functions are dendrite-specific rather than a random representation of the GO functional categories reported in the whole transcriptome. In fact, the reported dendrite-prevalent GO

categories are involved in neurons development and functions, which are consistent with previous findings and functional classification of dendritic mRNAs [16-18].

Functional categories of mouse dendritic transcripts

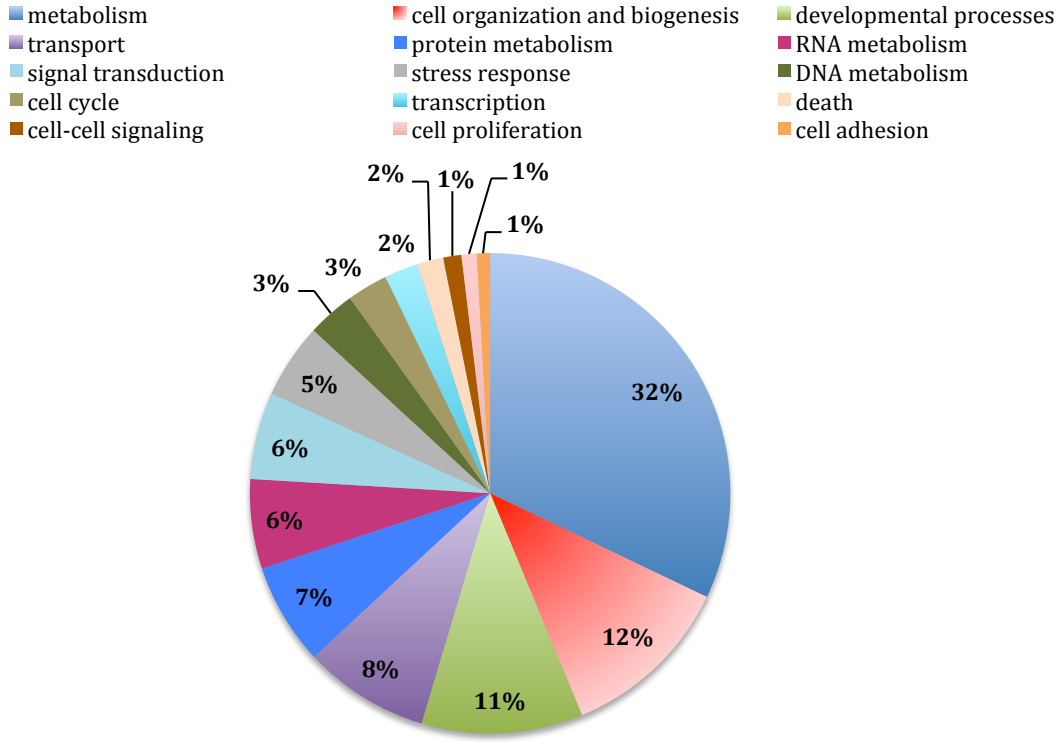
Figure 3.3 illustrates the main functional categories found in mouse. It highlights how the dendritic genes are governed by functions connected to metabolism as well as transport, developmental processes and cell adhesion. Particularly the metabolism and cell adhesion functional categories were significantly enriched in mouse dendrites compared to the whole transcriptome (Chi-square, FDR<0.01).

Functional categories of rat dendritic transcripts

Figure 3.4 illustrates the main functional categories found in rat. As in mouse, metabolism was significantly enriched in rat dendrites compared to the whole transcriptome (Chi-square, FDR <0.01). In general most of the functional categories found in rat dendritic genes were also detected in mouse dendritic transcripts. Nevertheless, the cell adhesion functional category, which is significant in mouse, was not represented in the rat dendritic genes but instead the cell-cell signaling category was particularly prominent (Chi-square, FDR <0.01) in rat dendritic transcripts but not in mouse. Actually these two functional categories are interconnected and important for proper neuronal growth,

synaptogenesis and signaling [19]. Their different level of representation in the rat and mouse dendrites GO categories highlights variation in their level of expression which bypasses our stringent cut off value in one species but not the other (and vice versa). A recent study done in rat and mouse on more than 30 adhesion G protein couple receptors supports our idea [20]. In this investigation, both species present the same patterns of gene expression in several brain tissues but the relative expression value for a given G protein couple receptors transcript was significantly different in one species versus the other. Additionally it is known that most of the cell-adhesion and cell-cell signaling genes, such as Cadherins and Protocadherins, are differentially expressed in different brain regions or neurons and their level of expression is calcium and phosphorylation-dependent [21-23] which might also be species-dependent. Indeed, recent discoveries of multiple Protocadherins (~50) and their wide heterodimers combinations (~2500) provide high degree of variation in these receptors that might contribute to species-specific synaptic functions [24, 25].

Functional Categories: Mouse Transcriptome



Functional Categories: Mouse dendritic transcripts

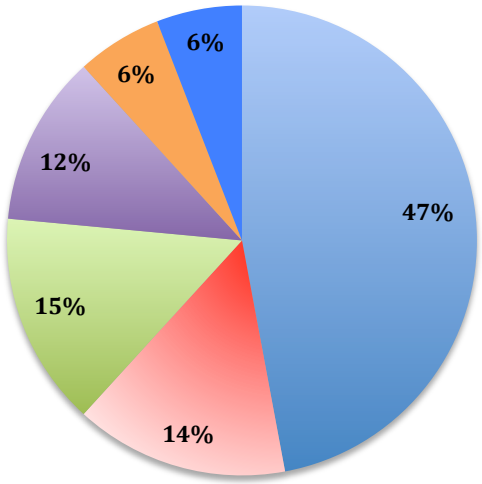
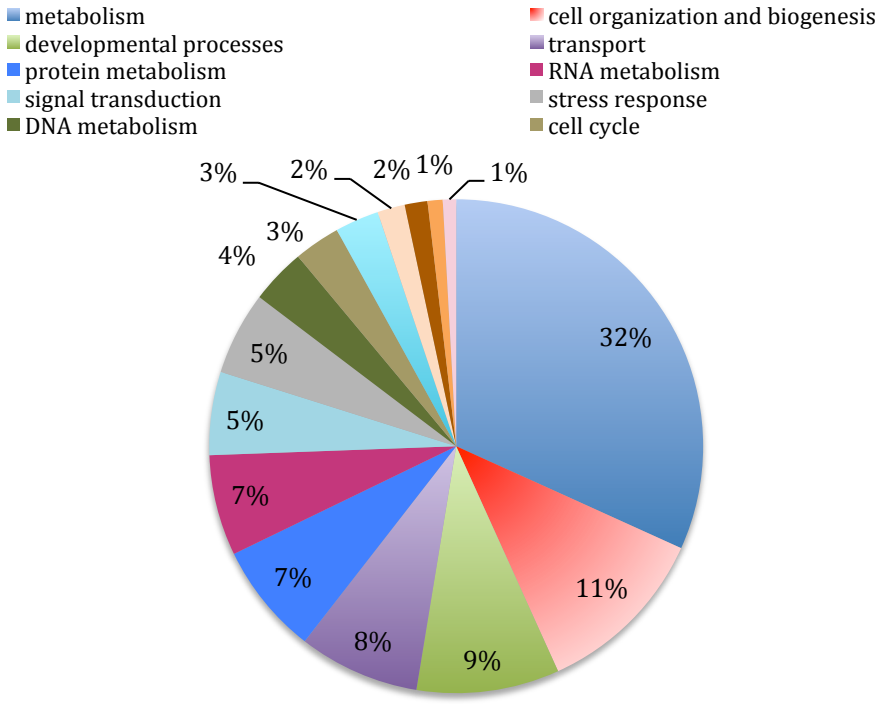


Figure 3.3: Mouse Functional categories

Functional Categories: Rat Transcriptome



Functional Categories: Rat dendritic transcripts

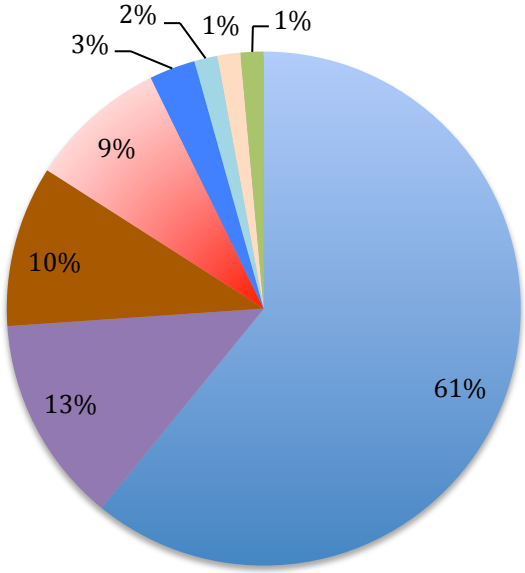


Figure 3.4: Rat Functional categories

Dendritic transcripts do not exhibit an unusually long half-life

In order to be locally translated in dendrites, mRNA must be transported over long distances while remaining stable. One can wonder if these dendritically localized mRNA possessed a longer half-life than the general transcriptome trend. This could potentially support their long distance trafficking and localized translation upon a given stimuli. I compared our mouse dendritic transcripts against a publically available National Institute of Health, NIH (http://lgsun.grc.nia.nih.gov/mRNA/) database for mRNA half-life in mouse [26]. This comparison did not show any trend toward a longer half-life for subcellularly localized transcripts (t-test with $p>0.05$), at least based on our mouse data. This result suggested that the regulation of subcellularly localized mRNAs is not simply based on its intrinsic stability but is likely linked to complex events involving interaction between *trans*-acting elements and *cis*-acting elements.

Diverse *trans*-factors interact with rat and mouse dendritic transcripts.

How do mRNAs get to their final destination? It is clear that some kind of spatio-temporal mechanism regulates local mRNAs translation in dendrites [27, 28], but the details of the putative mechanism are not well known. Many studies have suggested that a key factor in mRNA trafficking is based on their interaction with RNA binding proteins (RBPs) as well as other messenger ribonucleoproteins (mRNPs) to form large complexes called RNA granules [29, 30]. The formation of

these granules is promoted by an interaction between the “zipcodes” *cis*-acting sequences, that mainly lie in the 3’UTR region within the mRNA, and their corresponding *trans*-acting RBPs [31, 32]. The identification of RBPs has been a very challenging and tedious task as the transported mRNAs are often associated with large multi-protein complexes and are in low abundance. A large number of proteomic and biochemical studies have shown that the composition of RNA granules is far from being homogenous [30, 33]. These contain different RBPs such as Staufen, zipcode-binding protein 1 (ZBP1) and heteronuclear RNP-A2 (hnRNP-A2) as well as hundreds of mRNA species. Some identified mRNA species include calcium/ calmodulin-dependent protein kinase II (*CaMKII*) alpha, activity-regulated cytoskeletal-associated (*Arc*), beta-actin, the noncoding BC1, glutamate receptor subtypes and translational machinery components, such as ribosome subunits and elongation factor 1a [34-36]. All these studies suggest that the composition of these granules is not universal and could vary depending on the location in the cell, the cellular activity, and the stage of transport or anchoring required. Using *in situ* hybridizations, several experiments that examined some RBPs and/or their mRNA have reported the presence of a non-uniform and punctate distribution patterns in the subcellular regions of neurons [18, 33, 37], which may be associated with local translation.

Here, using this large-scale *in situ* data, I conducted an analysis to identify the relationship between dendritic gene expression patterns and their potential link to specific RBPs (see Materials and Methods). I hypothesized that transcripts sharing similar patterns of distribution might share common RBPs. First, I

surveyed the transcripts localized to dendrites and grouped them into two classes: those showing punctate localization within dendrites and those showing uniform distribution within the dendrites. This clustering was based on sorting the data via the average pixel intensity variance to mean ratio recorded for each image and then refined with a visual inspection as an additional level of filtering. A detailed description and illustration of this procedure is available in the Material and Methods section. The two main types of distribution patterns detected in both mouse and rat dendrites are illustrated in the Figure 3.5(A-B), and the identity of the transcripts investigated in mouse and rat respectively is provided in Tables 3.3 and 3.4. Second, I searched within both classes for each transcript sequence's affinity to RNA binding proteins using a publicly available RNA Binding Protein database (http://rbpdb.ccbr.utoronto.ca/) [38]. Further detail on this screen is available in the Materials and Methods section. As discussed below, the resulting RBPs identified to specifically interact with the dendritic transcripts are all previously hypothesized to be involved in the regulation of gene expression.

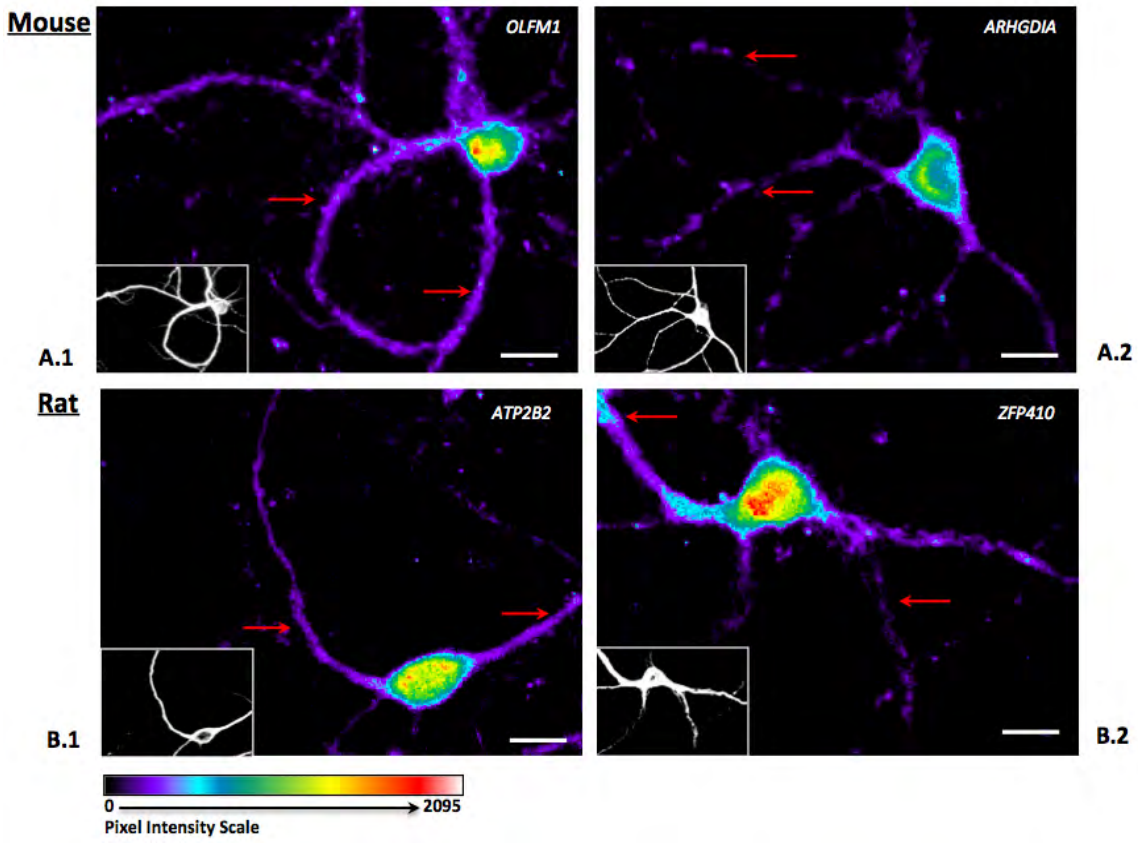


Figure 3.5: *In situ* hybridization reveals different patterns of localization in neuronal dendrites.

Fluorescent Microscopy evaluation of biotin-conjugated oligoprobes on paraformaldehyde fixed 14-day cultured rat and mouse cortical neurons hybridized with 9 biotin-conjugated oligoprobes detected with streptadivin-Alexa568. For each image, the small bottom left corner panels represent MAP2 immuno-staining. Scale bar = 20µm. Various distribution patterns are highlighted with red arrows.

(A), In mouse:
(A.1) Probe against *OLFM1* transcript illustrates a uniform distribution in dendrites
(A.2) Probe against *ARHGDIA* transcript illustrates a punctate distribution in dendrites.
(B), In rat:
(B.1) Probe against *ATP2B2* transcript illustrates a uniform distribution in dendrites
(B.2) Probe against *ZFP410* transcript illustrates a punctate distribution in dendrites.

Table 3.3: List of mouse mRNA examined for patterns of distribution in dendrites.

Gene Symbol	Mouse Refseq	Distribution in dendrites
Glp2r	NM_175681	Uniform
Pcdh10	NM_011043	Uniform
Rsp4	NM_009094	Uniform
Vim	NM_011701	Uniform
Glp2r	NM_175681	Uniform
Pcdh17	NM_001013753	Uniform
Rnf111	NM_033604	Uniform
Ak7	XM_994344	Uniform
Atp5d	NM_172294	Uniform
Dpysl2	NM_009955	Uniform
Rbfox1	NM_183188	Uniform
Sirt1	NM_001159589	Uniform
Syt6	NM_018800	Uniform
Arl6ip5	NM_022992	Uniform
H13	NM_010376	Uniform
Olf1273	NM_146442	Uniform
Syt4	NM_009308	Uniform
Myo5a	NM_010864	Uniform
Appbp2	NM_025825	Uniform
Olfm1	NM_019498	Uniform
Rpl23	XM_001477371	Punctate
Snrpn	XM_001480670	Punctate
Kif15	NM_010620	Punctate
Usp9x	NM_009481	Punctate
Opa1	NM_133752	Punctate
Trpv5	NM_001007572	Punctate
Atp5a1	NM_007505	Punctate
Fmnl1	NM_001077698	Punctate
Jub	NM_010590	Punctate
Lypla1	NM_008866	Punctate
Mdh2	NM_008617	Punctate
Arhgef7	NM_017402	Punctate
Mapk1	NM_001038663	Punctate
Rpl6	NM_011290	Punctate
Tmeff1	NM_021436	Punctate
Vmn2r57	NM_177764	Punctate
Dync1i1	NM_010063	Punctate
Nefl	NM_010910	Punctate
Cacng5	NM_080644	Punctate
Slc25a40	NM_178766	Punctate
Sept9	NM_001113488	Punctate
Gnao1	NM_010308	Punctate
Syngr1a	NM_207708	Punctate
Syngr1b	NM_207708	Punctate
Arhgdia	NM_133796	Punctate
Dnajc5	NM_016775	Punctate
Kcna2	NM_008417	Punctate
Gnb1	NM_010312	Punctate
H2afz	XM_001480384	Punctate
Irgm1	NM_008326	Punctate
Atp2b2	NM_009723	Punctate
Mtap2	NM_008632	Punctate

Table 3.4: List of rat mRNA examined for patterns of distribution in dendrites.

Gene Symbol	Rat Refseq	Distribution in dendrites
Sept9	NM_176856.1	Uniform
Ak7	XM_234507.4	Uniform
Arhgef7	NM_053740.1	Uniform
Atp2b2	NM_012508.3	Uniform
Atp5d	NM_139106.1	Uniform
Cacng5	NM_080693.1	Uniform
Dpysl2	XM_573810.2	Uniform
Dync1i1	NM_019234.1	Uniform
Glp2r	NM_021848.1	Uniform
Irgm	NM_001012007.1	Uniform
Kif15	NM_181635.2	Uniform
Lypla1	NM_013006.1	Uniform
Nefl	NM_031783.1	Uniform
Olfm1	NM_053573.1	Uniform
Olr259	NM_001000222.1	Uniform
Pcdh17	XM_224389.4	Uniform
Prkch	NM_031085.2	Uniform
Rpl23	NM_001007599.1	Uniform
Rps20	NM_001007603.1	Uniform
Rsp4	NM_001007600.1	Uniform
Sirt1	XM_228146.4	Uniform
Syt6	NM_022191.1	Uniform
Arl6ip5	NM_023972.2	Uniform
Atp5g3	NM_053756.1	Uniform
Gnao1	NM_017327.1	Uniform
H2afz	NM_022674.1	Uniform
Mapk1	NM_053842.1	Uniform
Mapre3	NM_001007656.1	Uniform
Myo5a	NM_022178.1	Uniform
Rpl6	NM_053971.1	Uniform
Snrpn	NM_031117.1	Uniform
Vgf	NM_030997.1	Uniform
Vim	NM_031140.1	Uniform
Trmt112	XM_215167.2	Uniform
Atp5a1	NM_023093.1	Uniform
Dnajc5	NM_024161.2	Punctate
Gnb1	NM_030987.1	Punctate
H13	NM_001107789.1	Punctate
Spag7	XM_001079933.1	Punctate
Stmn3	NM_024346.1	Punctate
Syngr1	NM_019166.1	Punctate
Syt4	NM_031693.1	Punctate
Usp9x	XM_343766.3	Punctate
Appbp2	XM_001081113.1	Punctate
Grina	NM_153308.1	Punctate
Jub	NM_053503.1	Punctate
Opa1	NM_133585.2	Punctate
Ppm1e	NM_198773.1	Punctate
Tmeff1	NM_023020.1	Punctate
Vmn2r57	NM_173130.1	Punctate
Zfp410	XM_234409	Punctate
Fmnl1	XM_001081542.1	Punctate

Different types of RBPs are found in both punctate and uniformly distributed mouse dendritic transcripts.

The sequence analysis of the subset of transcripts differentially distributed in mouse dendrites (Table 3.3) clearly distinguished some RBPs that had higher affinity for one group of spatial pattern over another. Table 3.5 summarizes the level of incidence, in percentage, of the RBPs associated with each kind of dendritic spatial pattern in the mouse transcripts. Some RBPs like the small nuclear ribonucleoprotein A, SNRPA, Ataxin-2 binding protein 1, A2BP1 (also known as RNA binding protein fox-1 homolog, Rbfox1), or Zinc finger protein 36, ZFP36 (also known as Tristetraprolin, TTP), were equally represented in both punctate and uniformly distributed transcripts. These three different RBPs were shown to be involved in mRNA splicing, stability, regulation of mRNA poly (A) tail shortening and to participate in mRNA transport [39-42]. Other RBPs were more specific to one spatial class compared to the other. The arginine/serine-rich splicing factor proteins SRFS1 and the embryonic lethal, abnormal vision, Drosophila)-like 2 ELAVL2 (also known as Hu antigen B) protein, which are implicated in alternative splicing, stabilization and/or enhanced translation of ARE-containing mRNA [43, 44] were particularly prevalent in the punctate transcripts class. In contrast, the arginine/serine-rich splicing factor proteins SRFS2 and KH-type splicing regulatory protein (KHSRP, also known as MARTA1/2 in rat), which have also been involved in alternative pre-mRNA splicing and mRNA localization [45, 46], were more frequently associated with

the uniformly distributed transcripts. The occurrence of SRFS1 and SFRS2 (also known as ASF/SF2 and SC35 respectively) in two different patterns of transcripts distribution in dendrites might relate to their known differences in RNA binding specificities and their antagonist functions particularly in mRNA splicing regulation [47-49]. Likewise, ELVAL2 and KHSRP, known for their affinity to ARE-containing mRNAs, could be interacting with different co-transfactors that might result in the differential patterns of distribution detected in our investigations [46, 50].

Table 3.5: Level of RBPs' occurrence in mouse punctate and uniformly-distributed dendritic genes

RBP	Motif Sequence	Motif Length	Uniform (n=21)	Puncta (n=31)
ELAVL2	UUUUAUUUA/AACCUUUUUUUUCU	9	19.05%	41.94%
SFRS1	AAGACAGAGC	10	9.52%	29.03%
A2BP1	UGCAUG	6	23.81%	29.03%
ZFP36	AAAAAAAAAAG/AAAAAGGAAAG	11	28.57%	32.26%
Rna15	UGUGUAUUCUCC	12	0.00%	3.23%
sus	UCAGGAGUCU	10	0.00%	3.23%
ZRANB2	AGGUAA	6	0.00%	3.23%
ybx2*	AACAUC	6	0.00%	0.00%
sap-49	GUGUGA	6	4.76%	3.23%
SFRS7	AGACUACGAG	10	4.76%	3.23%
SNRPA	GGGUAUGCUG	10	28.57%	25.81%
NCL	UGCCCAGAAGG	11	9.52%	6.45%
QKI	UACUAAC	6	4.76%	0.00%
pum	AAUAUAAAUAUAUAUA/UUUUUAAAUAUAAAAA	16	9.52%	3.23%
EIF4B	GUUGGAA	7	9.52%	0.00%
SFRS2	UGUUCGAGUA	10	23.81%	9.68%
KHSRP	CCCCCCCC	8	19.05%	0.00%

Different types of RBPs are found in punctate and uniformly distributed rat dendritic transcripts.

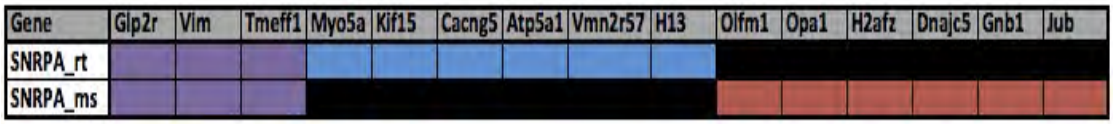
The sequence analysis of the subset of transcripts differentially distributed in rat dendrites (Table 3.4) also distinguished some RBPs that had higher affinity toward one group of spatial pattern over another. Table 3.6 summarizes the level of incidence, in percentage, of the RBPs that scored the highest affinities toward the rat transcripts. A total of 94% (16 out of 17) of these rat RBPs were also reported previously in mouse and all of them (100%) have been implicated in either mRNA splicing, stability, poly (A) tail shortening and/or mRNA transport. On one hand, as seen in the mouse, A2BP1 and ZFP36 showed equal affinity for both punctate and uniformly distributed rat transcripts; and, ELVAL2 was particularly prevalent in the rat punctate transcripts. On the other hand many other RBPs had different level of occurrence in rat and mouse transcripts (Table 3.6 versus Table 3.5) such as SNRPA and YbX2 (Y-box-binding protein 2), which are clearly more prevalent in the rat punctate and uniformly distributed transcripts respectively. One reason behind these differences in distribution patterns could be due to species-specific nucleotide sequence variations in orthologous genes. Indeed the identity of the transcripts associated with the RBPs does not match perfectly between species. For instance, 67% of the transcripts associated with

SNRPA are non-homologous between rat and mouse. This example is illustrated in Table 3.7 and further detail on this between species analysis will be addressed in the sections below and in Table 3.14. These results suggest that transcript' localization patterns might be based upon species-specific interconnections between transcript's identity and the matching RBPs.

Table 3.6: Level of RBPs' occurrence in rat punctate and uniformly distributed dendritic genes

RBP	Motif Sequence	Motif Length	Uniform (n=34)	Puncta (n=17)
ELAVL2	UUUUAUUUA/AACCUUUUUUUUCU	9	11.76%	41.18%
SNRPA	GGGUAUGCUG	10	14.71%	35.29%
QKI	UACUAAC	6	0.00%	11.76%
SFRS1	AAGACAGAGC	10	11.76%	23.53%
Rna15	UGUGUAUUCUCC	12	0.00%	5.88%
SFRS2	UGUUCGAGUA	10	17.65%	23.53%
KHSRP	CCCCCCCC	8	2.94%	5.88%
pum	AAUAUAAAUAUAUAUA/UUUUUAAAUAUAAAAA	16	5.88%	5.88%
SFRS7	AGACUACGAG	10	5.88%	5.88%
ZFP36	AAAAAAAAAAG/AAAAAGGAAAG	11	20.59%	17.65%
NCL	UGCCCAGAAGG	11	2.94%	0.00%
sus	UCAGGAGUCU	10	2.94%	0.00%
A2BP1	UGCAUG	6	23.53%	17.65%
sap-49	GUGUGA	6	5.88%	0.00%
ZRANB2	AGGUAA	6	8.82%	0.00%
EIF4B	GUUGGAA	7	11.76%	0.00%
ybx2-a	AACAUC	6	14.71%	0.00%

Table 3.7: Rat and mouse transcripts with high binding affinity to the RNA binding protein SNRPA



Transcripts were investigated in rat (_rt) and mouse (_ms), via the RBPdatabase, for their sequence motif for SNRPA (n=15). Only 3 out of 15 transcripts were homologous between mouse and rat (marked in purple), six transcripts were rat-specific for SNRPA (marked in blue) and six transcripts in were mouse-specific for SNRPA (marked in red)

Regulation of localized gene expression in dendrites: a link to MicroRNAs

MicroRNAs (miRNAs) are particularly known to regulate gene expression by repressing mRNA translation via their interaction with Argonaute proteins, or by increasing mRNA degradation rate. MicroRNAs target their transcripts mainly through binding to secondary structures within the 3'UTR. Their recognition binding sites are usually short, between 6 to 8 nucleotides [51]. The first indication of miRNAs involvement in synaptic plasticity came from Drosophila, where CaMKII expression was silenced by a miRNA-mediated repression [52]. Further investigations have shown that, during long-term memory formation, a tight connection exists between the down-regulation of microRNAs (i.e. miR-138 and miR-124) and the increased expression of certain transcripts (i.e. *CaMKII* and *CREB*) [52-54]. It is now clearly established that gene silencing by miRNAs is active in neurons and participates in neural development and synaptic plasticity [55].

In order to get an insight into the regulation of subcellular gene expression via miRNAs silencing, I searched within our curated dendritic mRNA list for miRNA

affinity using the comprehensive miRWalk database of predicted as well as validated microRNA targets (http://www.ma.uni-heidelberg.de/apps/zmf/mirwalk/) [56]. I only considered experimentally validated microRNAs, available through miRWalk [56], as significant candidates for further investigations (see Materials and Methods). Note that the level of experimental certainty for the various miRNA function and target information for the entire 2044 miRNAs set provided by miRWalk is variable. Therefore, the results here need to be interpreted with caution. Nevertheless, the miRWalk database represents one of the most comprehensive integrated information resources for miRNA function and interaction and therefore I used it as a discovery tool for linking miRNA function to dendritic genes.

Mouse dendritic transcripts are likely to be regulated by microRNAs

Within the mouse dendritic transcripts almost 12% were linked to validated miRNAs. A total of 93 microRNAs were identified from the miRWalk database. A vast majority of these (>90%), such as miR16 and miR124, were reported to be functionally important in regulating gene expression in the brain. Additionally, some of these miRNAs, like miR138 and miR375, have been previously described in the dendrites [56-60]. Table 3.8 details the identities of these microRNAs and their mouse mRNA targets.

Rat dendritic transcripts are likely to be regulated by microRNAs

Close to 10% of rat dendritic transcripts were linked to validated miRNAs. A total of 98 microRNAs were identified from the miRWalk database. As in mouse, a vast majority of these rat microRNAs (>86%) were shown to be relevant for gene expression in the brain [56]. Table 3.9 details the identities of these microRNAs and their rat mRNA targets.

As the vast majority (~88% on average) of the miRNAs reported here in mouse and rat were previously shown to play a key role in the regulation of gene expression in the brain, this reinforces the potential role of their dendritic transcripts targets (listed in Tables 3.8 and 3.9) in neuronal functions (i.e. synaptic plasticity and neurogenesis). Additionally, 40% of all miRNAs that were annotated in the miRWalk database to be functionally relevant in the CNS were linked to our rat and mouse dendritic transcripts. Based on this fairly large proportion (40%) of CNS miRNAs targeting dendritic RNA, I postulate that miRNAs function in the CNS could be mediated through modulation of dendritic function. Finally, as only a relatively small number of the reported dendritic transcripts (~11% on average) were linked to miRNAs, this suggests that most of remaining transcripts may be post-transcriptionally regulated by other *trans*-factors, such as the RBPs discussed above and/or by *cis*-factors which will be examined in the next section.

Table 3.8: List of microRNAs associated with mouse dendritic transcripts

Gene	EntrezID	MicroRNA Name
ATF3	11910	miR-200b,miR-214
BASP1	70350	miR-9
CDH2	12558	miR-141,miR-155,miR-200a,miR-200b,miR-200c,miR-203,miR-205,miR-23b,miR-29c,miR-369-5p,miR-370,miR-375,miR-429,miR-542-5p
IRGM	15944	miR-196a,miR-196b
JUB	16475	miR-125a-3p,miR-125a-5p,miR-125b,miR-125b-3p,miR-125b-5p
LHX3	16871	miR-17,miR-181b
LYPLA1	18777	miR-138
MAPK1	26413	let-7a,let-7b,let-7c,let-7d,let-7e,let-7f,let-7i,miR-100,miR-101a,miR-101b,miR-124,miR-125a-5p,miR-130a,miR-132,miR-138,miR-139-5p,miR-140,miR-146b,miR-150,miR-17,miR-181b,miR-185,miR-19a,miR-204,miR-21,miR-214,miR-215,miR-218,miR-223,miR-23a,miR-23b,miR-24,miR-25,miR-26b,miR-27a,miR-27b,miR-290-3p,miR-290-5p,miR-291a-3p,miR-291a-5p,miR-291b-3p,miR-291b-5p,miR-292-5p,miR-295,miR-298,miR-29a,miR-29b,miR-29c,miR-301a,miR-30a,miR-30c,miR-30e,miR-320,miR-324-5p,miR-327,miR-328,miR-331-3p,miR-331-5p,miR-338-3p,miR-338-5p,miR-342-5p,miR-34a,miR-374,miR-382,miR-383,miR-409-3p,miR-485,miR-494,miR-497,miR-543,miR-7a,miR-9,miR-92a,miR-98,miR-99a,miR-99b
MDH2	17448	miR-23b
MRPS30	59054	miR-30e
MTAP2	17756	miR-17,miR-181b,miR-26a
PCDH18	73173	miR-9
PDZD7	0	let-7a,let-7b,let-7c,let-7d,let-7e,let-7f,let-7i,miR-100,miR-101a,miR-101b,miR-122,miR-124,miR-125a-5p,miR-130a,miR-132,miR-138,miR-139-5p,miR-150,miR-155,miR-181b,miR-185,miR-204,miR-206,miR-21,miR-214,miR-215,miR-218,miR-223,miR-23a,miR-23b,miR-24,miR-25,miR-26b,miR-27a,miR-290-3p,miR-290-5p,miR-292-5p,miR-298,miR-29a,miR-29b,miR-29c,miR-301a,miR-30a,miR-30c,miR-30e,miR-320,miR-324-5p,miR-327,miR-328,miR-331-3p,miR-331-5p,miR-338-3p,miR-338-5p,miR-342-5p,miR-34a,miR-374,miR-375,miR-382,miR-383,miR-409-3p,miR-485,miR-494,miR-497,miR-543,miR-7a,miR-9,miR-92a,miR-98,miR-99a,miR-99b
POU4F2	18997	miR-1,miR-130a,miR-19a,miR-23a,miR-23b,miR-26a,miR-34a
PRPH1	19132	miR-1,miR-183,miR-96
SIRT1	93759	miR-134,miR-199a-5p,miR-34a,miR-9
SYT4	20983	let-7a,let-7b,let-7c,let-7d,let-7e,let-7f,let-7i,miR-100,miR-101a,miR-101b,miR-124,miR-125a-5p,miR-130a,miR-132,miR-138,miR-139-5p,miR-150,miR-181b,miR-185,miR-204,miR-21,miR-214,miR-215,miR-218,miR-223,miR-23a,miR-23b,miR-24,miR-25,miR-26b,miR-27a,miR-290-3p,miR-290-5p,miR-292-5p,miR-298,miR-29a,miR-29b,miR-29c,miR-301a,miR-30a,miR-30c,miR-30e,miR-320,miR-324-5p,miR-327,miR-328,miR-331-3p,miR-331-5p,miR-338-3p,miR-338-5p,miR-342-5p,miR-34a,miR-374,miR-382,miR-383,miR-409-3p,miR-485,miR-494,miR-497,miR-543,miR-7a,miR-9,miR-92a,miR-98,miR-99a,miR-99b
VIM	22352	let-7a,let-7b,let-7c,let-7d,let-7e,let-7f,let-7i,miR-100,miR-101a,miR-101b,miR-124,miR-125a-5p,miR-130a,miR-132,miR-138,miR-139-5p,miR-150,miR-181b,miR-185,miR-204,miR-21,miR-214,miR-215,miR-218,miR-223,miR-23a,miR-23b,miR-24,miR-25,miR-26b,miR-27a,miR-290-3p,miR-290-5p,miR-292-5p,miR-298,miR-29a,miR-29b,miR-29c,miR-301a,miR-30a,miR-30c,miR-30e,miR-320,miR-324-5p,miR-327,miR-328,miR-331-3p,miR-331-5p,miR-338-3p,miR-338-5p,miR-342-5p,miR-34a,miR-374,miR-382,miR-383,miR-409-3p,miR-485,miR-494,miR-497,miR-543,miR-7a,miR-9,miR-92a,miR-98,miR-99a,miR-99b
VSNL1	26950	let-7a,let-7b,let-7c,let-7d,let-7e,let-7f,let-7g,let-7i,miR-100,miR-101a,miR-101b,miR-124,miR-125a-5p,miR-126-3p,miR-126-5p,miR-130a,miR-132,miR-138,miR-139-5p,miR-150,miR-181b,miR-185,miR-196a,miR-204,miR-21,miR-214,miR-215,miR-218,miR-223,miR-23a,miR-23b,miR-24,miR-25,miR-26b,miR-27a,miR-290-3p,miR-290-5p,miR-292-5p,miR-298,miR-29a,miR-29b,miR-29c,miR-301a,miR-30a,miR-30c,miR-30e,miR-31,miR-320,miR-324-5p,miR-327,miR-328,miR-331-3p,miR-331-5p,miR-338-3p,miR-338-5p,miR-342-5p,miR-34a,miR-374,miR-382,miR-383,miR-409-3p,miR-485,miR-494,miR-497,miR-543,miR-7a,miR-9,miR-92a,miR-98,miR-99a,miR-99b

Table 3.9: List of microRNAs associated with rat dendritic transcripts

Gene	Entrez ID	MicroRNA Name
CALM3	24244	miR-1
GNB1	24400	let-7a, let-7b, let-7c, let-7d, let-7e, let-7f, let-7i, miR-100, miR-101a, miR-101b, miR-124, miR-125a-5p, miR-129, miR-130a, miR-132, miR-138, miR-139-5p, miR-150, miR-181b, miR-185, miR-204, miR-21, miR-214, miR-215, miR-218, miR-223, miR-23a, miR-23b, miR-24, miR-25, miR-26b, miR-27a, miR-290, miR-292-5p, miR-298, miR-29a, miR-29b, miR-29c, miR-301a, miR-30a, miR-30b-5p, miR-30c, miR-30e, miR-320, miR-324-5p, miR-327, miR-328, miR-331, miR-338, miR-342-5p, miR-34a, miR-374, miR-382, miR-383, miR-409-3p, miR-485, miR-494, miR-497, miR-543, miR-664, miR-7a, miR-9,miR-92a, miR-98, miR-99a, miR-99b
GRB2	81504	miR-183, miR-433, miR-9
GRIA2	29627	miR-181b, miR-30a, miR-30b-3p, miR-30b-5p, miR-30c, miR-30c-1, miR-30c-2, miR-30d, miR-30e
JUB	85265	miR-125a-3p, miR-125a-5p, miR-125b, miR-125b-3p, miR-125b-5p
LYPLA1	25514	miR-138
MAPK1	116590	let-7a, let-7b, let-7c, let-7d, let-7e, let-7f, let-7i, miR-100, miR-101a, miR-101b, miR-124, miR-125a-5p, miR-129, miR-130a, miR-132, miR-138, miR-139-5p, miR-150, miR-181b, miR-185, miR-204, miR-21, miR-214, miR-215, miR-218, miR-223, miR-23a, miR-23b, miR-24, miR-25, miR-26b, miR-27a, miR-290, miR-292-5p, miR-298, miR-29a, miR-29b, miR-29c, miR-301a, miR-30a, miR-30b-5p, miR-30c, miR-30e, miR-320, miR-324-5p, miR-327, miR-328, miR-331, miR-338, miR-342-5p, miR-34a, miR-374, miR-382, miR-383, miR-409-3p, miR-485, miR-494, miR-497, miR-543, miR-664, miR-7a, miR-9, miR-92a, miR-98, miR-99a, miR-99b
MAPK3	50689	let-7a, let-7b, let-7c, let-7d, let-7e, let-7f, let-7i, miR-125a-3p, miR-125a-5p, miR-125b, miR-125b-3p, miR-125b-5p, miR-132, miR-143, miR-146a, miR-146b, miR-15b, miR-181a, miR-205, miR-21, miR-212, miR-221, miR-222, miR-375, miR-455, miR-9
MDH2	81829	miR-23b
MECP2	29386	miR-212
PKM2	25630	miR-133a, miR-133b, miR-326
PRKCH	81749	miR-151, miR-200a, miR-216a, miR-99a
SCD2	83792	miR-195
SYT4	64440	let-7a, let-7b, let-7c, let-7d, let-7e, let-7f, let-7i, miR-100, miR-101a, miR-101b, miR-124, miR-125a-5p, miR-129, miR-130a, miR-132, miR-138, miR-139-5p, miR-150, miR-181b, miR-185, miR-204, miR-21, miR-214, miR-215, miR-218, miR-223, miR-23a, miR-23b, miR-24, miR-25, miR-26b, miR-27a, miR-290, miR-292-5p, miR-298, miR-29a, miR-29b, miR-29c, miR-301a, miR-30a, miR-30b-5p, miR-30c, miR-30e, miR-320, miR-324-5p, miR-327, miR-328, miR-331, miR-338, miR-342-5p, miR-34a, miR-374, miR-382, miR-383, miR-409-3p, miR-485, miR-494, miR-497, miR-543, miR-664, miR-7a, miR-9, miR-92a, miR-98, miR-99a, miR-99b
VEGFA	83785	miR-150, miR-184, miR-31
VIM	81818	let-7a, let-7b, let-7c, let-7d, let-7e, let-7f, let-7i, miR-100, miR-101a, miR-101b, miR-124, miR-125a-5p, miR-129, miR-130a, miR-132, miR-138, miR-139-5p, miR-150, miR-181b, miR-185, miR-204, miR-21, miR-214, miR-215, miR-218, miR-223, miR-23a, miR-23b, miR-24, miR-25, miR-26b, miR-27a, miR-290, miR-292-5p, miR-298, miR-29a, miR-29b, miR-29c, miR-301a, miR-30a, miR-30b-5p, miR-30c, miR-30e, miR-320, miR-324-5p, miR-327, miR-328, miR-331, miR-338, miR-342-5p, miR-34a, miR-374, miR-382, miR-383, miR-409-3p, miR-485, miR-494, miR-497, miR-543, miR-664, miR-7a, miR-9, miR-92a, miR-98, miR-99a, miR-99b

Bold = miRNAs reported and experimentally validated to be involved in Regulating gene expression in the brain (in either human, mouse and/or rat). Blue = Long Term Potentiation (LTP) function. Underscore = Dendritic functions. Orange = No evidence for regulating gene expression in the brain.

Cis-acting elements associated with dendritic transcripts

As discussed above, mRNA transportation and regulation is mediated by post-transcriptional events involving both cis and trans-acting determinants [61, 62]. However, the mechanism and targeting specificity of these mRNAs is still unclear (reviewed in [63]). Motifs imbedded in the transcripts' non-coding regions, particularly in the 3' UTR, have been most closely implicated in mRNA transport processes [64, 65]. More recent investigations from the Eberwine and Kim labs have highlighted the importance of introns retention in dendritic mRNA transcripts [66-68]. I therefore asked if the selected dendritic transcripts, from this *in situ* hybridization study, might exhibit within their 3'UTR or intronic sequence enrichment for functional motifs compared to the general whole transcriptome.

ID elements retention favoring subcellular localization of transcripts in dendrites

Retained intronic SINE (Short Interspersed Repetitive Elements) retrotransposons, called ID elements, were shown in a recent study from our labs [67] to mediate dendritic localization of some mRNAs in rat neurons. ID elements are thought to derive from the BC1 RNA gene [69] and to have expanded in the rodent genome but at different rates depending on the species. The genomic copy numbers of ID elements are estimated to vary widely from ~300 copies in guinea pigs to over 150,000 copies in rats [70]. In collaboration with Mugdha

Khaladkar, Ph.D., we re-annotated ID element numbers in both mouse and rat genome based on computational sequence analysis detailed in the Materials and Methods section. Given the established implication of ID elements in dendritic targeting of some transcripts [67], I asked if the occurrence of ID elements in the subset of dendritic transcripts is higher than the average number in the whole transcriptome.

Our data showed a significantly higher incidence of ID elements per gene in mouse dendritic transcripts at a rate of 0.703 per dendritic gene versus 0.592 per gene from whole genome (Chi square pVal<0.005). Not surprisingly similar results were found in the rat transcripts. I found a number of 1.413 ID elements per gene in rat dendritic transcripts compared to 1.103 per gene from whole genome (Chi square pVal<0.001). These results are consistent with the concept of ID elements within retained intronic sequences for mediating, in part, rat dendritic localization as highlighted recently by P. Buckley and colleagues [67]. It is worth noticing that our search focused particularly on the ID elements detected in the sense DNA strand orientation, as these seem to be the most functionally relevant for transcript localization [67]. More than 1/3 of these sense ID elements were experimentally recovered in both rat and mouse dendrites via RNAseq separately assayed in our labs. Theses results not only reinforce previous findings but also highlight that the functional effect of ID element in targeting subcellular transcripts may occur in both mice and rats.

Cis-acting motifs regulating transcripts localization and gene expression in dendrites exhibit a high prevalence of nucleotide repetition in the UTR region.

Cis-acting "zipcode" elements have been implicated in mRNA localization, stabilization, and regulation of local translation. The prediction of these motifs has been very challenging as it is largely driven by the mRNA secondary structure rather than its primary nucleotide sequence. None of the motif-finding algorithms developed thus far have proven fully efficient. While the identity of all the *cis*-factors is still unknown, some have been clearly identified. Table 1.1 in the Chapter 1 highlights a few of the *cis*-factors reported to be involved in mRNA sub-cellular localization. At the gene expression regulation level, sequence elements rich in adenosine and uridine, called AU-rich elements (AREs) in mammalian cells, are known to affect mRNAs stability and target them for rapid degradation, thereby regulating their translational rate [71, 72]. Particularly in neurons, the regulation of mRNA stability and its translation are key components to allow fast and efficient cellular responses to incoming stimuli.

Using the NIH (http://lgsun.grc.nia.nih.gov/mRNA/) database, I searched for the occurrence of some motifs well known for their role in gene expression regulation: AREs (ARE1-ARE4) and CpG in the 3'UTR and 5'UTR regions. By comparing how often these motifs appeared in mouse dendritic transcripts to the background of the whole mouse transcriptome, I recovered a significant enrichment in the dendritic transcripts (Chi-Square test with FDR<0.1) for the following motifs: 5'UTR-CpG, 3'UTR-ARE4 and 3'UTR-CpG. These results

support earlier findings on the implication of these motifs in gene expression regulation [71, 73, 74] and suggest that further investigations on dendritic mRNA regulation via AREs and/or CpG motifs could shed light on important function regulation in the brain. Particularly, most studies on CpG repeats have mainly focused on functionality in the 5'UTR [75, 76], whereas our study suggests that the 3'UTR region could be equally important. In fact, further studies on 3'UTR-CpGs could potentially link this region to transcripts trafficking or microRNAs regulation as was suggested in recent investigations [77-79].

3.3.5 Comparative analysis of the subcellular localization of mRNA transcripts in rat and mouse via *in situ* hybridization

Our previous investigation via microarrays, detailed in Chapter 2, reported a high degree of evolutionary divergence (81%) between rat and mouse dendritic transcriptomes. What causes these differences in transcript localization between species? Potential answers could be based upon gene specific sequences, the different cellular functions of these genes, or simply the result of divergent targeting mechanisms used by one set of transcripts but not others.
Here, our *in situ* hybridization studies on some of these divergent genes may illuminate some of these points, parallel our previous findings in Chapter 2 and highlight species-specific differences in subcellular localization. The outcome of our *in situ* assays could also stress the importance of divergence in either *cis*- and/or *trans*-acting factors in determining species-specific neurons function

differences. To maintain consistency in this study, I followed the same analytical framework for the across species comparison as the one I used in the previous sections based on within species investigations.

Transcripts in rat and mouse dendrites exhibit species-specific localization.

Although my *in situ* hybridization study included 362 transcripts in rat and mouse neurons, I narrowed the between-species comparative study to 55% of the probes (n=200). The curated set of probes corresponds to known rat and mouse orthologous genes for which the *in situ* probes were sequence verified to have a high level of gene specificity in both species. Within this subset, 84.5% (n=169) of the transcripts resulted in significant dendritic localization either in rat or mouse or both species' dendrites. The following analysis will focus entirely on this selected group. Of these 169 transcripts, a majority of 58% (n=98) showed a significant dendritic localization in both rat and mouse, as their D/S ratio was above the rat and mouse control cut-off values respectively. To note is that the mean expression level of these shared targeted transcripts was significantly different ($p<0.0001$ with t-test), suggesting that the difference in gene expression is likely due to species-specific active RNA transport. Of the remaining differentially targeted transcripts, 13% (n=22) were specifically localized to rat dendrites and 29% (n=49) were specifically localized to mouse dendrites. The Venn diagram in Figure 3.6 illustrates this distribution.

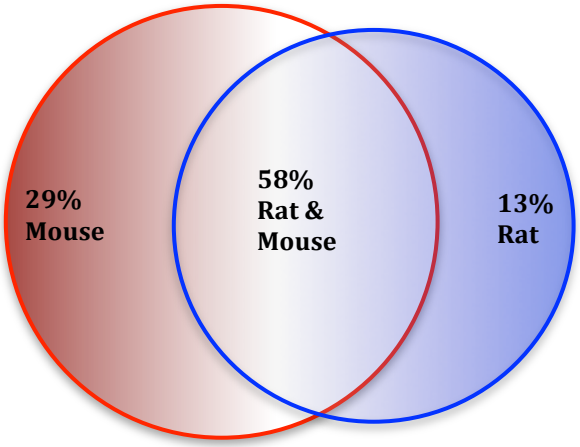


Figure 3.6: Venn diagram for distribution of transcripts investigated in rat and mouse dendrites.

Figure 3.7 displays the D/S ratio detected for all transcripts examined via *in situ* hybridization on rat and mouse neurons. These transcripts are displayed based on the sorting of the differential value between rat and mouse D/S ratio. This type of representation (Figure 3.7) allows visual differentiation of the cluster of transcripts with similar level of expression in rat and mouse dendrites (middle section of the graph, Figure 3.7) from the clusters with higher level of dendritic gene expression in one species versus the other (extreme left and right sections of the graph, Figure 3.7). It is worth noticing that certain gene families have some of members in one cluster and other members in another cluster. For instance in for Kinesin motor protein genes Kif17 and Kif15, the Kif17 exhibits a high level of expression in mouse dendrites only and the Kif15 exhibits a much higher (two-

fold) level of expression in rat dendrites compared to mouse dendrites. The Myosin motor protein genes Myo5a and Myo5b illustrate also this gene expression divergence between rat and mouse. Moreover, even within a group of transcripts with a high dendritic expression in both species, significant variations exist: e.g., Kif15 has ~2 times higher level of expression in rat compared to mouse dendrites while the potassium channel Kcnj3 exhibits a level of expression ~2 times higher in mouse compared to rat dendrites. Overall, the diversity of gene expression in rat and mouse dendrites suggests species-specific differences in mRNA transport mechanism, regulation or both. The detailed list of transcripts considered dendritically localized in rat and mouse, rat only or mouse only is available in Table 3.10 (located at the end of this chapter).

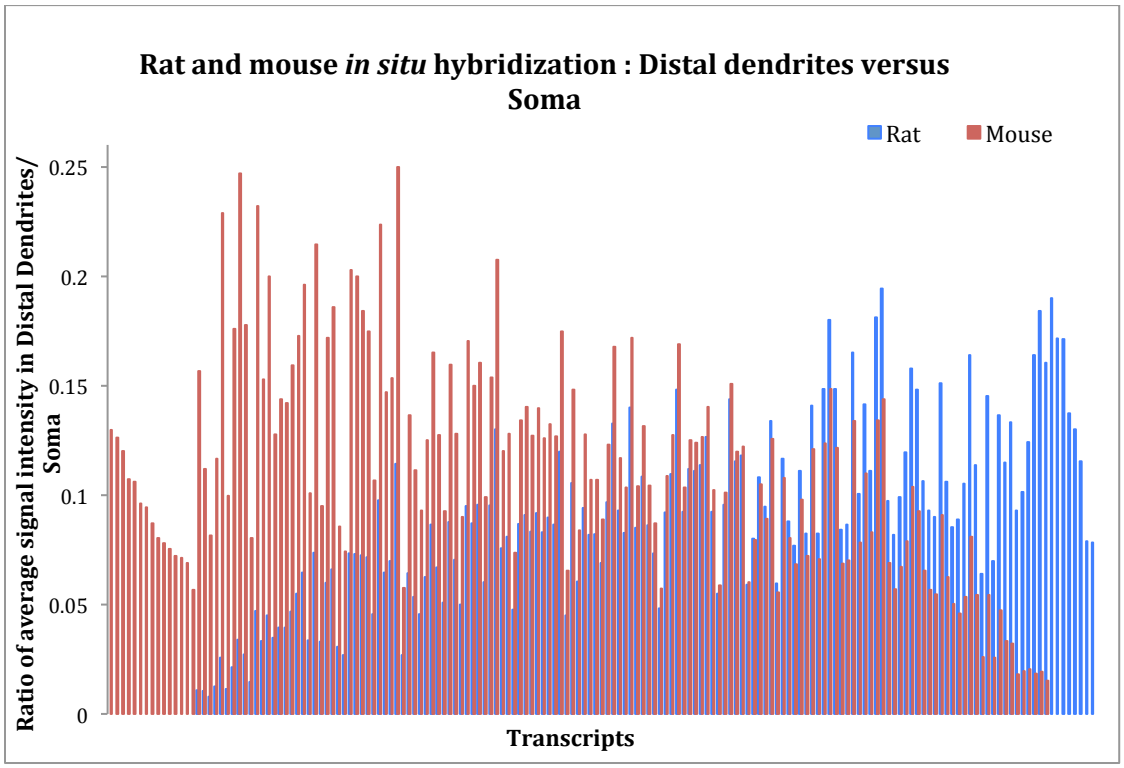


Figure 3.7: Distribution of the average signal intensity ratios of distal dendrites to soma in 169 transcripts examined via *in situ* hybridization in rat and mouse neurons.

Figure 3.8 visually illustrates, via *in situ* hybridization on rat and mouse neuronal cells, subcellular localization differences in three selected genes. For example, *Dpsysl2* (Fig. 3.8*A*) shows high signal in the cell soma of both the rat and mouse neurons, but high dendritic signal only in the rat neuron. In contrast, *Uba52* shows high dendritic signal in mouse but not rat (Fig. 3.8*B*), and *Atp2b2* shows consistently high dendritic signal in both mouse and rat (Fig. 3.8*C*).

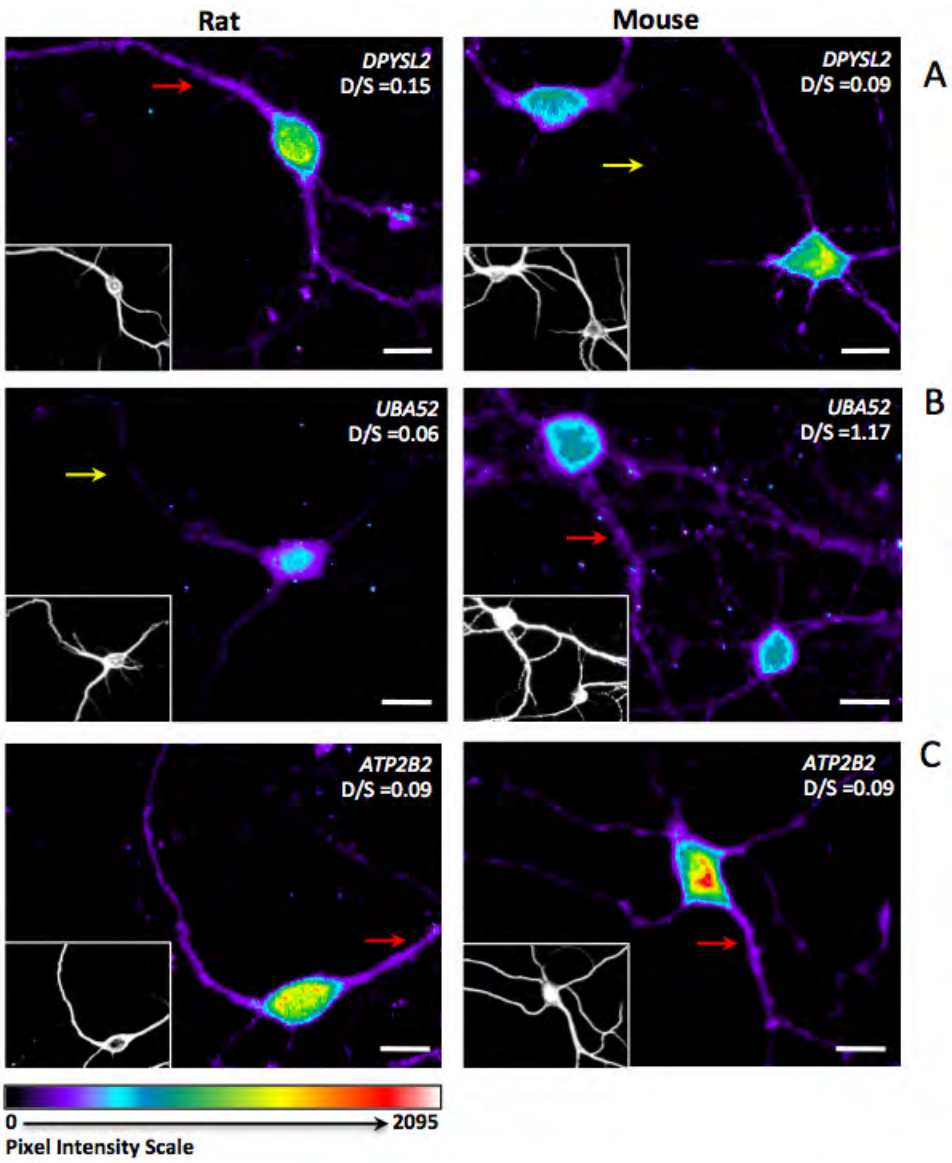


Figure 3.8: *In situ* hybridization shows inter-species differences in dendritic localization.

Fluorescent Microscopy evaluation of biotin-conjugated oligoprobes on paraformaldehyde fixed 14-day cultured rat and mouse cortical neurons hybridized with 3 biotin-conjugated oligoprobes detected with streptadivin-Alexa568. For each probe images set, the small bottom left corner panels represent MAP2 immuno-staining. Scale bars = 20µm.
(A), Probe against *Dpsysl2* transcript shows higher dendritic localization in rat neurons (Red arrow) than in mouse neurons (Yellow arrow).
(B), Probe against *Uba52* transcript shows higher dendritic localization in mouse neurons (Red arrow) than in rat neurons (Yellow arrow).
(C), Probe against *Atp2b2* transcript shows high dendritic localization in both rat and mouse neurons (Red arrows).

Functional annotation of dendritic transcripts examined in rat and mouse

In order to get an overview of general cellular functions of the dendritic transcripts investigated in both rat and mouse I performed, as previously, a Gene Ontology (GO) analysis on the comparative dendritic transcriptome of rat and mouse neurons (Figure 3.9) using DAVID [15]. In the subset of transcripts dendritically localized in rat and mouse (Figure 3.9A), I found the main functional categories as those reported in previous sections, such as metabolism, cell organization and transport. These were significantly enriched compare to both rat and mouse genomic background (Chi square test, pVal <0.0001 for Metabolism and pVal <0.02 for both the Cell Organization and the Transport functional categories). Given their common dendritic localization, I hypothesize that the function of these dendritic transcripts is conserved across species and therefore that they play a fundamental role in mammalian neurons. The functions of these particular genes are also likely to be evolutionarily conserved in other mammalian brains (i.e. *Nefl*, known to be involved in Parkinson's disease [80], or *Atp2b2*, known to be involved in ataxia in humans [81]) and therefore these could be good candidates for further investigation in primates.

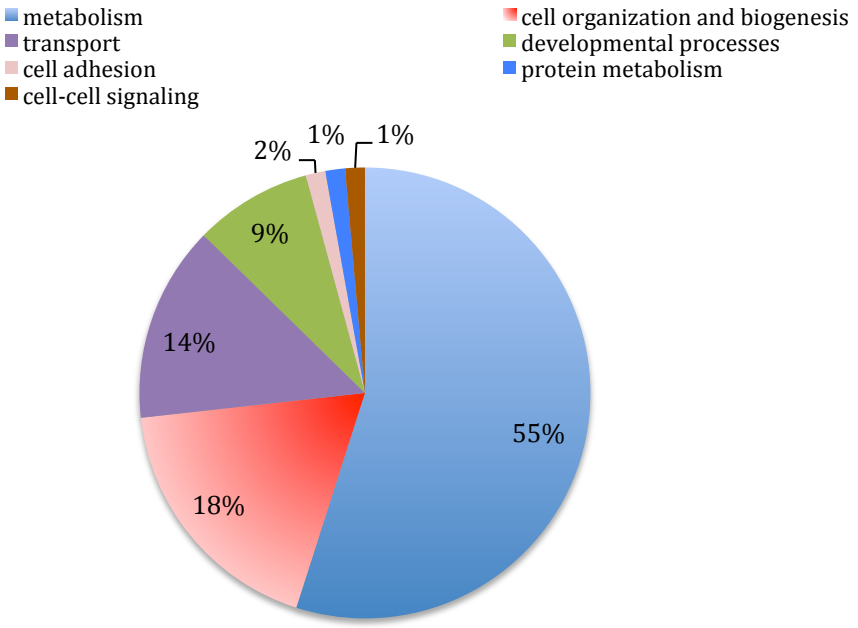
Other relevant GO categories involving signal transduction, apoptosis and RNA metabolism were distinguished within the species-specific dendritic transcripts. Transcripts involved in signal transduction were uniquely detected in mouse dendrites (Figure 3.9B), and many of these such as *Gnas*, *Rab14*, *Rhobtb1* and *Usp8* were shown to play a key role in synaptic plasticity [82-84].

Transcripts involved in regulating apoptosis (i.e. *Pabpn1*) and RNA processing and metabolism (like *Aars*, *Mecp2* and *Rsp6*) were more specific to rat dendrites (Figure 3.9C).

The differences seen between these two rodents in the nature and functions of some dendritic transcripts reinforce the idea that species-specific subcellular localization might be a key source behind species differences in brain's ability such as learning and memory.

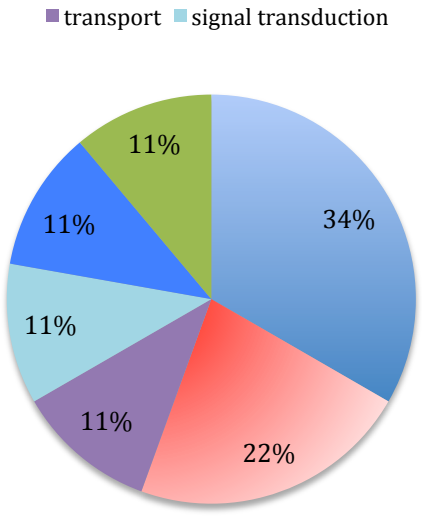
A)

Functional Categories: Transcripts commonly localized to rat and mouse dendrites (n=98).



B)

Functional Categories: Transcripts unique to mouse dendrites (n=49).



C)

Functional Categories: Transcripts unique to rat dendrites (n=22).

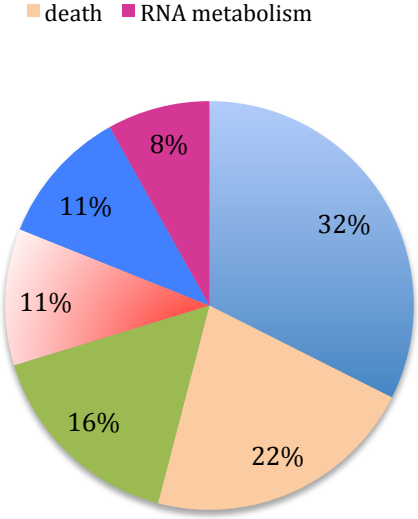


Figure 3.9: Functional categories for rat and mouse dendritic transcripts

Transcript coverage within dendrites: conserved and divergent mechanisms across species.

The diverse morphologies of dendritic branches influences how different neurons receive, filter, and consolidate electrical signals such as action potential propagation and information processing [85]. Thus variability of neuronal morphology clearly affects both the connectivity and the activity of the nervous system. Even though considerable progress has been made in identifying [86, 87] specific genes acting on dendritic morphogenesis in Drosophila, the exact molecular mechanisms controlling these processes remain poorly understood particularly in mammals. Do different RNAs localize to different branches of dendrites? If so, are they conserved across species? A comparative analysis of the differential distribution of individual transcripts within primary, secondary, and/or tertiary branches of dendrites, as illustrated in Figure 3.10, would provide some insight into the evolutionary conservation of certain genes or functions and their impact on neurogenesis and synaptic regulation.

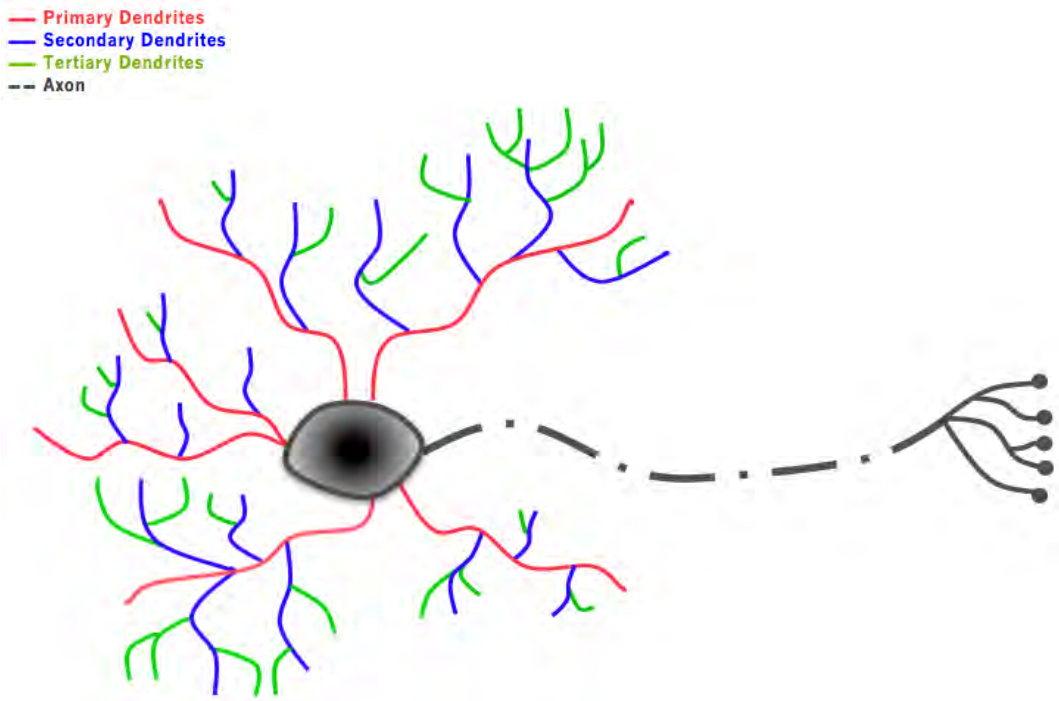


Figure 3.10: Illustration of primary, secondary and tertiary dendrites in a neuron.

To investigate the possible link between differential RNA localization within dendrites compartments and dendritic function, I categorized the occurrence of transcripts in primary, secondary and tertiary dendrites (see Figure 3.10). I investigated 29 randomly selected transcripts known to localize in rat and mouse dendrites and to belong to the main functional categories described in Figure 3.9A. Table 3.11 reports the detailed outcome of this screen. Overall 93% of the transcripts investigated showed a distal localization in at least the secondary dendrite branches. In both rat and mouse the transcripts mostly recovered in farther dendritic branches (i.e. tertiary branches) are involved in developmental processes (such as Dpysl2, Pou4f2, Lhx3 and Sirt1) and cell-cell adhesion (such as Pcdh17, Pcdh10 and Cdh2). This between-species

conservation of transcripts distribution in dendrites in those functional categories (75% conservation between rat and mouse) is not surprising as the involved transcripts are crucial for neuron growth and proper function in the brain [88-93]. On the other hand, the transcripts with the most divergence in their level of distribution in dendrites (i.e. localized to tertiary dendrites in one species but only to primary or secondary dendrites in the other, and vice-versa) belong to functions more amenable to species adaptation, such as metabolism and transport [94-96]. Additionally, I did not detect any significant nucleotide size difference between the transcripts differentially distributed within the dendrites (primary versus secondary versus tertiary) nor between rat and mouse dendrites.

Thus the observed mRNA dendritic compartment localization differences between rat and mouse suggest that either each species is using a different mRNA targeting mechanism, or that there are species-specific differences in the translation hot spots sites. These could define the scope of extended targeting in dendritic branches of a given transcript.

The differences in transcripts localization in dendrite branches might generate differences, both between and within species, in the way synaptic inputs are integrated and processed in the neurons [85], which might have a great effect on the firing properties of neurons and potentially could impact synaptic plasticity [97].

Table 3.11: Some rat and mouse transcripts show similar localization in dendrites while others show a divergent localization.

Functional Categories	Functional Description	Gene Symbol	Gene Description	Branching Rat dendrites	Branching Mouse dendrites
Cell organization	Motor Protein	*Myh2*	Myosin Heavy Chain II a	3	3
	Motor Protein	*Dynll1*	Dynein Light Chain LC8-Type 1	2	2
Cell-cell adhesion	Cadherin	*Pcdh17*	Protocadherin 17	3	3
	Cadherin	*Pcdh10*	Protocadherin 10	3	3
	Cadherin	*Cdh2*	Cadherin 2	2	3
Transport	MAPK signaling Pathway	*Mapk1*	Mitogen Activated Protein Kinase 1	3	3
	Motor Protein	*Kif15*	Kinesin Family Member 15	2	2
	Synaptic vesicle	*Syngr1*	Synaptogyrin 1	2	2
	Calcium ion binding	*Syt6*	Synaptotagmin VI	2	2
	Calcium ion binding	*Syt4*	Synaptotagmin IV	2	3
	WD40 repeat region	*Sec13*	SEC13 Homolog	2	3
Cell-Cell signaling	G-protein Complex	*Gnbpb1*	Guanine Nucleotide Binding Protein Beta1	3	3
	G-protein Complex	*Gnbpa*	Guanine Nucleotide Binding Protein Alpha0	2	2
	Intracellular signaling	*Fgf13*	Fibroblast Growth Factor 13	2	3
Developmental processes	Neurogenesis	*Dpysl2*	Dihydropyrimidinase-Like 2	3	2
	Neurogenesis	*Pou4f2*	POU Domain, Class 4, Transcription Factor 2	3	3
	Transcription Regulation	*Lhx3*	LIM Homeobox Protein 3	3	3
	Transcription Regulation	*Sirt1*	Silent Mating Type Information Regulation 1	3	3
	Transcription Regulation	*Atf3*	Activating Transcription Factor 3	1	1
Metabolism	Oxydative Phosphorylation	*Cox5b*	Cytochrome C Oxidase, Subunit Vb	2	3
	Oxydative Phosphorylation	*Atp5d*	ATP Synthase, Delta	2	2
	Pyruvate metabolism	*Pklr*	Pyruvate Kinase Liver And Red Blood Cell	2	3
	Pyruvate metabolism	*Mdh2*	Malate Dehydrogenase 2, NAD	2	3
Protein metabolism	Methylation	*Uba52*	Ubiquitin A-52 Residue Ribosomal Protein	2	2
	Translation elongation	*Rpl23*	Ribosomal Protein L23	2	2
	Translation elongation	*Rps15*	Ribosomal Protein S15	2	2
	Translation elongation	*Rpl6*	Ribosomal Protein L6	2	3
	Translation elongation	*Rspa*	Similar To 40S Ribosomal Protein SA	1	1

1	mRNA localized to primary dendrites
2	mRNA localized to primary and secondary dendrites
3	mRNA localized to primary, secondary and tertiary dendrites

Trans-acting factors and transcripts localization in rat and mouse: differences and similarity

Our above analysis, summarized in Tables 3.5 and 3.6 for mouse and rat respectively, highlighted the possible implication of different pools of RBPs depending on the transcript distribution patterns in mouse and rat dendrites. Some RBPs were particularly related to the punctate dendritic transcripts, while others were more specific to uniformly distributed transcripts. In general, the assessed RBPs have been linked to neuronal gene expression, mainly due to their role in mRNA transport and editing mechanisms [43, 98]. Comparing these two rodents, the RBPs SNRPA, A2BP1 and ZFP36 were represented regardless of the pattern of transcript distribution in dendrites. Additionally, for both species, whereas SRFS1 and ELAVL2 were particularly prevalent in the class of punctate transcripts, SRFS2 and KHSRP were more frequently associated with the uniformly distributed class of transcripts. However, a closer comparison between Tables 3.5 and 3.6 shows that the identities of these RBPs do not perfectly match between mouse and rat, suggesting that there might be a species-specific mechanism for regulating transcripts localization. In order to assess whether for homologous transcripts, a correlation exists between patterns of distribution and RBPs, I performed a more detailed comparison between rat and mouse dendritic transcripts (n=32 mRNAs, corresponding to the intersection between mouse and rat Tables 3.3 and 3.4). These are listed in Table 3.12. Overall, mouse transcripts exhibited a higher incidence of punctate patterns than uniform spatial distribution

with a 3:1 ratio while the opposite was detected in rat with a 1:3 ratio for punctate versus uniformly-distributed transcripts.

Table 3.12 List of rat and mouse transcripts examined for distribution patterns in dendrites.

Gene Symbol	MouseRefseq	RatRefSeq	Distribution in Mouse	Distribution in Rat
Pcdh17	NM_001013753	XM_224389.4	Uniform	Uniform
Glp2r	NM_175681	NM_021848.1	Uniform	Uniform
Vim	NM_011701	NM_031140.1	Uniform	Uniform
Sirt1	NM_001159589	XM_228146.4	Uniform	Uniform
Syt6	NM_018800	NM_022191.1	Uniform	Uniform
Arl6ip5	NM_022992	NM_023972.2	Uniform	Uniform
Olfm1	NM_019498	NM_053573.1	Uniform	Uniform
Myo5a	NM_010864	NM_022178.1	Uniform	Uniform
H13	NM_010376	NM_001107789.1	Uniform	Punctate
Syt4	NM_009308	NM_031693.1	Uniform	Punctate
Usp9x	NM_009481	XM_343766.3	Punctate	Punctate
Opa1	NM_133752	NM_133585.2	Punctate	Punctate
Dnajc5	NM_016775	NM_024161.2	Punctate	Punctate
Gnb1	NM_010312	NM_030987.1	Punctate	Punctate
Vmn2r57	NM_177764	NM_173130.1	Punctate	Punctate
Jub	NM_010590	NM_053503.1	Punctate	Punctate
Tmeff1	NM_021436	NM_023020.1	Punctate	Punctate
Fmnl1	NM_001077698	XM_001081542.1	Punctate	Punctate
Rpl23	XM_001477371	NM_001007599.1	Punctate	Uniform
Kif15	NM_010620	NM_181635.2	Punctate	Uniform
Snrpn	XM_001480670	NM_031117.1	Punctate	Uniform
Arhgef7	NM_017402	NM_053740.1	Punctate	Uniform
Atp2b2	NM_009723	NM_012508.3	Punctate	Uniform
Dync1i1	NM_010063	NM_019234.1	Punctate	Uniform
Sept9	NM_001113488	NM_176856.1	Punctate	Uniform
Cacng5	NM_080644	NM_080693.1	Punctate	Uniform
Lypla1	NM_008866	NM_013006.1	Punctate	Uniform
Nefl	NM_010910	NM_031783.1	Punctate	Uniform
Gnao1	NM_010308	NM_017327.1	Punctate	Uniform
Rpl6	NM_011290	NM_053971.1	Punctate	Uniform
H2afz	XM_001480384	NM_022674.1	Punctate	Uniform
Atp5a1	NM_007505	NM_023093.1	Punctate	Uniform

Table 3.13 summarizes of the different distribution patterns per transcript and highlights the heterogeneity seen in these patterns between rat and mouse.

Table 3.13: Distribution patterns in dendrites for homologous transcripts in rat and mouse.

mRNAs distribution patterns	Uniform in mouse	Punctate in mouse
Uniform in rat	25%	41%
Punctate in rat	6%	28%

A more detailed comparative list of rat and mouse transcripts investigated via the RBPdatabase and their resulting affinity for RBPs is provided in Table 3.14. This table also displays, in both species, the distribution patterns (i.e. uniform versus punctate) detected for a given transcript as well as its corresponding nucleotide size (bp). Even though a clear heterogeneity exists in rat and mouse distribution patterns, a mutual preference for certain type of RBPs was evidently seen in both rodents; close to 60% (19/32) of the transcripts were associated with at least one exact same RBP. The most frequently represented RBPs in both species were ZPF36, ELAVL2 and SNPRA. The remaining ~40% of transcripts were associated with different RBPs in rat and mouse. The fact homologous RBPs seem to be associated with most dendritic transcripts in a non-species-specific fashion underscores the evolutionary conservation of these proteins' putative functions in targeting and locally regulating mRNA in the neurons' processes. For homologous transcripts in rat and mouse dendrites, I could not distinguish any obvious relationship between distribution patterns (i.e., uniform versus punctuated), nucleotide size and the nature of the interacting

RBPs, which suggests that complex interactions and regulatory events exist between mRNAs, RBPs and other *trans*-acting factors such as microRNAs [60, 99, 100].

Table 3.14 Relationship between transcripts and their affinity for RBPs in rat and mouse dendrites.

Gene	Pattern		Transcript Size (bp)			ZFP36	ELAVL2	SNRPA	A2BP1	SFRS1	SFRS2	SFRS7	pum	ZRANB2	PABPC1	Rna15	KHSRP	EIF4B	NONO	QKI	ybx2-a	NOVA2	NCL	sap-49	sus
	Ms	Rt	Ms	Rt	Diff	Ms Rt	Ms Rt	Ms Rt	Ms Rt	Ms Rt	Ms Rt	Ms Rt	Ms Rt	Ms Rt	Ms Rt	Ms Rt	Ms Rt	Ms Rt	Ms Rt	Ms Rt	Ms Rt	Ms Rt	Ms Rt	Ms Rt	Ms Rt
Lypla1	P	U	2447	2356	91																				
Gnao1	P	U	2968	2068	900																				
Vmn2r57	P	P	2592	2890	298																				
Sirt1	U	U	3400	640	2760																				
Dync1i1	P	U	2718	2658	60																				
Snrpn	P	U	2076	1428	648																				
Arhgef7	P	U	4665	4379	286																				
Kif15	P	U	4840	4214	626																				
Syt4	U	P	3901	3877	24																				
Opa1	P	P	5948	3200	2748																				
Dnajc5	P	P	3674	937	2737																				
H13	U	P	5202	1392	3810																				
Pcdh17	U	U	9509	4661	4848																				
Jub	P	P	3503	3261	242																				
Gnb1	P	P	3143	1544	1599																				
Sept9	P	U	3796	3873	77																				
Usp9x	P	P	11903	12133	230																				
Tmeff1	P	P	2442	2506	64																				
Atp5a1	P	U	2443	1871	572																				
Myo5a	U	U	11684	5620	6064																				
Atp2b2	P	U	6922	4228	2694																				
Glp2r	U	U	4873	1653	3220																				
Vim	U	U	1834	1796	38																				
Olfm1	U	U	1048	2759	1711																				
H2afz	P	U	957	811	146																				
Cacng5	P	U	3678	828	2850																				
Nefl	P	U	2014	2032	18																				
Arl6ip5	U	U	1339	567	772																				
Fmnl1	P	P	3774	3659	115																				
Rpl23	P	U	1064	522	542																				
Rpl6	P	U	1282	962	320																				
Syt6	U	U	4416	1850	2566																				

Legend

	Mouse (Ms)
	Rat (Rt)
	Rat & Mouse
P	**Punctate**
U	**Uniform**
Diff	**Mouse - Rat transcript size difference (bp)**

MicroRNAs are likely involved in gene expression regulation in rodent dendrites

miRNAs are involved in diverse aspects of development, maintenance, and disease, and are largely evolutionarily conserved in mammals. Genome-wide studies on various organisms including viruses, plants, worms to humans have revealed thousands of miRNAs. These were shown to be involved in different aspects of development, cellular functions, and disease, and are largely evolutionarily conserved in mammals. Particularly hundreds of conserved brain-expressed miRNAs in both mouse and human have been implicated in mammalian neuronal development or function. Here, I assessed the evolutionary pattern of miRNAs that are putatively associated with our dendritic transcripts.

The within-species examination of miRNAs associated with dendritic transcripts revealed close to a hundred experimentally validated miRNAs (Tables 3.8 and 3.9). These were reported to post-transcriptionally regulate 12% (n=19) and 10% (n=16) of the transcripts investigated here in mouse and rat dendrites respectively [56]. At the between species level, 11% (19 out of 169) of the transcripts that were investigated in both mouse and rat via *in situ* hybridization were linked to 72 known miRNAs in mouse and/or rat as detailed in Table 3.15.

Table 3.15: List of transcripts investigated in rat and mouse and their microRNAs

Gene Symbol	MicroRNA Name
Jub	miR-125a-3p, miR-125a-5p, miR-125b, miR-125b-3p, miR-125b-5p
Lypla1	miR-138
Mapk1	let-7a, let-7b, let-7c, let-7d, let-7e, let-7f, let-7i, miR-100, miR-101a, miR-101b, miR-124, miR-125a-5p, miR-129, miR-130a, miR-132, miR-138, miR-139-5p, miR-150, miR-181b, miR-185, miR-204, miR-21, miR-214, miR-215, miR-218, miR-223, miR-23a, miR-23b, miR-24, miR-25, miR-26b, miR-27a, miR-290, miR-292-5p, miR-298, miR-29a, miR-29b, miR-29c, miR-301a, miR-30a, miR-30b-5p, miR-30c, miR-30e, miR-320, miR-324-5p, miR-327, miR-328, miR-331, miR-338, miR-342-5p, miR-34a, miR-374, miR-382, miR-383, miR-409-3p, miR-485, miR-494, miR-497, miR-543, miR-664, miR-7a, miR-9, miR-92a, miR-98, miR-99a, miR-99b
Mdh2	miR-23b
Syt4	let-7a, let-7b, let-7c, let-7d, let-7e, let-7f, let-7i, miR-100, miR-101a, miR-101b, miR-124, miR-125a-5p, miR-129, miR-130a, miR-132, miR-138, miR-139-5p, miR-150, miR-181b, miR-185, miR-204, miR-21, miR-214, miR-215, miR-218, miR-223, miR-23a, miR-23b, miR-24, miR-25, miR-26b, miR-27a, miR-290, miR-292-5p, miR-298, miR-29a, miR-29b, miR-29c, miR-301a, miR-30a, miR-30b-5p, miR-30c, miR-30e, miR-320, miR-324-5p, miR-327, miR-328, miR-331, miR-338, miR-342-5p, miR-34a, miR-374, miR-382, miR-383, miR-409-3p, miR-485, miR-494, miR-497, miR-543, miR-664, miR-7a, miR-9, miR-92a, miR-98, miR-99a, miR-99b
Vim	let-7a, let-7b, let-7c, let-7d, let-7e, let-7f, let-7i, miR-100, miR-101a, miR-101b, miR-124, miR-125a-5p, miR-129, miR-130a, miR-132, miR-138, miR-139-5p, miR-150, miR-181b, miR-185, miR-204, miR-21, miR-214, miR-215, miR-218, miR-223, miR-23a, miR-23b, miR-24, miR-25, miR-26b, miR-27a, miR-290, miR-292-5p, miR-298, miR-29a, miR-29b, miR-29c, miR-301a, miR-30a, miR-30b-5p, miR-30c, miR-30e, miR-320, miR-324-5p, miR-327, miR-328, miR-331, miR-338, miR-342-5p, miR-34a, miR-374, miR-382, miR-383, miR-409-3p, miR-485, miR-494, miR-497, miR-543, miR-664, miR-7a, miR-9, miR-92a, miR-98, miR-99a, miR-99b
Atf3	miR-200b,miR-214
Basp1	miR-9
Cdh2	miR-141,miR-155,miR-200a,miR-200b,miR-200c,miR-203,miR-205,miR-23b,miR-29c,miR-369-5p,miR-370,miR-375,miR-429,miR-542-5p
Irgm	miR-196a,miR-196b
Lhx3	miR-17,miR-181b
Mtap2	miR-17,miR-181b,miR-26a
Pou4f2	miR-1,miR-130a,miR-19a,miR-23a,miR-23b,miR-26a,miR-34a
Prph1	miR-1,miR-183,miR-96
Sirt1	miR-134,miR-199a-5p,miR-34a,miR-9
Vsnl1	let-7a,let-7b,let-7c,let-7d,let-7e,let-7f,let-7g,let-7i,miR-100,miR-101a,miR-101b,miR-124,miR-125a-5p,miR-126-3p,miR-126-5p,miR-130a,miR-132,miR-138,miR-139-5p,miR-150,miR-181b,miR-185,miR-196a,miR-204,miR-21,miR-214,miR-215,miR-218,miR-223,miR-23a,miR-23b,miR-24,miR-25,miR-26b,miR-27a,miR-290-3p,miR-290-5p,miR-292-5p,miR-298,miR-29a,miR-29b,miR-29c,miR-301a,miR-30a,miR-30c,miR-30e,miR-31,miR-320,miR-324-5p,miR-327,miR-328,miR-331-3p,miR-331-5p,miR-338-3p,miR-338-5p,miR-342-5p,miR-34a,miR-374,miR-382,miR-383,miR-409-3p,miR-485,miR-494,miR-497,miR-543,miR-7a,miR-9,miR-92a,miR-98,miR-99a,miR-99b
Mecp2	miR-212
Prkch	miR-151, miR-200a, miR-216a, miR-99a
Scd2	miR-195

Black = MicroRNAs experimentally validated in rat and mouse*
Red = MicroRNAs experimentally validated in mouse*
Blue = MicroRNAs experimentally validated in rat*
* Data extracted from miRWalk (http://www.ma.uni-heidelberg.de/apps/zmf/mirwalk/index.html)

Close to 74% of the transcripts associated with microRNAs were dendritically localized in both rat and mouse. The exhaustive repartition of the 19 transcripts investigated in rat and mouse and potentially associated with microRNAs is displayed in the Table 3.16 below.

Table 3.16: Repartition of transcripts investigated, via *in situ* hybridization, in rat and mouse neurons (n=19) with known microRNAs association.

	Rat only	Mouse only	Rat and mouse
Proportion of mRNA localized to dendrites and associated with microRNAs	10.5%	16.8%	73.7%

MicroRNAs, including the 72 reported above, may be sequence conserved between species but their mode of expression and mRNA targets can diverge through evolution. For instance, two microRNAs specifically expressed in the mammalian nervous system: miR-9 and miR-124 share 100% of their nucleotide sequences identity among many species, but their expression level and their mRNA targets are not all identical between species [101]. In our data, miR-9 and miR-124 seemed associated in both rodents to targeting and regulating Mapk1, Vim and Syt4 transcripts, but seemed associated with additional transcripts only in mouse: miR-9 were additionally linked to Sirt1 and Basp1 and miR-124 were linked to Vsnl1. The differences in the mRNAs targets for these same miRNAs suggest that these might exhibit, in a species-dependent fashion, diverse post-transcriptional regulation events which could impact many key neuronal functions such as synaptic plasticity, neurogenesis or dendrites morphogenesis. Indeed,

several investigations have shown that specific microRNAs functions vary in a context-dependent manner and/or on depending on the organism and its developmental stage [101-103]. For instance, miR-9 exhibits a high level of functional specialization across vertebrates and is spatially restricted to different areas of central nervous system in Xenopus [104], zebra fish, chicken and mouse [105-107]. Thus gene expression regulation by miRNAs might contribute uniquely to the evolution of the complex nervous system.

Retained ID elements act as *Cis*-elements in the subcellular localization of transcripts to rat and mouse dendrites

Our microarray analysis on mechanically dissected rat and mouse dendritic transcripts, detailed in Chapter 2, showed a significant enrichment of the intron-retained ID elements in the localized transcripts of both species. Similarly, the above *in situ* hybridization study on *cis*-elements involved in dendritic subcellular localization of transcripts within rat and mouse independently, pointed towards a significant enrichment of ID elements. Hence, I postulated that this enrichment should also be detected in the subset of transcripts shown to be localized in both rat and mouse dendrites via *in situ* hybridization because these species may share ID-dependent targeting mechanisms.

Similar to what was uncovered in my previous analysis, I found that ID elements occur at significantly higher rates among those transcripts commonly

targeted to rat and mouse dendrites when compared to the entire transcriptome of each species. More specifically, I found 0.65 ID elements per mouse dendritic gene compared to 0.59 ID elements per gene in the mouse genome (Chi square pVal<0.08) and 1.5 ID elements per rat dendritic gene compared to 1.1 ID elements per gene in the rat genome (Chi square pVal<0.001). Additionally, the majority of dendritic transcripts with retained ID elements were homologous with 84% (21 out of 26) and 58.3% (21 out of 36) in mouse and rat respectively. These are mainly involved in neurons processes growth, cell-cell adhesion, synaptic vesicles and ion binding. The remaining dendritic transcripts with retained ID elements, 16% (5 out of 26) and 41.7% (15 out of 36) in mouse and rat respectively, were more species-specific (i.e. non-homologous between rat and mouse) and relayed mainly to metabolic, transport and signal transduction functions.

For dendrites gene expression levels, I did not detect any significant difference in levels neither within the clusters of ID elements retained transcripts homologous and non-homologous in rat and mouse, nor between these and the clusters of transcripts without ID elements retention. This finding suggests that the major role of ID elements retention could be to support transcripts localization in dendrites [67] but not necessarily to locally increase the level of transcripts gene expression.

3.4 DISCUSSION

The main purpose of this study was to survey, via *in situ* hybridization, patterns of mRNA localization in rat and mouse dendrites and to uncover factors that might contribute to this localization, including species-specific factors. I examined nearly 400 of the most representative dendrite specific genes in rat and mouse selected from our previous microarrays study (Chapter 2). The majority of these transcripts displayed, with high confidence, localization in dendrites. Image analysis results highlighted differences and similarities in transcripts localization between rat and mouse dendrites suggesting that species-specific mechanism could be acting on the dendritic transcripts subcellular localization and/or gene expression regulation.

3.4.1 *Cis* and *trans*-factors associated with sub-cellularly localized transcripts.

Our study supports previous findings on mRNA localization and gene expression regulation in dendrites [9, 11, 14, 18, 108-110] and extends them by adding novel evolutionary concepts to these mechanisms. For instance, not only did it reveal a high incidence of the introns-retained ID elements in rat dendritic transcripts as was reported previously in our labs [67], but also showed that this retention is widespread in mouse, which provides additional evidence for the functional relevance of these retrotransposons in transcripts targeting to dendrites thereby regulating their level of local expression. Our investigation also

highlighted the prevalence of other *cis*-elements involved in the stabilization and regulation of dendritic transcripts such as the 3'UTR ARE4 motif (A/U 12 mers non-specifically repeated) [26, 71, 111]. This finding supports J. Mattick 's proposition on the major role of RNA-based processes in promoting phenotypic variation and complexity in higher eukaryotes [112]. Moreover, both rat and mouse dendritic transcripts showed a significant connection with *trans*-acting RNA-binding proteins that are mainly implicated in alternative splicing, stability and regulation of translation such as the ZFP36, A2BP1 or ELAVL2 protein [39, 41, 43, 113]. Particularly ELAVL2, known to exhibit high binding affinity to ARE-containing mRNAs [111], was the RBP that interacted the most frequently with the dendritic transcripts. This result is consistent with our finding of high ARE4 incidence within the 3'UTR of the investigated dendritic transcripts (Chi-Square, FDR<0.09) [111] and supports the fact that ELAVL2 might play a key role in regulating the translation a majority of dendritic transcripts and might be essential for securing a tight control of gene expression in neurons.

Further computational investigations on a wider range of motifs including secondary or even tertiary structure motifs will most likely recover additional *cis*-elements, and will allow more advanced predictions of RNA-binding proteins affinities [114, 115]. These results will guide the design of future functional experiments that will validate the candidate *cis* or *trans*-elements predicted from these motifs. These experiments could allow testing of whether certain RBPs are associated with a specific group of functionally related mRNAs as was proposed by J. D. Keene [116]. An additional question that could be addressed is whether

a relationship exists between post-translational modifications of RBPs (i.e. palmitoylation, acetylation or phosphorylation), their mRNAs targets and microRNA-mediated translational repression. This might be the case of Sirt1 mRNA whose regulation depends upon the phosphorylation state of the RBP HuR [117] and the interaction with some miRNAs such as miR-134 [118]. Understanding the molecular functionality of these proteins could provide important insights into neurons differentiation, plasticity and signals' coordination and processing in the nervous system.

3.4.2 MicroRNAs may fine-tune the regulation of gene expression in subcellularly localized transcripts

Post-transcriptional regulatory events are required to control mRNA translational in the appropriate place and time. Transcripts, meant to be translated in dendrites, could be kept dormant in a microRNA-dependent fashion both during their trafficking and even after their delivery to the appropriate synapse. Upon stimulation, these silenced transcripts could get separated from the RISC complex in order to be translated. Recently, a study showed that Armitage, a RISC protein known to be involved in miRNA processing, was both detected in dendrites and degraded during memory formation [119]. In the past decade, more and more evidence links microRNAs (miRNA) to different neuronal functions ranging from early neurogenesis to synaptic plasticity, including miR-9, miR-26a, miR-124, miR125b, miR-128, miR-134, miR-137, miR-138, miR181b

and miR-218 to name a few [101, 120-122]. Our dendritically localized transcripts appeared to be linked to many validated brain-specific miRNAs (Table 3.14) such as miR16, recently involved in ARE-containing mRNAs decay [123]; miR-26a, reported to inhibit Map2 translation [124]; and, miR-181b, known to decrease neuronal outgrowth by interacting with GluR2 and Vsnl1 (Table 3.14) and to be upregulated in schizophrenia [125]. Additionally, our investigation highlighted a possible relationship between dendritically localized rat and mouse mRNA and experimentally-validated brain-specific miRNAs (Table 3.14). I speculate that some of these miRNAs, such as miR-9 and miR-124 [101], are likely involved in species-specific as well as transcript-specific post-transcriptional regulation, and play a major role in the proper development and evolution of complex nervous systems.

3.4.3 Evolution of mRNA subcellular localization

Our investigation underlined three main clusters of dendritic transcripts: one with a shared significant localization in both rat and mouse, one with a significant localization in mouse dendrites but not in rat and one with a significant localization in rat dendrites but not in mouse. Even within the cluster of transcripts that showed localization in the dendrites of both species, the quantified level of gene expression was significantly different between species (t-test, $p<0.0001$). This suggests that, even if mRNA localization is conserved at a broader level, divergence seems to occur between species at a finer scale. This

diversity in subcellular localization is likely to play a role in cellular functions and may contribute to the phenotypic divergence between these species including the difference in their learning and memory abilities [126-128]. Additionally, the divergence that was uncovered here between rat and mouse does not only seem to be associated with transcripts' presence in dendrites but also with their coverage and distribution within the dendritic branches. This suggests that multiple species-specific mechanisms and potentially specific combinations of transcript clusters could be involved in the calibration and coordination of these cellular events.

3.4.4 RNA localization and disease

Our data showed that subcellularly localized mRNAs are likely involved in various neuronal function including neurogenesis, vesicle trafficking, and cell-cell communication. It is possible that the integrity of the process of transcript localization and post-transcriptional regulation might be essential for ensuring appropriate neuronal cell growth, wiring, and brain function. Any dis-regulation of this process might have detrimental effects on the cell, or even the whole organism. Indeed, inappropriate targeting of mRNAs would lead to aberrant protein distributions within the cell, alteration of cell's normal composition and pathways. Support for this broad idea comes from previous studies, which show that several of the dendritically localized mRNAs in our study, such as *Atxn2*, *Nsf*, *Uchl1*, *Myo5a* or *Mecp2,* encode well-known regulators of neurogenesis and

synaptic function with reported relationship to neurological diseases such as Amyotrophic Lateral Sclerosis, Griscelli syndrome, Alzheimer, Parkinson and Huntington disease [129-137]. These examples highlight the potential utility of our database in identifying genes relevant to research into human disease. Additionally, as both rat and mouse *in situ* data will be available, this might guide researchers in their of choice of which model organism is more appropriate to address a given question. In sum, our findings suggest the possibility that some neurological disease pathways thought to be linked to gene mutations, protein aggregation and protein mis-folding, may in fact be due to RNA-protein interactions, RNA localization, and RNA post-transcriptional regulation.

3.4.5 Impacts of the *in situ* database usage

All of our results have been archived in a searchable relational database (http://kim.bio.upenn.edu/insitu/public/). One of the most useful tools to come out of our study is the *in situ* hybridization survey done in parallel in rat and mouse. The availability of this dataset in both organisms will allow researchers to compare and contrast neuronal gene expression in a closely related species using models of molecular evolution. Additionally, this database will augment existing mRNA transcript resources available mainly on tissue or brain sections such as the Allen Brain Atlas [138] or the Brain Maps atlas [139]. Since our study was done on primary cultures it provides additional details at the single-cell level in respect to spatial gene expression patterns in neurons. This investigation is

particularly important as molecular functions and their evolution may not be perceptible at the broader tissue level, as we showed in our previous dendrite transcriptome analysis detailed in Chapter 2.

Another research field that will benefit greatly from this dataset will be the study of RNA *cis* and *trans*-acting factors that determine the differential localization of different mRNA subpopulations. This current analysis identified several evolutionary conserved motifs and their potential RBPs but what I have presented in this chapter is most likely only the tip of the iceberg. The use of RNA motifs prediction algorithms on our extensive data set will allow revealing common sequence and/or structural elements within and between species that could be strong candidates for future functional study. As the rat and mouse genomic sequences are available, the assessment of sequence and structure conservation between and within these species will be facilitated. This assessment will accelerate progress in identifying conserved *cis*-elements and *trans*-acting machineries acting on the post-transcriptional regulation of our curated dendrites transcripts. Also, RBPs involved in transcripts localization in both rat and mouse, such as the evolutionary conserved ELAVL2 and ZFP36 [43, 140], will most likely have human homologs with similar functions and could be relevant candidate for further functional studies toward a better understanding of human brain gene expression regulation.

3.4.6 Future inquiries and conclusion

Does the mRNA localization assessed in this present study hold with similar pattern in other types of neurons? Do they follow similar fine-scale patterns as the one reported here? How do these events translate in more evolutionary distant organisms? What are the functional consequences of these events? These questions, and many more, remain to be addressed. Our current large-scale *in situ* hybridization study along with our dendrite transcriptome array analysis, described in Chapter 2, underscores that subcellular localization is a widely spread event acting on a large number of transcripts, if not the majority. This mechanism seems to involve, in a species-specific fashion, a combination of complex interactions between *cis* and *trans*-factors.

Thus the species-specific differences reported in this study suggest that neuronal function may be more heterogeneous than previously thought. That is, neurons of different species, even relatively closely related species, may have distinct dendritic physiology. Furthermore, it hints at the idea that brain function and organismal behavior may be modulated not only through neuro-anatomy, but also through changes in individual neurons.

3.5 MATERIALS AND METHODS

Selection of dendrite-specific probes for *in situ* hybridization

Microarray data from rat and mouse dissected dendrites (details in Chapter 2) was used to generate a list of the most representative genes in dendrites. This list was based on the highest expressed genes in rat and mouse dendrites. Genes consistently ranked in the top 5% throughout different arrays made with a pool of 100, 200 and 400 dendrites, were chosen as the most expressed. From that pool, 362 candidate dendrites-specific mRNAs were selected to design unique 25-oligomer probes with a Biotin label at their 5'end.

Selection of soma-specific probes for *in situ* hybridization

In order to define a reference for the signal intensity detected by *in situ* hybridization in dendrites versus soma, I selected 12 soma specific genes in both rat and mouse using also our arrays data obtained from dissected dendrites and cell soma. This selection was based on the union of the lowest expressed genes in rat and mouse dendrites together with the highest expressed genes in soma, ranked by the bottom 5% and the top 5% respectively. A majority of these genes are known to be located in the nucleus and cytoplasm of the cell [141-149]. I further refined this list of genes by verifying their low level of expression in dendrites using RNAseq data and choosing the bottom 30%. Seven out of the

twelve genes verified this condition (Table 3.17). I also designed, as described above unique 25-mer probes for these genes and called them “Control probes”.

Table 3.17: List of the 7 soma-specific control probes

Symbol	Accession	Description	Sequence (5'-3')
Rif1	NM_175238	Rap1 interacting factor 1	TCTGGACCAATCTGAAGATCTCTGC
Daglb	NM_144915	Diacylglycerol lipase, beta	CCCCCGGAGCATGATGGCTAACAGG
Lama1	NM_008480.2	Laminin, alpha 1	CATTTAATAGAGGTATCTGTTGTCC
Taf6l	NM_146092.1	TAF6-like RNA polymerase II	TATAAACCGGAAGTGTGGGGCGCCA
Cdca1	NM_023284.3	Cell division cycle associated 1	CAGATTTAATCTTGTGGATTTCTTG
Tex21	XR_034206.1	Testis expressed gene 21	CATAAATTTGGACGACAAGCTTCAG
Pbsn	NM_017471.2	Probasin	AAACATTCAAAGATGCACTAATACC

Design of *in situ* hybridization probes

In order to generate probes with the highest specificity against their targeted transcript and species of interest, I used the sequence of the 50-mer probe provided in the Sentrix BeadArrayTM Technology, Illumina® (RatRef-12 V1.0 and Mouse-6 V1.1 BeadChips, Illumina®) as a starting point. For each probe, all possible 25-mer combinations were computed and blasted simultaneously against both the mouse and rat genome. The final probe selection favored the ones with the best blast scores with a high identity against the transcript of interest (between 75-100%). This identity was either based on a contiguous match of the probe or on a mismatch with few gaps and/or insertions or deletions. The exact percent identity values for each species are available on http://kim.bio.upenn.edu/insitu/public/. Probes selection also took into account intermediate level of GC content (40-60%) to allow a more stable DNA/RNA

binding, as well as to allow low secondary structure in order to avoid self-annealing or hairpin formation. Further details on all probe sequence and gene descriptions, as well as relevant links to other databases, are available on http://kim.bio.upenn.edu/insitu/public/

Culturing conditions

Cortical primary cultures from mouse E18 (C57BL/6, Charles River Laboratories, Inc.) and rat E19 (Sprague-Dawley Charles River Laboratories, Inc.) were plated at 100,000/ml in neurobasal medium (Invitrogen) with B-27 supplement (Sigma) on 12-mm round German Spiegelglas coverslips (Bellco Glass) and grown for 14 days [150]. Mouse and rat embryonic samples used for primary cultures were developmentally matched based on the protocol provided by Charles River Laboratories (http://www.criver.com/SiteCollectionDocuments/rm_rm_d_pregnant_rodent.pdf).

***In situ* hybridization protocol and imaging**

Species-specific biotin-labeled 25 DNA-oligomer probes were custom-made (Sigma-Genosys®). 14 day-old primary rat and mouse cortical neurons were fixed for 15 minutes in 4% paraformaldehyde, washed in 1X PBS and permeabilized with 0.2% TritonX-100 for 10min at room temperature (RT). Cells were prehybridized for 3h at 36°C with 50% formamide, 1X Denhardt's solution, 4X SSC, 10mM DTT, 0.1% CHAPS, 0.1% Tween-20, 500µg/ml yeast tRNA, 500µg/ml salmon sperm DNA. *In situ* hybridization was performed for 16h at

36°C with 15ng/µl probe in prehybridization buffer. After probe hybridization, Rabbit anti-MAP2 (Microtubule Associated Protein 2) primary antibody (1:1000) was added to cells for 1h at RT followed by addition of secondary antibodies Alexa 488 goat anti-rabbit antibody (1:750) and Alexa 568 streptavidin conjugated (1:750) for 1h at RT. The co-staining for MAP2 was performed for two main reasons: First, MAP2 is known to be a marker for dendrites, second MAP2 is conserved in mammals and its expression is known to coincide with the maturation of neuronal morphology and thus could be used as reference baseline for the maturity of both rat and mouse neurons fixed after 14 days in culture [151-153]. DAPI staining was performed before mounting the slides. The samples were visualized by fluorescent microscopy with 2 different exposure times: 200 and 500 ms (Axiovert 200M Inverted Fluorescent Microscope – Zeiss Inc., 20x Objective). All experiments on a given gene as well as the related imaging were performed simultaneously in rat and mouse to avoid potential experimental batch bias and to allow a more consistent and homogenous comparison between these species.

Image processing (in collaboration with Jai-Yoon Sul, Ph.D.)

The collected images were processed via a custom-made Metamorph® image analysis software program. This program includes a manual selection for background subtraction as well as several steps of image-mask generation based on Map2 and DAPI staining that allow extraction of the pixel intensities for the corresponding regions of interest. Figure 3.11 provides a detailed illustration

of this procedure. The three main regions of interest included: the soma, which was defined as double the size of the nucleus (DAPI) minus the central region related to the nucleus (Cytosol Soma Mask in Figure 3.11); the proximal dendrites, which were outlined by a 2.5 times expansion of the soma (Proximal Dendrites Mask in Figure 3.11); and the distal dendrites, which corresponded to all the remaining branched areas that showed positive Map2 staining (Distal Dendrites Mask in Figure 3.11). These three masks were generated for both exposure times (200ms and 500ms) and were used subsequently for processing the data and extracting pixel intensity values.

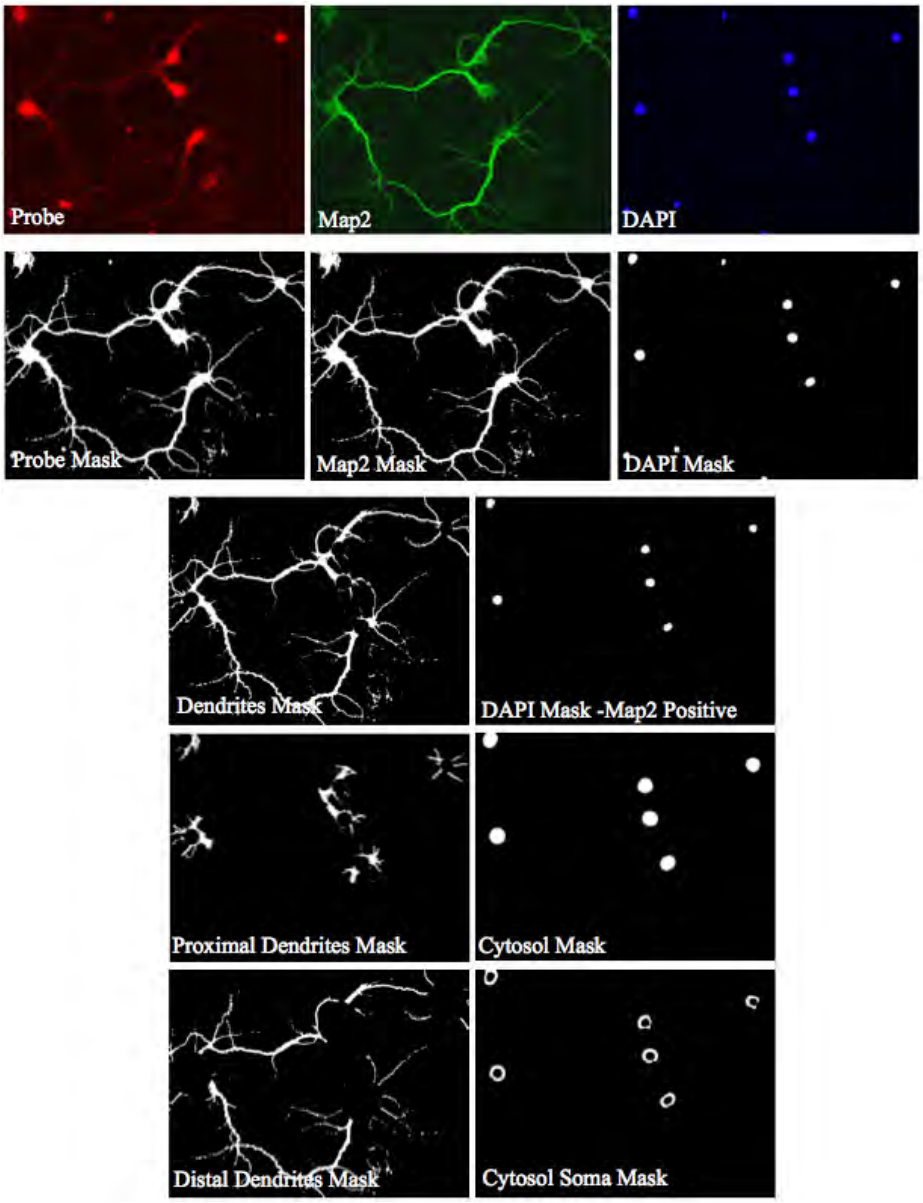


Figure 3.11: Different masks generated for the *in situ* hybridization image analysis.

The raw images Probe, Map2, DAPI are binarized to generate the Probe Mask, Map2 Mask and DAPI Mask. Subsequently the DAPI Mask is overlapped with Map2 Mask to get ride off any non-neuronal cell such as the Glia cells, then its area is expanded 2 times to simulate the cell body area (Cytosol Mask). The final soma area includes the Cytosol Mask minus the DAPI Mask. The Dendrites Mask is generated by the overlap between the Probe and Map2 Masks minus the Cytosol Mask. Additional masks delimited the proximal dendritic region, which was defined by 2.5 times the soma radius, and the distal dendritic region, which corresponded to the whole Dendrites Mask minus the Proximal Dendrites Mask. These dendrites masks allow to distinguish transcripts pixel intensity in proximal versus distal dendrites.

Data processing (in collaboration with Stephen Fisher, Ph.D.)

To avoid any image shift between the different channels (Map2 and Probe), all images were cross-correlated and normalized based on the pixel intensity values at the long exposure time of the probe channel compared to the Map2 channel using MATLAB [154]. Once the image alignment was verified, all images were processed via a custom-made MATLAB imaging program [154] that took in consideration the image-masks generated previously in Metamorph®. Based on the soma/DAPI mask (Cytosol Soma Mask in Figure 3.11), we recovered for each image field an average of 10 cells. As result of this data processing, several statistics on each region of interest (soma, proximal and distal dendrites, at both the short and long exposure time) were computed and recorded in Excel. For instance, these include the average median pixel intensity and its corresponding standard deviation, the Signal to Noise ratio (average pixel intensity/standard deviation) and its corresponding Log value (referred as Dispersion Index).

Data analysis

Reference set up

It is essential to be able to define the relative level of a given transcript in dendrites compared to the soma. For each of the control probes and in each species, I computed the average of the pixel intensity detected in dendrites versus soma and defined it as the ratio Dendrites/Soma (D/S). I then computed, for each species, a “control-cut-off” value corresponding to the overall average (X) of all control probes plus two times the standard deviation (SD). I found for

the rat a cut-off value of 0.06 (with X= 0.04 and SD=0.01) and for the mouse a cut-off value of 0.054 (with X= 0.034 and SD= 0.01). These cut-off values were subsequently used as a reference baseline for all remaining *in situ* signals wherever the signal intensity in dendrites versus soma was computed within each species.

Validation of replicability

Considering the extent of this project, it was not possible to perform a large number of biological replicates for all the genes investigated in this study (~400). Thus, I followed a different strategy to verify the quality and reproducibility of these experiments. For a subset of genes within each species, I generated several biological replicates (2-12 replica), computed their D/S ratio and standard error. All seven soma control probes were replicated between 5 to 12 times and had an average standard error of 25%. Additionally, three probes taken from the list of candidate dendrite-specific genes (eukaryotic translation elongation factor 2, *EEF2; Calmodulin 2, CALM2 and the* coiled-coil domain containing 65, Ccd65) were replicated 3 times and showed consistent D/S values with an average standard error of 20.6%. The consistency seen in these biologically replicated samples strengthens the confidence in the quality and reliability of the data extracted from the remaining non-replicated samples.

RNA binding proteins analysis

All the cultures used for *in situ* hybridization were grown and fixed by the same experimental conditions (See “Culturing conditions” and “*in situ* hybridization protocol” in the above sections). Additionally, within each experimental batch and for each new experiment and animal culture date, four blank control coverslips were included to test the potential culture auto-fluorescence, Map2, Probe and DAPI channels. These followed the exact same hybridization conditions as the other coverslips but did not include the component of interest to be tested (i.e. Map2 anti-body, probe, DAPI). When imaged, these controls did not show any signal intensity, neither uniform nor punctate. Additionally, in the set of probes that were biologically replicated (see the above section “Validation of replicability”), a very good repeatability of smooth staining for the targeted mRNA was detected across replicates when there was no reason to expect punctuation. I classified, in each species, transcripts localized to the dendrites according to whether they formed punctate or uniform distributions. The punctate versus uniform gene classification was first based on sorting the data using the average pixel intensity variance to mean ratio recorded for each image and second a further level of data filtering was done via a visual inspection. This inspection involved scoring each image from 1 to 4 depending on the level of punctuation: a score of 1 corresponds to a uniform smooth probe signal in all the dendrites, a score of 2 to a smooth probe signal even if uneven in some areas of the dendrite, a score of 3 to a discontinuous signal intensity (“punctuated”) in the probe signal

in most cells' dendrites, and a score of 4 to multiple discontinuities in the probe signal along the vast majority of cells' dendrites (see Figure 3.12).

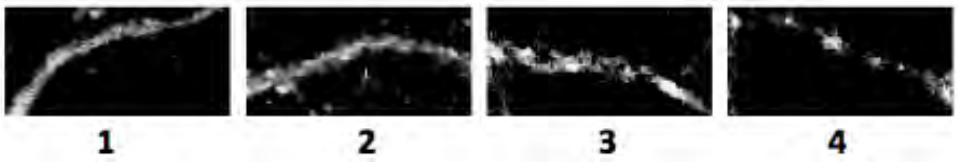


Figure 3.12: Scale for scoring transcripts distribution patterns in dendrites.

I then selected 52 mouse and rat dendrites transcripts, taken from the top 30% punctuated and uniform clusters, and searched within the sequence of each candidate transcript for its RNA binding protein affinity. This investigation was carried out via the publically available RNA binding protein database (http://rbpdb.ccbr.utoronto.ca/) [38]. This database relies on acquiring, within an input nucleotide sequence (in our case transcript sequence taken from either mouse or rat dendrites), the RNA binding protein affinity based upon experimentally proven data in human, mouse, worm and fly. I only considered the relative scores of at least 80% (80-100%) in the Position Weight Matrix (PWM). From these, I selected the two highest overall scores that relate to nucleotide motifs equal or higher than six bases.

MicroRNA analysis

In order to investigate the possible regulation of dendritic gene expression via miRNAs silencing, I searched within our curated dendritic mRNA list for miRNA affinity using the comprehensive miRWalk database (http://www.ma.uni-heidelberg.de/apps/zmf/mirwalk/, version of March 15th, 2011) [56]. The miRWalk

consists of two modules: a predicted targets module and a validated targets module. I mainly focused my search on the validated module. The miRNAs provided in this module were all experimentally validated and reported in the literature (listed in the database). These validations include laser capture combined with multiplex real-time RT (reverse transcription) PCR to quantify microRNAs in specific regions of interest [124]; delivery of microRNAs in vivo by use of recombinant adeno-associated virus (rAAV) with associated functional readout [155]; and targeted deletion, by injection of a retrovirus expressing Cre recombinase, against a given miRNA locus that contains LoxP sites insertion [156]. Additionally, the miRWalk validated targets module provides experimentally validated miRNA interaction information associated to genes, pathways, diseases, organs, cell lines and OMIM disorders. This module includes validated information on 2044 miRNAs from human, mouse, and rat and reports more than 67598 relationship associated to 3821 genes, 375 pathways, 549 diseases, 468 organs, 74 cell lines and 2033 OMIM disorders.
Further detail on this database is available via the following link:
http://www.ma.uni-heidelberg.de/apps/zmf/mirwalk/documentation.html

ID elements mapping in mouse and rat (in collaboration with Mugdha Khaladkar, Ph.D.)

RepeatMasker [157] annotations were downloaded from UCSC Genome Browser for mouse (mm9) and rat (rn4) genome assemblies [158]. The genomic coordinates for all SINE ID elements were extracted and analyzed for overlap

with the Refseq annotated mRNAs in both sense (S) as well as anti-sense (AS) orientation. The presence of ID elements in the intronic, exonic or UTR regions was noted. The 3'UTR region for this purpose was extended 1000 bp downstream of the annotated 3'UTR end in order to account for longer un-annotated 3'UTRs. Overall, there were 33,406 ID elements in mouse (S: 15414, AS: 17992) and 61,311 ID elements in rat (S: 27396, AS: 33915) that mapped to the Refseq mRNAs.

Database generation (in collaboration with Stephen Fisher, Ph.D.)

A publically available database was constructed in order to keep record of this large-scale *in situ* dataset and to become a potential public resource and guideline for future investigations: http://kim.bio.upenn.edu/insitu/public/.
This database not only provides a listing of the transcripts investigated and their corresponding images and pixel intensities, but also includes many additional features, like data filtering/sorting options, full MySQL queries, and the ability to download the raw Tiff images. As researchers will have access to our raw data, they could perform their own analyses to address to their inquiries more appropriately.

Table 3.1: List of transcripts dendritically localized in mouse

Gene Symbol	Mouse RefSeq	Gene Description
Ppia	XM_913899	peptidylprolyl isomerase A
Gria2	NM_001083806	glutamate receptor, ionotropic, AMPA2 (alpha 2)
Mecp2	NM_001081979	methyl CpG binding protein 2
Omd	NM_012050	osteomodulin
Zfp410	NM_144833	zinc finger protein 410
Hsp1	NM_010480	heat shock protein 90, alpha (cytosolic), class A member 1
Mtap1a	NM_032393	microtubule-associated protein 1 A
Krtap6type2	NM_010673	keratin associated protein 6-2
Cfl1	NM_007687	cofilin 1, non-muscle
Slc25a40	NM_178766	solute carrier family 25, member 40
Dync1i1	NM_010063	dynein cytoplasmic 1 intermediate chain 1
V1ra4	NM_053219	vomeronasal 1 receptor, A4
Fmnl1	NM_001077698	formin-like 1
Mtap2	NM_008632	microtubule-associated protein 2
Atp5b	NM_016774	ATP synthase, H+ transporting mitochondrial F1 complex, beta subunit
Olfr259	NM_146735	olfactory receptor 483
Atf3	NM_007498	activating transcription factor 3
Vapb	NM_019806	vesicle-associated membrane protein, associated protein B and C
Dhx8	NM_144831	DEAH (Asp-Glu-Ala-His) box polypeptide 8
Slc22a15	NM_001039371	solute carrier family 22 (organic anion/cation transporter), member 15
Orai3	NM_198424	ORAI calcium release-activated calcium modulator 3
Ifitm7	XM_977607	predicted gene, EG665536; interferon induced transmembrane protein 7
Kcna2	NM_008417	potassium voltage-gated channel, shaker-related subfamily, member 2
Timm23	XM_001476967	translocase of inner mitochondrial membrane 23 homolog
Atp5d	NM_172294	sulfatase 1
LOC366974	NM_133726	suppression of tumorigenicity 13
Gltscr1	NM_001081418	glioma tumor suppressor candidate region gene 1
Cdh16	NM_007663	cadherin 16
Govn5	NM_177764	vomeronasal 2, receptor 57
H2a	XM_001480384	H2A histone family, member Z; predicted gene 6722; predicted gene 8203
Sdha	NM_023281	succinate dehydrogenase complex, subunit A, flavoprotein (Fp)
Pklr	NM_013631	pyruvate kinase liver and red blood cell
Arhgef7	NM_017402	Rho guanine nucleotide exchange factor (GEF7)
LOC362994	NM_153518	coiled-coil domain containing 65
Gnas	NM_001077507	GNAS (guanine nucleotide binding protein, alpha stimulating) complex locus
Def8	NM_054046	differentially expressed in FDCP 8
Jub	NM_010590	ajuba
Fkbp2	NM_008020	FK506 binding protein 2
Lypla1	NM_008866	lysophospholipase 1
Mtrr	NM_172480	5-methyltetrahydrofolate-homocysteine methyltransferase reductase
Sat2	NM_026991	spermidine/spermine N1-acetyl transferase 2
Stip1	NM_016737	stress-induced phosphoprotein 1
Rps15	NM_009091	ribosomal protein S15
Zmpste24	NM_172700	zinc metallopeptidase, STE24 homolog (S. cerevisiae)
Trmt112	NM_026306	RIKEN cDNA 0610038D11 gene; predicted gene 13072; predicted gene 7828
Alg11	NM_183142	asparagine-linked glycosylation 11 homolog (yeast, alpha-1,2-mannosyltransferase)
Olfr386	XM_981170	olfactory receptor 384; olfactory receptor 386
Rpl23	XM_001477371	similar to ribosomal protein L23
Hp1	XM_001475747	similar to Serine/arginine repetitive matrix protein 2

4922501c03Rik	XM_001000297	RIKEN cDNA 4922501C03 gene
Kif15	NM_010620	kinesin family member 15
LOC498015	NM_145823	phosphatidylinositol transfer protein, cytoplasmic 1
Hsp40	NM_016775	DnaJ (Hsp40) homolog, subfamily C, member 5
LOC683007	NM_177389	melanoma inhibitory activity 3
Aptase	NM_018794	ATPase, H+ transporting, lysosomal accessory protein 1
Hnrpk	NM_025279	heterogeneous nuclear ribonucleoprotein K; predicted gene 7964
LOC500054	NM_028370	protection of telomeres 1B
Lhx3	NM_001039653	LIM homeobox protein 3
Atp2b2	NM_009723	ATPase, Ca++ transporting, plasma membrane 2
Spnb3	NM_021287	spectrin beta 3
Ftl1	XM_001478411	ferritin light chain 2; similar to Ferritin light chain 1 (Ferritin L subunit 1)
Rhobtb1	NM_001081347	Rho-related BTB domain containing 1
Dpysl2	NM_009955	dihydropyrimidinase-like 2
Strbp	NM_009261	spermatid perinuclear RNA binding protein
LOC687463	NM_007790	predicted gene 8892; structural maintenace of chromosomes 3
Slmap	NM_032008	sarcolemma associated protein
Gnbp	NM_010312	guanine nucleotide binding protein (G protein), beta 2
Ubiqfep6	NM_177823	ubiquitin associated and SH3 domain containing, A
Myh2	NM_001039545	myosin, heavy polypeptide 2,1, skeletal muscle, adult
Usp8	NM_019729	ubiquitin specific peptidase 8
Gdia	NM_133796	Rho GDP dissociation inhibitor (GDI) alpha
Arl6ip5	NM_022992	ADP-ribosylation factor-like 6 interacting protein 5
Ef2	NM_007907	predicted gene 13050; eukaryotic translation elongation factor 2
Dynll2	NM_001168472	dynein light chain LC8-type 2
Cox5b	XM_001475417	cytochrome c oxidase, subunit Vb
Mapk1	NM_001038663	mitogen-activated protein kinase 1
Nefl	NM_010910	neurofilament, light polypeptide
LOC682988	NM_026039	mediator of RNA polymerase II transcription, subunit 18 homolog (yeast)
Snrpn	XM_001480670	small nuclear ribonucleoprotein N
Rspa	XM_916601	ribosomal protein SA pseudogene; similar to 40S ribosomal protein SA (p40)
Gabrg1	NM_010252	gamma-aminobutyric acid (GABA) A receptor, subunit gamma 1
Kif17	NM_010623	kinesin family member 17
Eif3s9	XM_001479308	eukaryotic translation initiation factor 3, subunit B
Fgf13	NM_010200	fibroblast growth factor 13
Atp5a1	NM_007505	ATP synthase, H+ transporting, mitochondrial F1 complex, alpha subunit, isoform 1
Prph1	NM_001163588	peripherin
Rbfox1	NM_183188	RNA binding protein, fox-1 homolog (C. elegans) 1
Eif4a2	NM_001123038	eukaryotic translation initiation factor 4A2
Marcksl1	NM_010807	MARCKS-like 1; predicted gene 9106
Rnf111	NM_033604	predicted gene 6162; similar to ring finger 111; ring finger 111
Pap	NM_011036	regenerating islet-derived 3 beta
Rpl35a	NM_001130484	ribosomal protein L35A
Trpv5	NM_001007572	transient receptor potential cation channel, subfamily V, member 5
Med14	NM_001048208	mediator complex subunit 14
Kcnk9	NM_001033876	potassium channel, subfamily K, member 9
Cacna1a	NM_007578	calcium channel, voltage-dependent, P/Q type, alpha 1A subunit
Prdx2	NM_011563	peroxiredoxin 2
Usp9x	NM_009481	ubiquitin specific peptidase 9, X chromosome
Syt6	NM_018800	synaptotagmin VI
Suclg1	NM_019879	succinate-CoA ligase, GDP-forming, alpha subunit
Cacng5	NM_080644	calcium channel, voltage-dependent, gamma subunit 5
Irgm	NM_008326	immunity-related GTPase family M member 1
Rpl6	NM_011290	ribosomal protein L6
9130011E15Rik	NM_198296	RIKEN cDNA 9130011E15 gene

Dynll1	XM_892488	similar to cytoplasmic dynein light chain 1
Olr1273	NM_146442	olfactory receptor 934
Mdh2	NM_008617	malate dehydrogenase 2, NAD (mitochondrial)
Dcb1	NM_019967	deleted in bladder cancer 1 (human)
Vim	NM_011701	vimentin
Apc2	NM_007459	adaptor protein complex AP-2, alpha 2 subunit
LOC500420	XM_001475329	RIKEN cDNA 2610029I01 gene
Pcdh10	NM_011043	protocadherin 10
Cdh2	NM_007664	cadherin 2; similar to N-cadherin
Hspa13	NM_030201	heat shock protein 70 family, member 13
Scd2	NM_009128	stearoyl-Coenzyme A desaturase 2
Rps4	NM_009094	ribosomal protein S4, X-linked
Tpt1	XM_001480658	tumor protein, translationally-controlled 1 pseudogene
Sar1b	NM_025535	SAR1 gene homolog B (S. cerevisiae)
Hspa12b	NM_028306	heat shock protein 12B
Tmem50a	NM_027935	transmembrane protein 50A
Ak7	XM_994344	adenylate kinase 7
LOC688007	XM_001480879	cDNA sequence BC030499
Tmeff1	NM_021436	transmembrane protein with EGF-like and two follistatin-like domains 1
Gnbp	NM_010308	guanine nucleotide binding protein, alpha O
Syt4	NM_009308	synaptotagmin IV
Rps23	NM_001100608	ribosomal protein S23
Cdh9	NM_177224	similar to chromodomain helicase DNA binding protein 9
Slc25a4	NM_007450	solute carrier family 25 (mitochondrial carrier adenine translocator) member 4
Tmem30a	NM_133718	transmembrane protein 30A
Aes	NM_001114179	mitochondrial carrier triple repeat 6
Pcdh17	NM_001013753	protocadherin 17
Rpl36a	XM_001480124	similar to large subunit ribosomal protein L36a
Eif4g2	NM_013507	eukaryotic translation initiation factor 4, gamma 2
Sec13	NM_024206	SEC13 homolog (S. cerevisiae)
Lypla1	NM_008866	lysophospholipase 1
Pou4f2	NM_138944	POU domain, class 4, transcription factor 2
Kcnj3	NM_008426	potassium inwardly-rectifying channel, subfamily J, member 3
Pcdhgb5	NM_033577	protocadherin gamma subfamily B, 5
Cort	NM_007745	cortistatin
Olfm1	NM_019498	olfactomedin 1
LOC362306	NM_007476	predicted gene 5823; ADP-ribosylation factor 1; predicted gene 8230
Mageb1	NM_031171	melanoma antigen, family B, 1; melanoma antigen, family B, 2
LOC364964	NM_172707	protein phosphatase 1, catalytic subunit, beta isoform
Basp1	NM_027395	brain abundant, membrane attached signal protein 1
Becn1	NM_026562	similar to cyclin N-terminal domain containing 1
Gnbp	NM_001160016	guanine nucleotide binding protein (G protein), beta 1
Sept9	NM_001113488	septin 9
Dst	NM_134448	dystonin; hypothetical protein LOC100047109
Syngr1	NM_207708	synaptogyrin 1
Uba52	NM_019883	ubiquitin A-52 residue ribosomal protein fusion product 1
Ttc21b	NM_001047604	tetratricopeptide repeat domain 21B
Rprd1b	NM_027434	regulation of nuclear pre-mRNA domain containing 1B
LOC499390	NM_016680	splicing factor, arginine/serine-rich 16
Nbea	NM_030595	neurobeachin
Rad1	NM_011232	RAD1 homolog (S. pombe)
Ptms	NM_026988	parathymosin
Opa1	NM_133752	similar to optic atrophy 1 (autosomal dominant)
Pcbp1	NM_011865	poly(rC) binding protein 1
Appbp2	NM_025825	amyloid beta precursor protein (cytoplasmic tail) binding protein 2
Rps3	NM_012052	ribosomal protein S3

Pcdh18	NM_130448	protocadherin 18
Rab14	NM_026697	RAB14, member RAS oncogene family
H13	NM_010376	histocompatibility 13
Glp2r	NM_175681	similar to glucagon-like peptide 2 receptor; glucagon-like peptide 2 receptor
Spag7	XM_001471841	hypothetical protein LOC100048812; sperm associated antigen 7
Myo5a	NM_010864	myosin VA
Sirt1	NM_001159589	sirtuin 1 (silent mating type information regulation 2, homolog) 1 (S. cerevisiae)
Ldhb	NM_008492	lactate dehydrogenase B
Ilkap	NM_023343	integrin-linked kinase-associated serine/threonine phosphatase 2C
Vsnl1	NM_012038	visinin-like 1
Prelid1	XM_001476721	PRELI domain containing 1; similar to Preli

Red color = Transcripts dendritically localized with low confidence
Black color = Transcripts dendritically localized with high confidence

Table 3.2: List of transcripts dendritically localized in rat

Gene Symbol	Rat RefSeq	Gene Description
Canx	NM_172008.1	calnexin (Canx)
Rhobtb1	XM_001074130.1	Rho-related BTB domain containing 1 (predicted), transcript variant 1
Laptm4a	NM_199384.1	lysosomal-associated protein transmembrane 4A (Laptm4a)
Rnf111	XM_001055544.1	ring finger protein 111 (Rnf111_predicted)
Atp6v1b2	NM_057213.2	ATPase, H transporting, lysosomal V1 subunit B2 (Atp6v1b2)
RGD1304782	XM_230798.4	similar to RIKEN cDNA 2610304G08 gene (RGD1304782_predicted)
Sirt5	NM_001004256.1	sirtuin 5 (silent mating type information regulation 2 homolog) 5 (Sirt5)
RGD1306538	NM_001014024.1	similar to hypothetical protein MGC13024 (RGD1306538)
Atp5j	NM_053602.1	ATP synthase, H+ transporting, mitochondrial F0 complex, subunit F6 (Atp5j)
Vdac2	NM_031354.1	voltage-dependent anion channel 2 (Vdac2)
RGD1559672	XM_001078915.1	similar to Translocase of inner mitochondrial membrane 23 homolog
Mtap2	NM_013066.1	microtubule-associated protein 2 (Mtap2)
Uba52	NM_031687.2	ubiquitin A-52 residue ribosomal protein fusion product 1 (Uba52)
LOC691644	XM_001077494.1	Myosin heavy chain, skeletal muscle, adult2 (Myosin heavy chain IIa), MyHCIIa
Rps20	NM_001007603.1	ribosomal protein S20 (Rps20)
LOC683007	XM_001064186.1	similar to melanoma inhibitory activity 3 (LOC683007)
LOC364750	XR_008099.1	similar to ribosomal protein S14 (LOC364750)
Slc3a1	NM_017216.1	solute carrier family 3, member 1 (Slc3a1)
Dynll1	NM_053319.2	dynein light chain LC8-type 1 (Dynll1)
Rps6	NM_017160.1	ribosomal protein S6 (Rps6)
Syt4	NM_031693.1	synaptotagmin IV (Syt4)
Brinp2	NM_173115.1	BMP/retinoic acid-inducible neural-specific protein 2 (Brinp2)
Sec13l1	NM_001006978.1	SEC13-like 1 (S. cerevisiae) (Sec13l1)
Appbp2	XM_001081113.1	amyloid beta precursor protein (cytoplasmic tail) binding protein 2 (Appbp2)
RGD1561319	XM_575592.2	similar to Pol(yrC)-binding protein 1 (Alpha-CP1) (hnRNP-E1) (
Atp5c1	NM_053825.1	ATP synthase H+ transporting mitochondrial F1 complex gamma polypeptide1
RGD1565811	XM_001054521.1	similar to OL-protocadherin isoform (RGD1565811_predicted)
Lhx3	XM_001059910.1	LIM homeobox protein 3 (Lhx3)
RAMP4	NM_030835.2	ribosome associated membrane protein 4 (RAMP4)
Olfm1	NM_053573.1	olfactomedin 1 (Olfm1)
Sipa1l1	NM_139330.1	signal-induced proliferation-associated 1 like 1 (Sipa1l1)
Atp5g3	NM_053756.1	ATP synthase H+ transporting mitochondrial F0 complex SU c (9) isoform3
Atrn	NM_031351.1	attractin (Atrn)
Tpt1	NM_053867.1	tumor protein, translationally-controlled 1 (Tpt1)
LOC499390	NM_001024294.1	similar to SWAP2 (LOC499390)
Cyfip2	XM_220333.4	cytoplasmic FMR1 interacting protein 2 (Cyfip2_predicted)
Opcml	NM_053848.1	opioid binding protein/cell adhesion molecule-like (Opcml)
Opa1	NM_133585.2	optic atrophy 1 homolog (human) (Opa1)
Atp2a2	XM_001079470.1	ATPase, Ca++ transporting, cardiac muscle, slow twitch 2 (Atp2a2)
H13	XM_230734.4	histocompatibility 13 (H13_predicted)
LOC500054	NM_001024322.1	similar to POT1-like telomere end-binding protein (LOC500054)
Glp2r	NM_021848.1	glucagon-like peptide 2 receptor (Glp2r)
Myo5a	NM_022178.1	myosin Va (Myo5a)
MGC105647	NM_001007008.1	similar to Nur77 downstream protein 2 (MGC105647)
Klhl2	XM_214331.4	kelch-like 2, Mayven (Drosophila) (Klhl2_predicted)
Pebp1	NM_017236.1	phosphatidylethanolamine binding protein 1 (Pebp1)
RGD1562118	XR_008143.1	similar to peroxiredoxin 2 isoform b (RGD1562118_predicted)

Adcy5	NM_022600.1	adenylate cyclase 5 (Adcy5)
LOC286914	NM_173130.1	putative pheromone receptor (Go-VN5) (LOC286914)
Rpl3	NM_198753.1	ribosomal protein L3 (Rpl3)
Psma5	NM_017282.1	proteasome (prosome, macropain) subunit, alpha type 5 (Psma5)
Atxn2	XM_001079639.1	ataxin 2 (Atxn2_predicted)
Pigp	XM_213650.3	phosphatidylinositol glycan, class P (Pigp_predicted)
Rab5c	XM_001081435.1	RAB5C, member RAS oncogene family (Rab5c_predicted)
RGD1564725	XM_341459.2	similar to hypothetical protein B230397C21 (RGD1564725_predicted)
Rps4x	NM_001007600.1	ribosomal protein S4, X-linked (Rps4x)
Atp5a1	NM_023093.1	ATP synthase, H+ transporting, mitochondrial F1 complex, subunit a, isoform 1
Atf3	NM_012912.1	activating transcription factor 3 (Atf3)
Fgf13	NM_053428.1	fibroblast growth factor 13 (Fgf13)
Jub	NM_053503.1	ajuba homolog (Xenopus laevis) (Jub)
LOC362994	NM_001014203.1	similar to NYD-SP28 protein (LOC362994)
Mapk1	NM_053842.1	mitogen activated protein kinase 1 (Mapk1)
Mor1	NM_031151.2	malate dehydrogenase, mitochondrial (Mor1)
Stmn2	NM_053440.2	stathmin-like 2 (Stmn2)
LOC500420	XM_575783.2	similar to CG12279-PA (LOC500420)
H2afz	NM_022674.1	H2A histone family, member Z (H2afz)
Sdcbp	NM_031986.1	syndecan binding protein (Sdcbp)
LOC682988	XM_001065118.1	mediator of RNA polymerase II transcription, SU 18 homolog, variant2
Mtap1a	NM_030995.1	microtubule-associated protein 1 A (Mtap1a)
RGD1306344	XM_230878.4	similar to Ab1-133 (RGD1306344)
Snrpn	NM_031117.1	small nuclear ribonucleoprotein N (Snrpn)
Arhgef7	NM_053740.1	Rho guanine nucleotide exchange factor 7 (Arhgef7)
Atp1b2	NM_012507.2	ATPase, Na+/K+ transporting, beta 2 polypeptide (Atp1b2)
Gnb1	NM_030987.1	guanine nucleotide binding protein, beta 1 (Gnb1)
Ap2a2	NM_031008.1	adaptor protein complex AP-2, alpha 2 subunit (Ap2a2)
Tmeff1	NM_023020.1	transmembrane protein with EGF-like and two follistatin-like domains 1
Kcnj3	NM_031610.1	potassium inwardly-rectifying channel, subfamily J, member 3 (Kcnj3)
Stxbp1	NM_013038.3	syntaxin binding protein 1 (Stxbp1)
Basp1	NM_022300.1	brain abundant, membrane attached signal protein 1 (Basp1)
Rpl23	NM_001007599.1	ribosomal protein L23 (Rpl23)
Gria2	NM_017261.1	glutamate receptor, ionotropic, AMPA2 (Gria2)
Ak7	XM_234507.4	adenylate kinase 7 (Ak7_predicted)
Fmnl1	XM_001081542.1	formin-like 1 (Fmnl1_predicted)
Slc25a4	NM_053515.1	solute carrier family 25 (mitochondrial carrier adenine translocator) member4
LOC680231	XR_005671.1	similar to chromodomain helicase DNA binding protein 9 (LOC680231)
Eif4a2	NM_001008335.1	eukaryotic translation initiation factor 4A2 (Eif4a2)
Cox5b	NM_053586.1	cytochrome c oxidase subunit Vb (Cox5b)
EF2	XR_006068.1	similar to Elongation factor 2 (EF-2) (LOC305181)
RGD1560170	XM_228713.4	cofactor required for Sp1 transcriptional activation, subunit 2, 150kDa
Olr259	NM_001000222.1	olfactory receptor 259 (Olr259_predicted)
Grina	NM_153308.1	glutamate receptor ionotropic N-methyl D-asparate-associated prot.1 (Grina)
Wsb2	NM_001007616.1	WD repeat and SOCS box-containing 2 (Wsb2)
Stmn3	NM_024346.1	stathmin-like 3 (Stmn3)
Scd2	NM_031841.1	stearoyl-Coenzyme A desaturase 2 (Scd2)
Atp2b2	NM_012508.3	ATPase, Ca++ transporting, plasma membrane 2 (Atp2b2)
Syngr1	NM_019166.1	synaptogyrin 1 (Syngr1)
LOC362306	XR_006311.1	hypothetical LOC362306 (LOC362306)
Arl6ip5	NM_023972.2	ADP-ribosylation factor-like 6 interacting protein 5 (Arl6ip5)
Lyl1	XM_001071678.1	lymphoblastic leukemia derived sequence 1 (Lyl1)
Becn1	NM_001034117.1	beclin 1 (coiled-coil, myosin-like BCL2-interacting protein), transcript variant 2

Cacng5	NM_080693.1	calcium channel, voltage-dependent, gamma subunit 5 (Cacng5)
Rps21	NM_031111.1	ribosomal protein S21 (Rps21)
Pklr	NM_012624.2	pyruvate kinase, liver and red blood cell (Pklr)
Apba1	NM_031779.1	amyloid beta (A4) precursor protein-binding, family A, member 1 (Apba1)
Sirt1	XM_228146.4	sirtuin 1 (silent mating type information regulation 2, homolog 1
Slc7a8	NM_053442.1	solute carrier family 7 (cationic amino acid transporter), member 8
Cdh16	NM_001012055.1	cadherin 16 (Cdh16)
Rps15	NM_017151.2	ribosomal protein S15 (Rps15)
Grb2	NM_030846.2	growth factor receptor bound protein 2 (Grb2)
Pkm2	NM_053297.1	pyruvate kinase, muscle (Pkm2)
Calm3	NM_012518.2	calmodulin 3 (Calm3)
Syn2	NM_019159.1	synapsin II (Syn2), transcript variant 2
Vgf	NM_030997.1	VGF nerve growth factor inducible (Vgf)
Gnai1	NM_013145.1	guanine nucleotide binding protein, alpha inhibiting 1 (Gnai1)
RGD1563644	XR_008237.1	similar to Cofilin, non-muscle isoform (Cofilin-1) (RGD1563644_predicted)
Lypla1	NM_013006.1	lysophospholipase 1 (Lypla1)
Atp5d	NM_139106.1	ATP synthase H+ transporting mitochondrial F1 complex, delta subunit(Atp5d)
LOC366974	XR_007583.1	similar to suppression of tumorigenicity 13 (LOC366974)
LOC683125	XM_001064575.1	similar to 40S ribosomal protein SA (p40) (34/67 kDa laminin receptor)
Tmem50a	XM_001067840.1	transmembrane protein 50A (Tmem50a_predicted)
Cdh2	NM_031333.1	cadherin 2 (Cdh2)
Rpl6	NM_053971.1	ribosomal protein L6 (Rpl6)
Ndufs6	NM_019223.1	NADH dehydrogenase (ubiquinone) Fe-S protein 6 (Ndufs6)
Dnajc5	NM_024161.2	DnaJ (Hsp40) homolog, subfamily C, member 5 (Dnajc5)
Nsf	NM_021748.1	N-ethylmaleimide sensitive fusion protein (Nsf)
RGD1564887	XM_574677.2	similar to 9130011E15Rik protein (RGD1564887_predicted)
RGD1564958	XM_214281.4	similar to glyceraldehyde-3-phosphate dehydrogenase (phosphorylating)
Vim	NM_031140.1	vimentin (Vim)
Dync1i1	NM_019234.1	dynein cytoplasmic 1 intermediate chain 1 (Dync1i1)
RGD1308082	NM_001009636.1	similar to px19-like protein (RGD1308082)
LOC686547	XM_001074235.1	TBC1 domain family member 4, Akt substrate of160kDa (AS160)
Mapre3	NM_001007656.1	microtubule-associated protein, RP/EB family, member 3 (Mapre3)
LOC288165	XM_001057918.1	similar to PEST-containing nuclear protein, transcript variant 1 (LOC288165)
Kcnk9	NM_053405.1	potassium channel, subfamily K, member 9 (Kcnk9)
Boll	XM_220155.4	bol, boule-like (Drosophila) (Boll_predicted)
Suclg1	NM_053752.1	succinate-CoA ligase, GDP-forming, alpha subunit (Suclg1)
RGD1309710	XM_215167.2	similar to RIKEN cDNA 0610038D11 (RGD1309710_predicted)
RGD1562629	XM_001059612.1	similar to neurobeachin (RGD1562629_predicted)
Hnrpab	NM_031330.1	heterogeneous nuclear ribonucleoprotein A/B (Hnrpab)
Gyg1	NM_031043.1	glycogenin 1 (Gyg1)
Fhod1	NM_001191600	FH1/FH2, Formin homology 2 domain-containing protein 1
Psmb2	NM_017284.1	proteasome (prosome, macropain) subunit, beta type 2 (Psmb2)
Tmem30a	NM_001004248.1	transmembrane protein 30A (Tmem30a)
Aco2	NM_024398.2	aconitase 2, mitochondrial (Aco2)
Spag7	XM_001079933.1	sperm associated antigen 7 (Spag7_predicted)
Aars	XM_214690.4	alanyl-tRNA synthetase (Aars)
Rpl30	NM_022699.2	ribosomal protein L30 (Rpl30)
RGD1562920	XM_214790.4	similar to Aig1 protein (RGD1562920_predicted)
Sept9	NM_176856.1	septin 9 (Sept9), transcript variant 2
Gabarapl2	NM_022706.2	GABA(A) receptor-associated protein like 2 (Gabarapl2)
Pabpn1	XM_001055786.1	poly(A) binding protein, nuclear 1 (Pabpn1)
Olr1273	NM_001000458.1	olfactory receptor 1273 (Olr1273_predicted)
Stx1b2	NM_012700.1	syntaxin 1B2 (Stx1b2)

Mecp2	NM_022673.1	methyl CpG binding protein 2 (Mecp2)
RGD1564579	XM_215057.4	similar to yippee-like 3 (RGD1564579_predicted)
Higd2a	XM_214433.3	HIG1 domain family, member 2A (Higd2a_predicted)
RGD1565122	XM_229983.4	similar to tetratricopeptide repeat domain 21B (RGD1565122_predicted)
Pdcd5	XM_001079809.1	programmed cell death 5 (Pdcd5_predicted)
Usp9x	XM_343766.3	ubiquitin specific peptidase 9, X chromosome (Usp9x_predicted)
Marcksl1	NM_030862.2	MARCKS-like 1 (Marcksl1)
LOC680448	XR_005755.1	similar to protocadherin gamma subfamily B, 5 (LOC680448)
RGD1308774	NM_001037186.1	similar to mitochondrial carrier family protein (RGD1308774)
Calm2	NM_017326.1	calmodulin 2 (Calm2)
Ccni	XM_001070498.1	cyclin I (Ccni_predicted)
Dpysl2	XM_573810.2	dihydropyrimidinase-like 2 (Dpysl2)
Dst	XM_001054738.1	dystonin (Dst_predicted)
Irgm	NM_001012007.1	immunity-related GTPase family, M (Irgm)
Syt6	NM_022191.1	synaptotagmin VI (Syt6)
Ftl1	NM_022500.3	ferritin light chain 1 (Ftl1)
Nefl	NM_031783.1	neurofilament, light polypeptide (Nefl)
Myo5b	NM_017083.1	myosin 5B (Myo5b)
Ppm1e	NM_198773.1	protein phosphatase 1E (PP2C domain containing) (Ppm1e)
Kif15	NM_181635.2	kinesin family member 15 (Kif15)
LOC688007	XM_001080857.1	hypothetical protein LOC688007 (LOC688007)
Rab43	NM_001024331.1	Ras-related protein RAB43 (Rab43)
Glt25d1	XM_214295.4	glycosyltransferase 25 domain containing 1 (Glt25d1_predicted)
Pou4f2	XM_344756.3	POU domain, class 4, transcription factor 2 (Pou4f2)
Gnao	NM_017327.1	guanine nucleotide binding protein, alpha o (Gnao)
Prkch	NM_031085.2	protein kinase C, eta (Prkch)
Vps35	XM_214646.4	vacuolar protein sorting 35 (Vps35_mapped)
Pcdh17	XM_224389.4	protocadherin 17 (Pcdh17_predicted)
Vegfa	NM_031836.1	vascular endothelial growth factor A (Vegfa)

Blue color = Transcripts dendritically localized with low confidence
Black color = Transcripts dendritically localized with high confidence

Table 3.10: List of transcripts examined in rat and mouse neurons.

Rat Symbol	Mouse Symbol	Description
Vsnl1	Vsnl1	visinin-like 1
Ilkap	Ilkap	integrin-linked kinase-associated serine/threonine phosphatase 2C
Ldhb	Ldhb	lactate dehydrogenase B
Rab14	Rab14	RAB14, member RAS oncogene family
Ptms	Ptms	parathymosin
Rad1	Rad1	RAD1 homolog
RGD1304782	Rprd1b	RIKEN cDNA 2610304G08 gene
LOC365181	LOC364964	Serine/threonine-protein phosphatase PP1-beta catalytic subunit
RGD1561997	Mageb1	Smage-1 protein
Cort	Cort	cortistatin
Rpl36a	Rpl36a	large subunit ribosomal protein L36a
Aes	Aes	amino-terminal enhancer of split
Sar1b	Sar1b	SAR1 gene homolog B
Stch	Hspa13	stress 70 protein chaperone, microsome-associated, 60kD human homolog
Dbccr1	Dcb1	deleted in bladder cancer chromosome region candidate 1
Cacna1a	Cacna1a	calcium channel, voltage-dependent, P/Q type, alpha 1A subunit
Trpv5	Trpv5	transient receptor potential cation channel, subfamily V, member 5
Pap	Pap	pancreatitis-associated protein
Rnf111	Rnf111	ring finger protein 111
Prph1	Prph1	peripherin 1
Eif3s9	Eif3s9	eukaryotic translation initiation factor 3, subunit 9
RGD1562511	Kif17	Kinesin 17
Arhgdia	Gdia	Rho GDP dissociation inhibitor
Usp8	Usp8	Ubiquitin specific protease 8
Gnb2	Gnbp	guanine nucleotide binding protein, beta polypeptide 2
Slmap	Slmap	sarcolemma associated protein
SmcD	SmcD	Structural maintenance of chromosome 3
Strbp	Strbp	spermatid perinuclear RNA binding protein
Rhobtb1	Rhobtb1	Rho-related BTB domain containing 1 , transcript variant 1
Spnb3	Spnb3	spectrin beta 3
Hnrpk	Hnrpk	heterogeneous nuclear ribonucleoprotein K
Atp6ap1	Aptase	ATPase, H+ transporting, lysosomal accessory protein 1
LOC303619	LOC635277	retinal degeneration B beta
RGD1307365	4922501c03Rik	KIAA1009 protein
Cbx3	Hp1	chromobox homolog 3
Stip1	Stip1	stress-induced phosphoprotein 1
Sat2	Sat2	spermidine/spermine N1-acetyl transferase 2
Mtrr	Mtrr	5-methyltetrahydrofolate-homocysteine methyltransferase reductase
RGD1564018	Lypla1	lysophospholipase I
Fkbp2	Fkbp2	FK506 binding protein 2
Gnas	Gnas	GNAS complex locus
Sdha	Sdha	succinate dehydrogenase complex, subunit A, flavoprotein
Gltscr1	Gltscr1	glioma tumor suppressor candidate region gene 1
RGD1559672	Timm23	Translocase of inner mitochondrial membrane 23 homolog
RGD1306538	Orai3	hypothetical protein MGC13024
Dhx8	Dhx8	DEAH
Vapb	Vapb	vesicle-associated membrane protein, associated protein B and C
Atp5b	Atp5b	ATP synthase, H+ transporting, mitochondrial F1 complex, beta polypeptide
Mtap2	Mtap2	microtubule-associated protein 2
RGD1563644	Cofilin	Cofilin, non-muscle isoform
Mtap1a	Mtap1a	microtubule-associated protein 1 A
Mecp2	Mecp2	methyl CpG binding protein 2

Gria2	Glur	glutamate receptor, ionotropic, AMPA2
LOC686547	Akt	TBC1 domain family member 4
Pabpn1	Pabpn1	poly
Rps6	Rps6	ribosomal protein S6
Sipa1l1	Sipa1l1	signal-induced proliferation-associated 1 like 1
LOC291964	Fhod1	FH1/FH2 domain-containing protein
Pkm2	Pk	pyruvate kinase, muscle
Prkch	Prkch	protein kinase C, eta
Ppm1e	Ppm1e	protein phosphatase 1E
Grina	Glur	glutamate receptor, ionotropic, N-methyl D-asparate-associated protein 1
Myo5b	Myo5b	myosin 5B
Vps35	Vps35	vacuolar protein sorting 35
Glt25d1	Glt25d1	glycosyltransferase 25 domain containing 1
Rab43	Rab43	Ras-related protein RAB43
RGD1564579	Ypel3	yippee-like 3
Aars	Aars	alanyl-tRNA synthetase
Mapre3	Map	microtubule-associated protein, RP/EB family, member 3
Pigp	Pigp	phosphatidylinositol glycan, class P
Atxn2	Atxn2	ataxin 2
RGD1308082	Prelid1	px19-like protein
Sirt1	Sirt1	sirtuin 1
Myo5a	Myo5a	myosin Va
Spag7	Spag7	sperm associated antigen 7
Glp2r	Glp2r	glucagon-like peptide 2 receptor
H13	H13	histocompatibility 13
Appbp2	Appbp2	amyloid beta precursor protein
RGD1561319	Pcbp1	Pol
Opa1	Opa1	optic atrophy 1 homolog
RGD1562629	Nbea	neurobeachin
LOC499390	LOC499390	SWAP2
RGD1565122	Ttc21b	tetratricopeptide repeat domain 21B
Uba52	Uba52	ubiquitin A-52 residue ribosomal protein fusion product 1
Syngr1	Syngr1	synaptogyrin 1
Dst	Dst	dystonin
Septin9	Septin9	septin 9
Gnb1	Gnbp	guanine nucleotide binding protein, beta 1
Becn1	Becn1	beclin 1
Basp1	Basp1	brain abundant, membrane attached signal protein 1
LOC362306	0610010I23Rik	hypothetical LOC362306
Olfm1	Olfm1	olfactomedin 1
LOC680448	Pcdhgb5	protocadherin gamma subfamily B, 5
Kcnj3	Kcnj3	potassium inwardly-rectifying channel, subfamily J, member 3
Pou4f2	Pou4f2	POU domain, class 4, transcription factor 2
Lypla1	Lypla1	lysophospholipase 1
Sec13l1	Sec13	SEC13-like 1
Pcdh17	Pcdh17	protocadherin 17
Tmem30a	Tmem30a	transmembrane protein 30A
Slc25a4	Slc25a4	solute carrier family 25
LOC680231	Chd9	chromodomain helicase DNA binding protein 9
Syt4	Syt4	synaptotagmin IV
Gnao	Gnbp	guanine nucleotide binding protein, alpha o
Tmeff1	Tmeff1	transmembrane protein with EGF-like and two follistatin-like domains 1
LOC688007	LOC668652	similar to Putative serine/threonine-protein kinase F31E3.2
Ak7	Ak7	adenylate kinase 7
Tmem50a	Tmem50a	transmembrane protein 50A
Tpt1	Tpt1	tumor protein, translationally-controlled 1
Rps4x	Rps4	ribosomal protein S4, X-linked

Scd2	Scd2	stearoyl-Coenzyme A desaturase 2
Cdh2	Cdh2	cadherin 2
RGD1565811	Pcdh10	OL-protocadherin isoform
LOC500420	2610029I01Rik	CG12279-PA
Ap2a2	Apc2	adaptor protein complex AP-2, alpha 2 subunit
Vim	Vim	vimentin
Mor1	Mdh2	malate dehydrogenase, mitochondrial
Olr1273	Olr1273	olfactory receptor 1273
RGD1564887	9130011E15Rik	9130011E15Rik protein
Dynll1	Dynll1	dynein light chain LC8-type 1
Rpl6	Rpl6	ribosomal protein L6
Irgm	Irgm	immunity-related GTPase family, M
Cacng5	Cacng5	calcium channel, voltage-dependent, gamma subunit 5
Suclg1	Suclg1	succinate-CoA ligase, GDP-forming, alpha subunit
Syt6	Syt6	synaptotagmin VI
Usp9x	Usp9x	ubiquitin specific peptidase 9, X chromosome
RGD1562118	Prdx2	peroxiredoxin 2 isoform b
Kcnk9	Kcnk9	potassium channel, subfamily K, member 9
RGD1560170	Med14	cofactor required for Sp1 transcriptional activation, subunit 2, 150kDa
Marcksl1	Marcksl1	MARCKS-like 1
Eif4a2	Eif4a2	eukaryotic translation initiation factor 4A2
Boll	A2bp1	bol, boule-like
Atp5a1	Atp5a1	ATP synthase, H+ transporting, mitochondrial F1 complex, alpha subunit, isoform 1
Fgf13	Fgf13	fibroblast growth factor 13
LOC683125	Rspa	40S ribosomal protein SA
Snrpn	Snrpn	small nuclear ribonucleoprotein N
LOC682988	Med18	mediator of RNA polymerase II transcription, subunit 18 homolog, transcript variant2
Nefl	Nefl	neurofilament, light polypeptide
Mapk1	Mapk1	mitogen activated protein kinase 1
Cox5b	Cox5b	cytochrome c oxidase subunit Vb
EF2		
LOC305181	Ef2	Elongation factor 2
Arl6ip5	Arl6ip5	ADP-ribosylation factor-like 6 interacting protein 5
LOC691644	Myh2	Myosin heavy chain, skeletal muscle, adult 2
Ndufs6	Ubiqfep6	NADH dehydrogenase
Dpysl2	Dpysl2	dihydropyrimidinase-like 2
Ftl1	Ftl1	ferritin light chain 1
Atp2b2	Atp2b2	ATPase, Ca++ transporting, plasma membrane 2
Lhx3	Lhx3	LIM homeobox protein 3
LOC500054	Pot1a	POT1-like telomere end-binding protein
LOC683007	Mia3	melanoma inhibitory activity 3
Dnajc5	Hsp40	Hsp40 homolog, subfamily C, member 5
Kif15	Kif15	kinesin family member 15
Rpl23	Rpl23	ribosomal protein L23
RGD1564725	Alg11	hypothetical protein B230397C21
RGD1309710	Trmt112	RIKEN cDNA 0610038D11
Rps15	Rps15	ribosomal protein S15
Jub	Jub	ajuba homolog
Ccdc65	Ccdc65	coiled-coil domain containing 65
Arhgef7	Arhgef7	Rho guanine nucleotide exchange factor 7
Pklr	Pklr	pyruvate kinase, liver and red blood cell
H2afz	H2a	H2A histone family, member Z
LOC286914	Govn5	putative pheromone receptor
Cdh16	Cdh16	cadherin 16
St13	St13	suppression of tumorigenicity 13

Atp5d	Atp5d	ATP synthase, H+ transporting, mitochondrial F1 complex, delta subunit
Atf3	Atf3	activating transcription factor 3
Olr259	Olfr259	olfactory receptor 259
Fmnl1	Fmnl1	formin-like 1
Dync1i1	Dync1i1	dynein cytoplasmic 1 intermediate chain 1
RGD1308774	Slc25a40	mitochondrial carrier family protein

Pink background = Transcripts dendritically localized with high confidence in mouse only
Blue background = Transcripts dendritically localized with high confidence in rat only
Green background = Transcripts dendritically localized with high confidence in mouse & rat

REFERENCES

1. Koentges, G., *Evolution of anatomy and gene control.* Nature, 2008. **451**(7179): p. 658-63.
2. Wu, C., et al., *Gene set enrichment in eQTL data identifies novel annotations and pathway regulators.* PLoS Genet, 2008. **4**(5): p. e1000070.
3. Wray, G.A., et al., *The evolution of transcriptional regulation in eukaryotes.* Mol Biol Evol, 2003. **20**(9): p. 1377-419.
4. Hovatta, I., et al., *DNA variation and brain region-specific expression profiles exhibit different relationships between inbred mouse strains: implications for eQTL mapping studies.* Genome biology, 2007. **8**(2): p. R25.
5. McClurg, P., et al., *Genomewide association analysis in diverse inbred mice: power and population structure.* Genetics, 2007. **176**(1): p. 675-83.
6. Jakobsson, M., et al., *Genotype, haplotype and copy-number variation in worldwide human populations.* Nature, 2008. **451**(7181): p. 998-1003.
7. Henrichsen, C.N., et al., *Segmental copy number variation shapes tissue transcriptomes.* Nature genetics, 2009. **41**(4): p. 424-9.
8. Cahan, P., et al., *The impact of copy number variation on local gene expression in mouse hematopoietic stem and progenitor cells.* Nature genetics, 2009. **41**(4): p. 430-7.

9. Crino, P.B. and J. Eberwine, *Molecular characterization of the dendritic growth cone: regulated mRNA transport and local protein synthesis.* Neuron, 1996. **17**(6): p. 1173-87.

10. Crino, P., et al., *Presence and phosphorylation of transcription factors in developing dendrites.* Proc Natl Acad Sci U S A, 1998. **95**(5): p. 2313-8.

11. Miyashiro, K., M. Dichter, and J. Eberwine, *On the nature and differential distribution of mRNAs in hippocampal neurites: implications for neuronal functioning.* Proc Natl Acad Sci U S A, 1994. **91**(23): p. 10800-4.

12. Burgin, K.E., et al., *In situ hybridization histochemistry of Ca2+/calmodulin-dependent protein kinase in developing rat brain.* The Journal of neuroscience : the official journal of the Society for Neuroscience, 1990. **10**(6): p. 1788-98.

13. Racca, C., A. Gardiol, and A. Triller, *Dendritic and postsynaptic localizations of glycine receptor alpha subunit mRNAs.* The Journal of neuroscience : the official journal of the Society for Neuroscience, 1997. **17**(5): p. 1691-700.

14. Steward, O., *mRNA localization in neurons: a multipurpose mechanism?* Neuron, 1997. **18**(1): p. 9-12.

15. Dennis, G., Jr., et al., *DAVID: Database for Annotation, Visualization, and Integrated Discovery.* Genome Biol, 2003. **4**(5): p. P3.

16. Zhong, J., T. Zhang, and L.M. Bloch, *Dendritic mRNAs encode diversified functionalities in hippocampal pyramidal neurons.* BMC neuroscience, 2006. **7**: p. 17.

17. Eberwine, J., et al., *Analysis of subcellularly localized mRNAs using in situ hybridization, mRNA amplification, and expression profiling.* Neurochem Res, 2002. **27**(10): p. 1065-77.

18. Poon, M.M., et al., *Identification of process-localized mRNAs from cultured rodent hippocampal neurons.* J Neurosci, 2006. **26**(51): p. 13390-9.

19. Fields, R.D. and K. Itoh, *Neural cell adhesion molecules in activity-dependent development and synaptic plasticity.* Trends in neurosciences, 1996. **19**(11): p. 473-80.

20. Haitina, T., et al., *Expression profile of the entire family of Adhesion G protein-coupled receptors in mouse and rat.* BMC neuroscience, 2008. **9**: p. 43.

21. Cremer, H., et al., *Inactivation of the N-CAM gene in mice results in size reduction of the olfactory bulb and deficits in spatial learning.* Nature, 1994. **367**(6462): p. 455-9.

22. Persohn, E., G.E. Pollerberg, and M. Schachner, *Immunoelectron-microscopic localization of the 180 kD component of the neural cell adhesion molecule N-CAM in postsynaptic membranes.* The Journal of comparative neurology, 1989. **288**(1): p. 92-100.

23. Pollerberg, G.E., M. Schachner, and J. Davoust, *Differentiation state-dependent surface mobilities of two forms of the neural cell adhesion molecule.* Nature, 1986. **324**(6096): p. 462-5.

24. Kohmura, N., et al., *Diversity revealed by a novel family of cadherins expressed in neurons at a synaptic complex.* Neuron, 1998. **20**(6): p. 1137-51.

25. Hynes, R.O., *Cell adhesion: old and new questions.* Trends in cell biology, 1999. **9**(12): p. M33-7.

26. Sharova, L.V., et al., *Database for mRNA half-life of 19 977 genes obtained by DNA microarray analysis of pluripotent and differentiating mouse embryonic stem cells.* DNA research : an international journal for rapid publication of reports on genes and genomes, 2009. **16**(1): p. 45-58.

27. Davis, L., G.A. Banker, and O. Steward, *Selective dendritic transport of RNA in hippocampal neurons in culture.* Nature, 1987. **330**(6147): p. 477-9.

28. Steward, O. and W.B. Levy, *Preferential localization of polyribosomes under the base of dendritic spines in granule cells of the dentate gyrus.* The Journal of neuroscience : the official journal of the Society for Neuroscience, 1982. **2**(3): p. 284-91.

29. Krichevsky, A.M. and K.S. Kosik, *Neuronal RNA granules: a link between RNA localization and stimulation-dependent translation.* Neuron, 2001. **32**(4): p. 683-96.

30. Anderson, P. and N. Kedersha, *RNA granules.* The Journal of cell biology, 2006. **172**(6): p. 803-8.

31. Huang, Y.S., et al., *Facilitation of dendritic mRNA transport by CPEB.* Genes & development, 2003. **17**(5): p. 638-53.

32. Rook, M.S., M. Lu, and K.S. Kosik, *CaMKIIalpha 3' untranslated region-directed mRNA translocation in living neurons: visualization by GFP linkage.* The Journal of neuroscience : the official journal of the Society for Neuroscience, 2000. **20**(17): p. 6385-93.

33. Kiebler, M.A. and G.J. Bassell, *Neuronal RNA granules: movers and makers.* Neuron, 2006. **51**(6): p. 685-90.

34. Ule, J. and R.B. Darnell, *RNA binding proteins and the regulation of neuronal synaptic plasticity.* Curr Opin Neurobiol, 2006. **16**(1): p. 102-10.

35. Wells, D.G., *RNA-binding proteins: a lesson in repression.* The Journal of neuroscience : the official journal of the Society for Neuroscience, 2006. **26**(27): p. 7135-8.

36. St Johnston, D., *Moving messages: the intracellular localization of mRNAs.* Nature reviews. Molecular cell biology, 2005. **6**(5): p. 363-75.

37. Sossin, W.S. and L. DesGroseillers, *Intracellular trafficking of RNA in neurons.* Traffic, 2006. **7**(12): p. 1581-9.

38. Cook, K.B., et al., *RBPDB: a database of RNA-binding specificities.* Nucleic acids research, 2011. **39**(Database issue): p. D301-8.

39. Gehman, L.T., et al., *The splicing regulator Rbfox1 (A2BP1) controls neuronal excitation in the mammalian brain.* Nature genetics, 2011. **43**(7): p. 706-11.

40. Hall, T.M., *Poly(A) tail synthesis and regulation: recent structural insights.* Current opinion in structural biology, 2002. **12**(1): p. 82-8.

41. Hall, T.M., *Multiple modes of RNA recognition by zinc finger proteins.* Current opinion in structural biology, 2005. **15**(3): p. 367-73.

42. Zhang, T., et al., *AU-rich element-mediated translational control: complexity and multiple activities of trans-activating factors.* Biochemical Society transactions, 2002. **30**(Pt 6): p. 952-8.

43. Soller, M., M. Li, and I.U. Haussmann, *Determinants of ELAV gene-specific regulation.* Biochemical Society transactions, 2010. **38**(4): p. 1122-4.

44. Cho, S., et al., *Interaction between the RNA binding domains of Ser-Arg splicing factor 1 and U1-70K snRNP protein determines early spliceosome assembly.* Proc Natl Acad Sci U S A, 2011. **108**(20): p. 8233-8.

45. Min, H., et al., *A new regulatory protein, KSRP, mediates exon inclusion through an intronic splicing enhancer.* Genes & development, 1997. **11**(8): p. 1023-36.

46. Gherzi, R., et al., *A KH domain RNA binding protein, KSRP, promotes ARE-directed mRNA turnover by recruiting the degradation machinery.* Molecular cell, 2004. **14**(5): p. 571-83.

47. Tacke, R. and J.L. Manley, *The human splicing factors ASF/SF2 and SC35 possess distinct, functionally significant RNA binding specificities.* The EMBO journal, 1995. **14**(14): p. 3540-51.

48. Gallego, M.E., et al., *The SR splicing factors ASF/SF2 and SC35 have antagonistic effects on intronic enhancer-dependent splicing of the beta-*

tropomyosin alternative exon 6A. The EMBO journal, 1997. **16**(7): p. 1772-84.

49. Crovato, T.E. and J. Egebjerg, *ASF/SF2 and SC35 regulate the glutamate receptor subunit 2 alternative flip/flop splicing.* FEBS letters, 2005. **579**(19): p. 4138-44.

50. Atlas, R., et al., *The insulin-like growth factor mRNA binding-protein IMP-1 and the Ras-regulatory protein G3BP associate with tau mRNA and HuD protein in differentiated P19 neuronal cells.* Journal of neurochemistry, 2004. **89**(3): p. 613-26.

51. Huntzinger, E. and E. Izaurralde, *Gene silencing by microRNAs: contributions of translational repression and mRNA decay.* Nature reviews. Genetics, 2011. **12**(2): p. 99-110.

52. Ashraf, S.I., et al., *Synaptic protein synthesis associated with memory is regulated by the RISC pathway in Drosophila.* Cell, 2006. **124**(1): p. 191-205.

53. Siegel, G., et al., *A functional screen implicates microRNA-138-dependent regulation of the depalmitoylation enzyme APT1 in dendritic spine morphogenesis.* Nature cell biology, 2009. **11**(6): p. 705-16.

54. Rajasethupathy, P., et al., *Characterization of small RNAs in Aplysia reveals a role for miR-124 in constraining synaptic plasticity through CREB.* Neuron, 2009. **63**(6): p. 803-17.

55. Kosik, K.S. and A.M. Krichevsky, *The message and the messenger: delivering RNA in neurons.* Science's STKE : signal transduction knowledge environment, 2002. **2002**(126): p. pe16.

56. Dweep, H., et al., *miRWalk - Database: Prediction of possible miRNA binding sites by "walking" the genes of three genomes.* Journal of biomedical informatics, 2011.

57. Schratt, G., *microRNAs at the synapse.* Nature reviews. Neuroscience, 2009. **10**(12): p. 842-9.

58. Millan, M.J., *MicroRNA in the regulation and expression of serotonergic transmission in the brain and other tissues.* Current opinion in pharmacology, 2011. **11**(1): p. 11-22.

59. Baudry, A., et al., *miR-16 targets the serotonin transporter: a new facet for adaptive responses to antidepressants.* Science, 2010. **329**(5998): p. 1537-41.

60. Abdelmohsen, K., et al., *miR-375 inhibits differentiation of neurites by lowering HuD levels.* Molecular and cellular biology, 2010. **30**(17): p. 4197-210.

61. Blichenberg, A., et al., *Identification of a cis-acting dendritic targeting element in MAP2 mRNAs.* The Journal of neuroscience : the official journal of the Society for Neuroscience, 1999. **19**(20): p. 8818-29.

62. Mori, Y., et al., *Two cis-acting elements in the 3' untranslated region of alpha-CaMKII regulate its dendritic targeting.* Nature Neuroscience, 2000. **3**(11): p. 1079-84.

63. Bramham, C.R. and D.G. Wells, *Dendritic mRNA: transport, translation and function.* Nature reviews. Neuroscience, 2007. **8**(10): p. 776-89.

64. de Moor, C.H., H. Meijer, and S. Lissenden, *Mechanisms of translational control by the 3' UTR in development and differentiation.* Seminars in cell & developmental biology, 2005. **16**(1): p. 49-58.

65. Andreassi, C. and A. Riccio, *To localize or not to localize: mRNA fate is in 3'UTR ends.* Trends in cell biology, 2009. **19**(9): p. 465-74.

66. Bell, T.J., et al., *Cytoplasmic BK(Ca) channel intron-containing mRNAs contribute to the intrinsic excitability of hippocampal neurons.* Proc Natl Acad Sci U S A, 2008. **105**(6): p. 1901-6.

67. Buckley, P.T., et al., *Cytoplasmic Intron Sequence-Retaining Transcripts Can Be Dendritically Targeted via ID Element Retrotransposons.* Neuron, 2011. **69**(5): p. 877-84.

68. Bell, T.J., et al., *Intron retention facilitates splice variant diversity in calcium-activated big potassium channel populations.* Proc Natl Acad Sci U S A, 2010. **107**(49): p. 21152-7.

69. Kim, J., et al., *Rodent BC1 RNA gene as a master gene for ID element amplification.* Proc Natl Acad Sci U S A, 1994. **91**(9): p. 3607-11.

70. Ono, T., et al., *Genomic organization and chromosomal distribution of rat ID elements.* Genes Genet Syst, 2001. **76**(4): p. 213-20.

71. Barreau, C., L. Paillard, and H.B. Osborne, *AU-rich elements and associated factors: are there unifying principles?* Nucleic acids research, 2005. **33**(22): p. 7138-50.

72. Espel, E., *The role of the AU-rich elements of mRNAs in controlling translation.* Seminars in cell & developmental biology, 2005. **16**(1): p. 59-67.

73. Naumann, A., et al., *A distinct DNA-methylation boundary in the 5'-upstream sequence of the FMR1 promoter binds nuclear proteins and is lost in fragile X syndrome.* American journal of human genetics, 2009. **85**(5): p. 606-16.

74. Sharma, R.P., D.P. Gavin, and D.R. Grayson, *CpG methylation in neurons: message, memory, or mask?* Neuropsychopharmacology : official publication of the American College of Neuropsychopharmacology, 2010. **35**(10): p. 2009-20.

75. Solvsten, C. and A.L. Nielsen, *FMR1 CGG repeat lengths mediate different regulation of reporter gene expression in comparative transient and locus specific integration assays.* Gene, 2011. **486**(1-2): p. 15-22.

76. Turner, J.D., et al., *Transcriptional control of the glucocorticoid receptor: CpG islands, epigenetics and more.* Biochemical pharmacology, 2010. **80**(12): p. 1860-8.

77. Illingworth, R., et al., *A novel CpG island set identifies tissue-specific methylation at developmental gene loci.* PLoS Biol, 2008. **6**(1): p. e22.

78. de Leon-Guerrero, S.D., G. Pedraza-Alva, and L. Perez-Martinez, *In sickness and in health: the role of methyl-CpG binding protein 2 in the central nervous system.* The European journal of neuroscience, 2011. **33**(9): p. 1563-74.

79. Wu, H., et al., *Genome-wide analysis reveals methyl-CpG-binding protein 2-dependent regulation of microRNAs in a mouse model of Rett syndrome.* Proc Natl Acad Sci U S A, 2010. **107**(42): p. 18161-6.

80. Abdo, W.F., et al., *CSF neurofilament light chain and tau differentiate multiple system atrophy from Parkinson's disease.* Neurobiology of aging, 2007. **28**(5): p. 742-7.

81. Tempel, B.L. and D.J. Shilling, *The plasma membrane calcium ATPase and disease.* Sub-cellular biochemistry, 2007. **45**: p. 365-83.

82. Pavlos, N.J., et al., *Quantitative analysis of synaptic vesicle Rabs uncovers distinct yet overlapping roles for Rab3a and Rab27b in Ca2+-triggered exocytosis.* The Journal of neuroscience : the official journal of the Society for Neuroscience, 2010. **30**(40): p. 13441-53.

83. Bruzzone, F., et al., *Expression of the deubiquitinating enzyme mUBPy in the mouse brain.* Brain Res Mol Brain Res, 2008. **1195**: p. 56-66.

84. Li, J.Y., et al., *Distribution and intraneuronal trafficking of a novel member of the chromogranin family, NESP55, in the rat peripheral nervous system.* Neuroscience, 2002. **110**(4): p. 731-45.

85. Magee, J.C., *Dendritic integration of excitatory synaptic input.* Nature reviews. Neuroscience, 2000. **1**(3): p. 181-90.

86. Parrish, J.Z., et al., *Genome-wide analyses identify transcription factors required for proper morphogenesis of Drosophila sensory neuron dendrites.* Genes & development, 2006. **20**(7): p. 820-35.

87. Gao, F.B., et al., *Genes regulating dendritic outgrowth, branching, and routing in Drosophila.* Genes & development, 1999. **13**(19): p. 2549-61.

88. Arimura, N., et al., *Role of CRMP-2 in neuronal polarity.* J Neurobiol, 2004. **58**(1): p. 34-47.

89. Williams, M.E., J. de Wit, and A. Ghosh, *Molecular mechanisms of synaptic specificity in developing neural circuits.* Neuron, 2010. **68**(1): p. 9-18.

90. Kim, S.Y., et al., *Non-clustered protocadherin.* Cell adhesion & migration, 2011. **5**(2): p. 97-105.

91. Michan, S., et al., *SIRT1 is essential for normal cognitive function and synaptic plasticity.* The Journal of neuroscience : the official journal of the Society for Neuroscience, 2010. **30**(29): p. 9695-707.

92. Thaler, J.P., et al., *LIM factor Lhx3 contributes to the specification of motor neuron and interneuron identity through cell-type-specific protein-protein interactions.* Cell, 2002. **110**(2): p. 237-49.

93. Badea, T.C., et al., *Distinct roles of transcription factors brn3a and brn3b in controlling the development, morphology, and function of retinal ganglion cells.* Neuron, 2009. **61**(6): p. 852-64.

94. da Fonseca, R.R., et al., *The adaptive evolution of the mammalian mitochondrial genome.* BMC genomics, 2008. **9**: p. 119.

95. Ballard, J.W. and R.G. Melvin, *Linking the mitochondrial genotype to the organismal phenotype.* Molecular ecology, 2010. **19**(8): p. 1523-39.

96. Reenan, R.A., *Molecular determinants and guided evolution of species-specific RNA editing.* Nature, 2005. **434**(7031): p. 409-13.

97. Raymond, J.L., S.G. Lisberger, and M.D. Mauk, *The cerebellum: a neuronal learning machine?* Science, 1996. **272**(5265): p. 1126-31.

98. Bolognani, F., et al., *Associative and spatial learning and memory deficits in transgenic mice overexpressing the RNA-binding protein HuD.* Neurobiology of learning and memory, 2007. **87**(4): p. 635-43.

99. Shibata, M., et al., *MicroRNA-9 regulates neurogenesis in mouse telencephalon by targeting multiple transcription factors.* The Journal of neuroscience : the official journal of the Society for Neuroscience, 2011. **31**(9): p. 3407-22.

100. Cohen, J.E., et al., *MicroRNA regulation of homeostatic synaptic plasticity.* Proc Natl Acad Sci U S A, 2011. **108**(28): p. 11650-5.

101. Gao, F.B., *Context-dependent functions of specific microRNAs in neuronal development.* Neural development, 2010. **5**: p. 25.

102. Cuperus, J.T., N. Fahlgren, and J.C. Carrington, *Evolution and functional diversification of MIRNA genes.* The Plant cell, 2011. **23**(2): p. 431-42.

103. Yuva-Aydemir, Y., et al., *MicroRNA-9: Functional evolution of a conserved small regulatory RNA.* RNA biology, 2011. **8**(4): p. 557-64.

104. Bonev, B., A. Pisco, and N. Papalopulu, *MicroRNA-9 reveals regional diversity of neural progenitors along the anterior-posterior axis.* Developmental cell, 2011. **20**(1): p. 19-32.

105. Darnell, D.K., et al., *MicroRNA expression during chick embryo development.* Developmental dynamics : an official publication of the American Association of Anatomists, 2006. **235**(11): p. 3156-65.

106. Kapsimali, M., et al., *MicroRNAs show a wide diversity of expression profiles in the developing and mature central nervous system.* Genome biology, 2007. **8**(8): p. R173.

107. Shibata, M., et al., *MicroRNA-9 modulates Cajal-Retzius cell differentiation by suppressing Foxg1 expression in mouse medial pallium.* The Journal of neuroscience : the official journal of the Society for Neuroscience, 2008. **28**(41): p. 10415-21.

108. Crino, J.E.a.P., *Analysis of mRNA Populations from Single Live and Fixed Cells of the Central Nervous System.* Current Protocols in Neuroscience, 2001. **5**(5.3).

109. Glanzer, J., et al., *RNA splicing capability of live neuronal dendrites.* Proc Natl Acad Sci U S A, 2005. **102**(46): p. 16859-64.

110. Paradies, M.A. and O. Steward, *Multiple subcellular mRNA distribution patterns in neurons: a nonisotopic in situ hybridization analysis.* J Neurobiol, 1997. **33**(4): p. 473-93.

111. Deschenes-Furry, J., N. Perrone-Bizzozero, and B.J. Jasmin, *The RNA-binding protein HuD: a regulator of neuronal differentiation, maintenance and plasticity.* BioEssays : news and reviews in molecular, cellular and developmental biology, 2006. **28**(8): p. 822-33.

112. Mattick, J.S., *RNA regulation: a new genetics?* Nature reviews. Genetics, 2004. **5**(4): p. 316-23.

113. Pascale, A., M. Amadio, and A. Quattrone, *Defining a neuron: neuronal ELAV proteins.* Cellular and molecular life sciences : CMLS, 2008. **65**(1): p. 128-40.

114. Hamilton, R.S. and I. Davis, *Identifying and searching for conserved RNA localisation signals.* Methods in molecular biology, 2011. **714**: p. 447-66.

115. Washietl, S., *Sequence and structure analysis of noncoding RNAs.* Methods in molecular biology, 2010. **609**: p. 285-306.

116. Keene, J.D., *RNA regulons: coordination of post-transcriptional events.* Nature reviews. Genetics, 2007. **8**(7): p. 533-43.

117. Abdelmohsen, K., et al., *Phosphorylation of HuR by Chk2 regulates SIRT1 expression.* Molecular cell, 2007. **25**(4): p. 543-57.

118. Gao, J., et al., *A novel pathway regulates memory and plasticity via SIRT1 and miR-134.* Nature, 2010. **466**(7310): p. 1105-9.

119. Ashraf, S.I. and S. Kunes, *A trace of silence: memory and microRNA at the synapse.* Curr Opin Neurobiol, 2006. **16**(5): p. 535-9.

120. Sempere, L.F., et al., *Expression profiling of mammalian microRNAs uncovers a subset of brain-expressed microRNAs with possible roles in murine and human neuronal differentiation.* Genome biology, 2004. **5**(3): p. R13.

121. Vreugdenhil, E. and E. Berezikov, *Fine-tuning the brain: MicroRNAs.* Frontiers in neuroendocrinology, 2010. **31**(2): p. 128-33.

122. Saba, R. and G.M. Schratt, *MicroRNAs in neuronal development, function and dysfunction.* Brain Res Mol Brain Res, 2010. **1338**: p. 3-13.

123. Jing, Q., et al., *Involvement of microRNA in AU-rich element-mediated mRNA instability.* Cell, 2005. **120**(5): p. 623-34.

124. Kye, M.J., et al., *Somatodendritic microRNAs identified by laser capture and multiplex RT-PCR.* RNA, 2007. **13**(8): p. 1224-34.

125. Beveridge, N.J., et al., *Dysregulation of miRNA 181b in the temporal cortex in schizophrenia.* Hum Mol Genet, 2008. **17**(8): p. 1156-68.

126. Routh, B.N., et al., *Anatomical and electrophysiological comparison of CA1 pyramidal neurons of the rat and mouse.* J Neurophysiol, 2009. **102**(4): p. 2288-302.

127. Frick, K.M., E.T. Stillner, and J. Berger-Sweeney, *Mice are not little rats: species differences in a one-day water maze task.* Neuroreport, 2000. **11**(16): p. 3461-5.

128. Snyder, J.S., et al., *Adult-born hippocampal neurons are more numerous, faster maturing, and more involved in behavior in rats than in mice.* J Neurosci, 2009. **29**(46): p. 14484-95.

129. Asaka, Y., et al., *Hippocampal synaptic plasticity is impaired in the Mecp2-null mouse model of Rett syndrome.* Neurobiology of disease, 2006. **21**(1): p. 217-27.

130. Satterfield, T.F., S.M. Jackson, and L.J. Pallanck, *A Drosophila homolog of the polyglutamine disease gene SCA2 is a dosage-sensitive regulator of actin filament formation.* Genetics, 2002. **162**(4): p. 1687-702.

131. Beretta, F., et al., *NSF interaction is important for direct insertion of GluR2 at synaptic sites.* Molecular and cellular neurosciences, 2005. **28**(4): p. 650-60.

132. Sakurai, M., et al., *Ubiquitin C-terminal hydrolase L1 regulates the morphology of neural progenitor cells and modulates their differentiation.* Journal of cell science, 2006. **119**(Pt 1): p. 162-71.

133. Correia, S.S., et al., *Motor protein-dependent transport of AMPA receptors into spines during long-term potentiation.* Nature Neuroscience, 2008. **11**(4): p. 457-66.

134. Elden, A.C., et al., *Ataxin-2 intermediate-length polyglutamine expansions are associated with increased risk for ALS.* Nature, 2010. **466**(7310): p. 1069-75.

135. Pountney, D.L., et al., *NSF, Unc-18-1, dynamin-1 and HSP90 are inclusion body components in neuronal intranuclear inclusion disease identified by anti-SUMO-1-immunocapture.* Acta neuropathologica, 2008. **116**(6): p. 603-14.

136. Gong, B. and E. Leznik, *The role of ubiquitin C-terminal hydrolase L1 in neurodegenerative disorders.* Drug news & perspectives, 2007. **20**(6): p. 365-70.

137. Miyata, M., et al., *A role for myosin Va in cerebellar plasticity and motor learning: a possible mechanism underlying neurological disorder in myosin Va disease.* The Journal of neuroscience : the official journal of the Society for Neuroscience, 2011. **31**(16): p. 6067-78.

138. Lein, E.S., et al., *Genome-wide atlas of gene expression in the adult mouse brain.* Nature, 2007. **445**(7124): p. 168-76.

139. Mikula, S., J.M. Stone, and E.G. Jones, *BrainMaps.org - Interactive High-Resolution Digital Brain Atlases and Virtual Microscopy.* Brains, minds & media : journal of new media in neural and cognitive science and education, 2008. **3**: p. bmm1426.

140. Blackshear, P.J., *Tristetraprolin and other CCCH tandem zinc-finger proteins in the regulation of mRNA turnover.* Biochemical Society transactions, 2002. **30**(Pt 6): p. 945-52.

141. Craig, J.M., W.C. Earnshaw, and P. Vagnarelli, *Mammalian centromeres: DNA sequence, protein composition, and role in cell cycle progression.* Experimental cell research, 1999. **246**(2): p. 249-62.

142. Pungaliya, P., et al., *TOPORS functions as a SUMO-1 E3 ligase for chromatin-modifying proteins.* Journal of proteome research, 2007. **6**(10): p. 3918-23.

143. Hardy, C.F., L. Sussel, and D. Shore, *A RAP1-interacting protein involved in transcriptional silencing and telomere length regulation.* Genes & development, 1992. **6**(5): p. 801-14.

144. Michel, B., P. Komarnitsky, and S. Buratowski, *Histone-like TAFs are essential for transcription in vivo.* Molecular cell, 1998. **2**(5): p. 663-73.

145. Nabetani, A., et al., *A conserved protein, Nuf2, is implicated in connecting the centromere to the spindle during chromosome segregation: a link*

between the kinetochore function and the spindle checkpoint. Chromosoma, 2001. **110**(5): p. 322-34.

146. Kasper, S. and R.J. Matusik, *Rat probasin: structure and function of an outlier lipocalin.* Biochimica et biophysica acta, 2000. **1482**(1-2): p. 249-58.

147. Moldovan, G.L., et al., *DNA polymerase POLN participates in cross-link repair and homologous recombination.* Molecular and cellular biology, 2010. **30**(4): p. 1088-96.

148. Frontini, M., et al., *NF-Y recruitment of TFIID, multiple interactions with histone fold TAF(II)s.* The Journal of biological chemistry, 2002. **277**(8): p. 5841-8.

149. Mantovani, R., *The molecular biology of the CCAAT-binding factor NF-Y.* Gene, 1999. **239**(1): p. 15-27.

150. Buchhalter, J.R. and M.A. Dichter, *Electrophysiological comparison of pyramidal and stellate nonpyramidal neurons in dissociated cell culture of rat hippocampus.* Brain Res Bull, 1991. **26**(3): p. 333-338.

151. Viereck, C., et al., *Phylogenetic conservation of brain microtubule-associated proteins MAP2 and tau.* Neuroscience, 1988. **26**(3): p. 893-904.

152. Tucker, R.P., L.I. Binder, and A.I. Matus, *Neuronal microtubule-associated proteins in the embryonic avian spinal cord.* The Journal of comparative neurology, 1988. **271**(1): p. 44-55.

153. Tucker, R.P., *The roles of microtubule-associated proteins in brain morphogenesis: a review.* Brain research. Brain research reviews, 1990. **15**(2): p. 101-20.

154. Mathworks, *MATLAB.*

155. Christensen, M., et al., *Recombinant Adeno-Associated Virus-Mediated microRNA Delivery into the Postnatal Mouse Brain Reveals a Role for miR-134 in Dendritogenesis in Vivo.* Frontiers in neural circuits, 2010. **3**: p. 16.

156. Magill, S.T., et al., *microRNA-132 regulates dendritic growth and arborization of newborn neurons in the adult hippocampus.* Proc Natl Acad Sci U S A, 2010. **107**(47): p. 20382-7.

157. Smit, A., Hubley, R and Green, P. *RepeatMasker Open-3.0.* 1996-2010; Available from: http://www.repeatmasker.org.

158. Fujita, P.A., et al., *The UCSC Genome Browser database: update 2011.* Nucleic acids research, 2011. **39**(Database issue): p. D876-82.

Chapter 4

SINGLE-CELL VARIABILITY: FICTION OR REALITY?

4.1 ABSTRACT

Even genetically identical single cells display variability in their gene expression and their responses to external stimuli. Several potential sources could create these variations, such as the inherent stochastic nature of biological processes or external environmental effects, resulting in the production of unique cell-specific proteomes. Because of the complexity of this event and its potential significant impact on cell's phenotype and even on the whole organism's phenotype, investigation of single-cell variability and cell-specific gene expression is becoming very popular in various disciplines. In order to characterize the small amount of mRNAs isolated from a single cell, the RNA must be amplified to the appropriate concentrations for analysis with currently available techniques.

In this study, I first validated, through a controlled single-cell dilution experiment, the reliability of the antisense RNA amplification (aRNA) procedure, regardless of the small amount of starting material and the number of RNA amplification rounds. Then, the combination of microarrays and RNAseq single-cell data from somas that were mechanically isolated from hippocampal neurons reinforced the idea that single-cell variability is biologically real. Additional advanced RNAseq analysis, complemented with *in situ* hybridization experiments,

suggested the presence of different 3’UTR isoforms of a given transcript between cells and within a single cell. We hypothesize that these different transcript isoforms could be involved in the regulation of gene expression and thus could be major players in the foundation of cell-to-cell variability.

4.2 INTRODUCTION

Biological data generated from tissue samples correspond to the average of multiple-thousands of heterogeneous cell populations. Thus, transcriptome measurement from a tissue reflects its average gene expression but cannot be considered a reflection of the state of each individual cell [1, 2]. In many ways, individual cells exhibit a large degree of variability. It has been proposed that this heterogeneity could be the result of stochastic noise in the gene expression of each individual cell. The amplitude and dynamics of gene expression are controlled by various internal and/or external factors, such as gene regulation, transcription rate and genetic or epigenetic factors [3]. Even in homologous tissues or cultures, cells can exhibit different sizes, protein levels and especially the amount of expressed mRNA or microRNA transcripts [4-7]. Additionally, Individual cells may differ functionally based on their localization and might respond differently to identical stimuli [8-10]. Given the large cellular heterogeneity of the CNS, it is clear that single-cell studies are a necessity. As modification in gene expression via the mRNA regulation often causes changes in protein concentrations, examination of the cell transcriptome is important to understand cellular sensitivity to internal and external stimuli.

The aim of this study is to profile the gene expression of individual neurons in order to investigate single-cell variability and the potential involvement of transcriptome markers in these differences. The amount of mRNA within a

single cell is estimated to be between 0.1 and 1 pg, making experimental assays challenging and therefore requiring amplification of the original single-cell sample. Currently, the most commonly used amplification procedures are quantitative reverse transcription (RT) followed by polymerase chain reaction (RT–qPCR) and amplified antisense RNA (aRNA). Variances and deviations in qPCR studies have been reported [11, 12], as this procedure is based on exponentially increasing the amount of nucleic acids. To overcome these issues, I used the aRNA procedure, which is based on T7 RNA polymerase amplification, and consequently, any skewing would be linear [13].

In this chapter, I report on three aspects of assessing single-cell transcriptome variability. First, to benchmark the variability of our aRNA single-cell studies I performed a control experiment to assess the variation associated with the aRNA amplification procedure as a function of the amount of starting material and the number of RNA amplifications. Second, using a combination of microarrays and RNAseq, I demonstrate that variability exists in the transcriptome across single-cell samples. Finally, I discuss preliminary results from a collaborative study with Dr. Miler Lee, where using the data I generated, he carried out computational analysis that suggested that individual cells might differentially use multiple 3'UTR isoforms. We hypothesized that different 3'UTR isoforms may be related to subcellular localization and I tested this idea by carrying out *in situ* hybridization on a small set of candidate genes using multiply labeled 3'UTR isoform probes.

4.3 RESULTS

4.3.1 aRNA amplification procedure shows high repeatability regardless of the amount of starting material and the number of amplification rounds.

In order to test the aRNA procedure, I simulated the mRNA content of a single cell in a controlled fashion and without any potential single-cell variability bias. My approach was based on extracting bulk mRNA from mouse cortex and generating serial dilutions to single-cell RNA levels. I performed a “mouse-control” experiment in which pooled mRNA from mouse cortex was diluted to single-cell RNA levels (0.1, 1, and 10 pg) and amplified in at least 3 replicates for 2 and 4 rounds (see Materials and Methods). The final aRNA products and a sample of non-amplified mRNA, taken from the bulk original mouse cortex mRNA, were assayed on the Affymetrix platform with a total of 17 arrays.

The mouse microarray data showed high correlation between all samples, with average correlation of 0.75. Figure 4.1A-B, illustrates all of the correlations between the 1-pg diluted sample, amplified for 4 rounds, and the other samples. In this example, the average correlation of the samples amplified for 2 rounds is slightly lower than the correlation of those that were amplified for 4 rounds ($r = 0.70$ versus 0.80, respectively). The correlation of 0.78 detected between the 1-pg-diluted sample, amplified 4 rounds, and the original non-amplified bulk mRNA sample (Figure 4.1A) underscores the high level of consistency and reproducibility of this amplification technique.

In conclusion, from the array analysis, regardless of the small amount of starting material and number of amplification rounds (0, 2, or 4 rounds), the aRNA procedure is a very reliable assay to examine a biological sample with limited starting material such as a single cell.

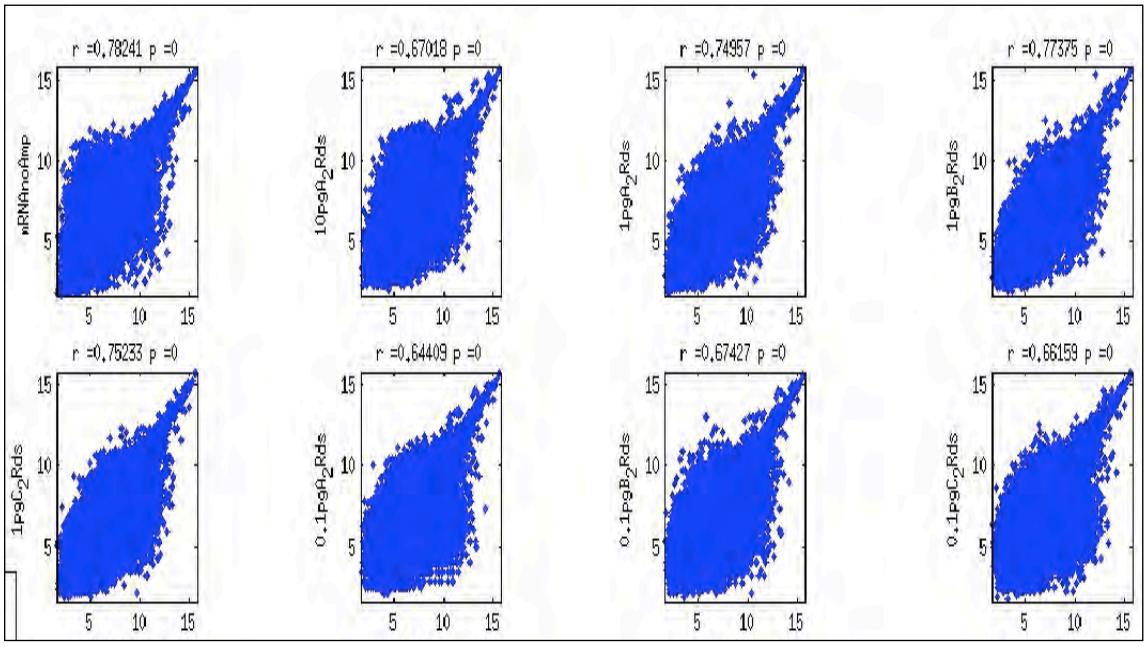


Figure 4.1A: Correlation between a sample of a 1-pg-dilution with 4 rounds of aRNA, on the X-axis, and the other tested samples, on the Y-axis: No aRNA-amplified, bulk mRNA (Plot in the top left corner) and 2 rounds of amplification with serial dilutions (0.1 pg, 1 pg and 10 pg) (All remaining plots).

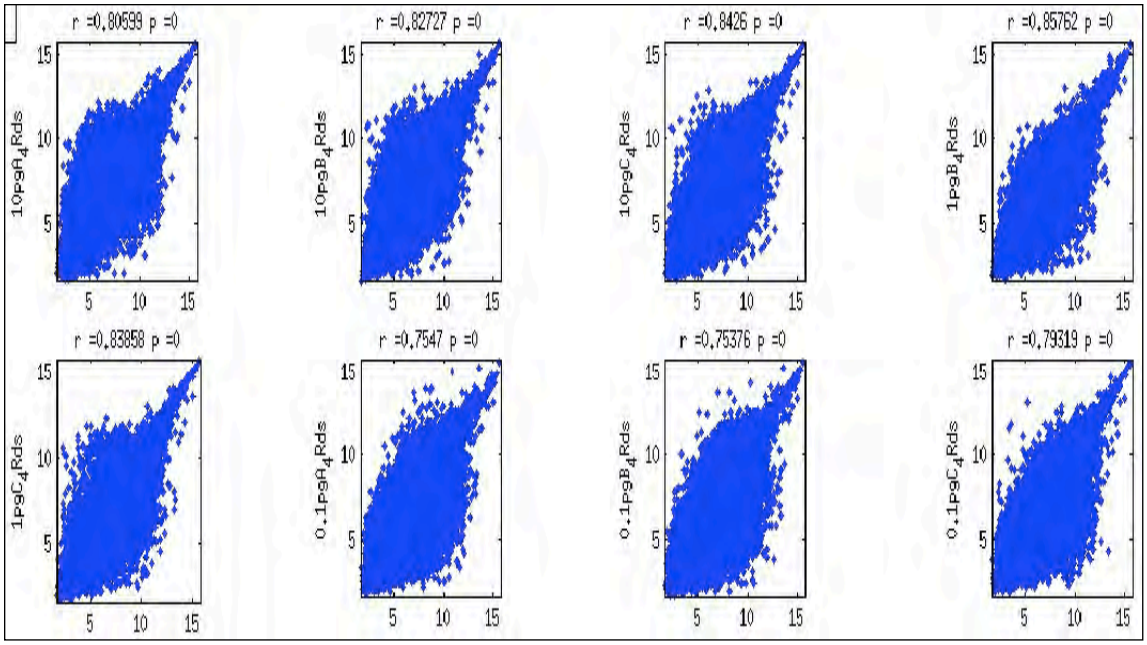


Figure 4.1B: Correlation between a sample of a 1-pg-dilution with 4 rounds of aRNA, on the X-axis, and the other tested samples, on the Y-axis: 4 rounds of amplification with serial dilutions (0.1 pg, 1 pg and 10 pg).

In order to additionally validate the accuracy of the aRNA procedure, I prepared and sequenced (RNAseq, Illumina, see Materials and Methods) a subset (n = 7) of these mouse "control" samples. Also, to account for any potential tissue type or species bias, I performed in parallel a similar "control" experiment using pooled mRNA from rat hippocampus. In this later case, I prepared and sequenced (RNAseq, Illumina) 5 "rat-control" samples (4 samples: aRNA amplified from a 1-pg hippocampal dilution and 1 sample: non-amplified mRNA from the rat hippocampus). The RNAseq data from both rat and mouse is currently being analyzed (Miler Lee, Ph.D.).

4.3.2 Real biological variability exits between the transcriptome of single-cell samples.

After we verified the accuracy of the aRNA procedure for faithfully reproducing the mRNA population of the original sample, we used the control data to assess the variability of the transcriptome of mechanically isolated cell somas from hippocampal neurons in rat and mouse. Samples were collected (n=3 cells per species) and aRNA-amplified for microarray analysis (see Materials and Methods for more details). The overall correlation between these different single cells was 0.56 (with an average of 0.53 between rat samples and 0.59 between mouse samples), which is approximately 25% lower than the average correlation of 0.72 obtained in the control experiment. Figure 4.2 illustrates one example of the correlation plot between two isolated somas in rat (r = 0.47).

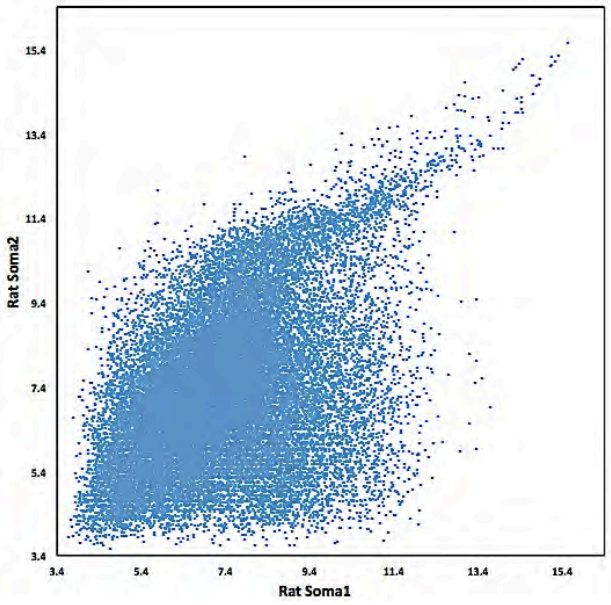


Figure 4.2: Correlation plot between gene expression (with a log2 scale) of two isolated rat somas (r = 0.47).

This outcome demonstrates that variability in RNA expression exists at the neuronal single-cell level, as described previously by Sul J.Y. et al., 2009 [14]. Additionally, our results show no significant difference in the variability amongst rat cells and mouse cells; that is, at this level of samples, there does not seem to be species-specific effects for level of single-cell variability. To assess the single-cell variability of individual genes, we computed the coefficient of variation (CV) for each gene for our single cell samples and other control samples. The ratio of the coefficient of variation of biological and control was computed and tested for significant difference using a F-test. With a FDR cutoff of 5%, we found 1134 genes significantly more variable for the biological single-cell samples and no genes that were significantly more variable in the control data (See Figure 4.3).

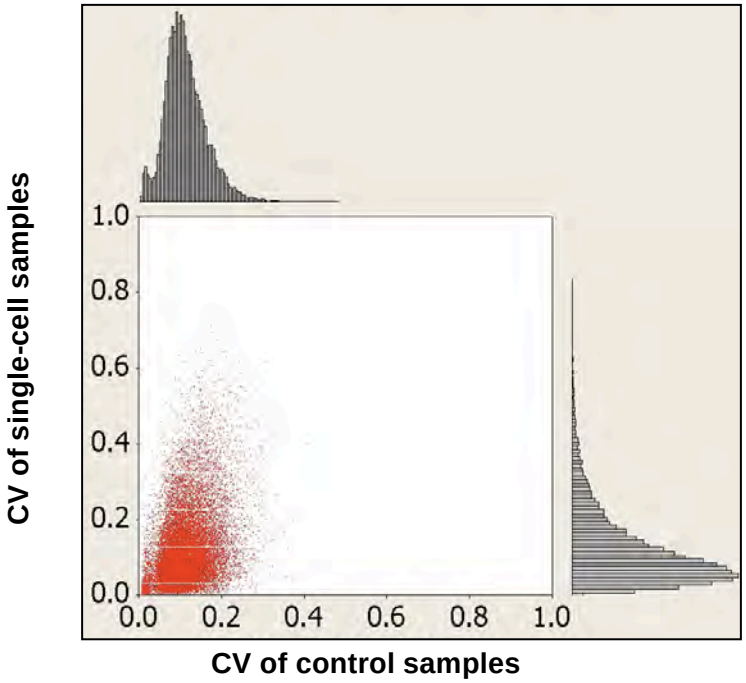


Figure 4.3: Relationship between the coefficient of variation (CV) of the biological single-cell samples and the control samples.

We next tested the potential level of divergence in the soma transcriptome between rat and mouse, as well as between two mouse strains (C57/BL6 and BALB/c). With the same approach described in Chapter 2, these soma results did not show any significant divergence either between or within species (ANOVA with FDR < 0.1%). This outcome strengthened our previous claim in Chapter 2 that transcriptome divergence is particularly captured at the subcellular level.

4.3.3 Polyadenylation variability: a potential signature of single-cell variability

To improve the characterization of the real biological variability detected at the single-cell level in the microarray analysis and to explore its foundation, I carried out single-cell RNAseq (see Materials and Methods), as this method has been shown to provide more robust and sensitive quantification of small sample sizes such as single cells. I mechanically isolated additional soma from rat hippocampal primary cultures (see Material and Methods), aRNA amplified them and prepared them for Illumina sequencing. A preliminary analysis of the RNAseq data from these single cells by Dr. Miler Lee suggested potential for single-cell level 3'UTR isoform variability (described briefly in this section). Based on these findings, I carried out *in situ* validation experiments for 3' UTR isoform use (described in the next section below).

Many investigations have shown that the same genes can exhibit a large amount of variability in their length forms and that these variances could have broad functional impact on the cell in localization, activation of immune cell

receptors and microRNA silencing [15-20]. Particularly in neurons, several transcripts, such as BDNF, Gria2 or IMPA1, have been widely investigated and demonstrated to carry different polyadenylation sites [17, 18, 21]. The resulting various transcript isoforms with various 3'UTR lengths were shown to have a critical functional impact on the cell. In particular, long isoforms were shown to favor transcript localization to dendrites.

Analysis of three sequenced soma samples from the rat hippocampus using an template end mapping strategy showed that ~60% of the genes show 3'UTR ends that are either shorter or longer than the previously annotated sites. In many cases, a single gene showed multiple 3'UTR isoforms. The different sequence alignment patterns, by read depth, are illustrated in Figure 4.4 (reproduced with Dr. Lee's permission).

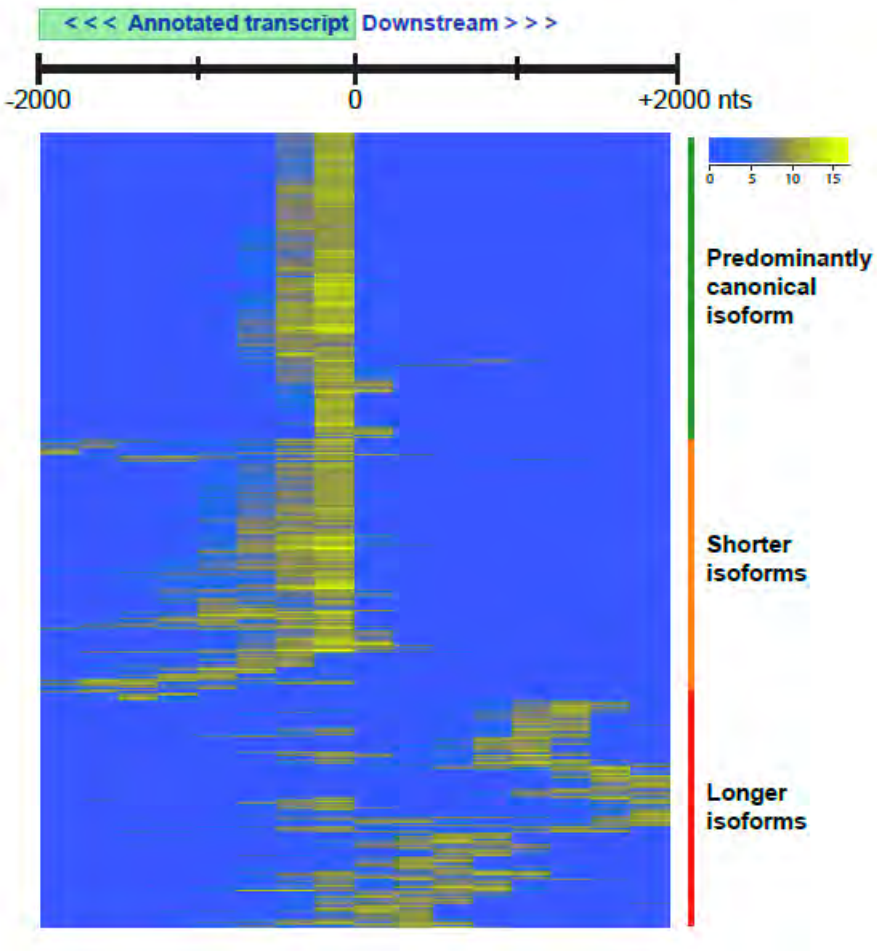


Figure 4.4: RNAseq read depth for 3 single cells showing transcripts with sequence alignment downstream the annotated Refseq boundary. (Scale: Blue (low) to Yellow (high) number of reads).

This finding suggested that within a cell, some transcripts could have different polyadenylation sites in their 3'UTR region and therefore exist in different lengths. The proportion of these transcripts' isoforms is also variable between cells, as 39% of genes vary in expression by at least 32-fold between any two somas, suggesting that there may be cell-to-cell variability in 3'UTR use (Figure 4.5; reproduced with Dr. Lee's permission). Figure 4.5A illustrates the

case of a transcript, the protein tyrosine phosphatase non-receptor type 2 (Ptpn2), showing consistent RNAseq read depths between 3 somas, which implies a single form of this transcript between cells. Figure 4.5B illustrates the case of another transcript, the zinc finger protein 148 (Zfp148), showing different levels of RNAseq read depths within and between 3 somas (Figure 4.5B), which in this case, implies different transcript length forms within and between cells. The transcription factor Zfp148 was shown in recent investigations [22, 23] to regulate certain events in development via transcriptional repression. We speculate that the existence of different isoforms of this transcript might play an important role in its complex regulatory functions.

Further computational and functional experiments are necessary to validate this idea and to validate that the transcripts detected here with different ends are due to real biology and are not sequencing artifacts.

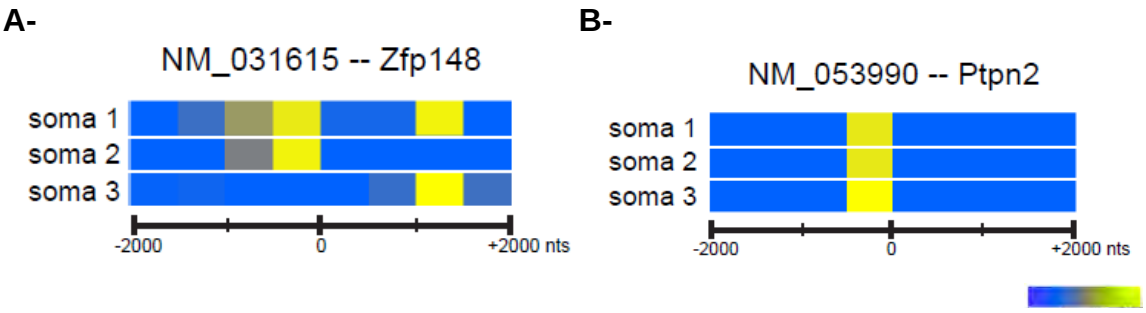


Figure 4.5A-B: Illustration, for a given transcript, of the different patterns of RNAseq reads detected within and between cells (Scale: Blue (low) to Yellow (high) number of reads)

Why do some transcripts carry certain isoforms but not others? The answer to this question could come from various sources. A preliminary way to

address the question could be based on the functional differences between the genes with various isoforms and those without. Indeed, a GO enrichment search showed that the genes with the most variable length forms had ion channel, ion binding, zinc finger protein and protein kinase functions. All of these are critical functions for neurons' excitability, rapid response to stimuli and establishment of synaptic plasticity. Conversely, the least variable genes showed a particular enrichment in mitochondrion, respiration and protein catabolism functions, which are more "housekeeping" functions.

4.3.4 *In situ* experiments toward validating the role of 3'UTR in single-cell variability

I performed *in situ* hybridization on selected candidate genes (Adcy5, CamK2a, CamK2d, Gabra4, Kcna1, Kcnd2, MAP2, Foxp1, Ufm1, Prkcz, Arfrp1 and Rab21) that are likely to possess several 3'UTR forms using differentially labeled probes for the canonical and long isoform (Alexa Fluor dyes 568 and 647, respectively. See Materials and Methods for more details). Based on the raw data, without any image processing, the *in situ* hybridization supports the co-existence of several length forms of a given transcript via a length variation in the 3'UTR region. This case is illustrated in the *in situ* hybridization of the *Ufm1* transcript in Figure 4.6. A basic visual inspection of the cells labeled against the canonical and long 3'UTR forms, clearly finds the presence of both forms of transcripts in neuronal soma (Figure 4.6A). Additionally, overlaying the images of

these probes seems to suggest that these transcripts are differentially distributed within a cell (Figure 4.6B, C1 and C2).

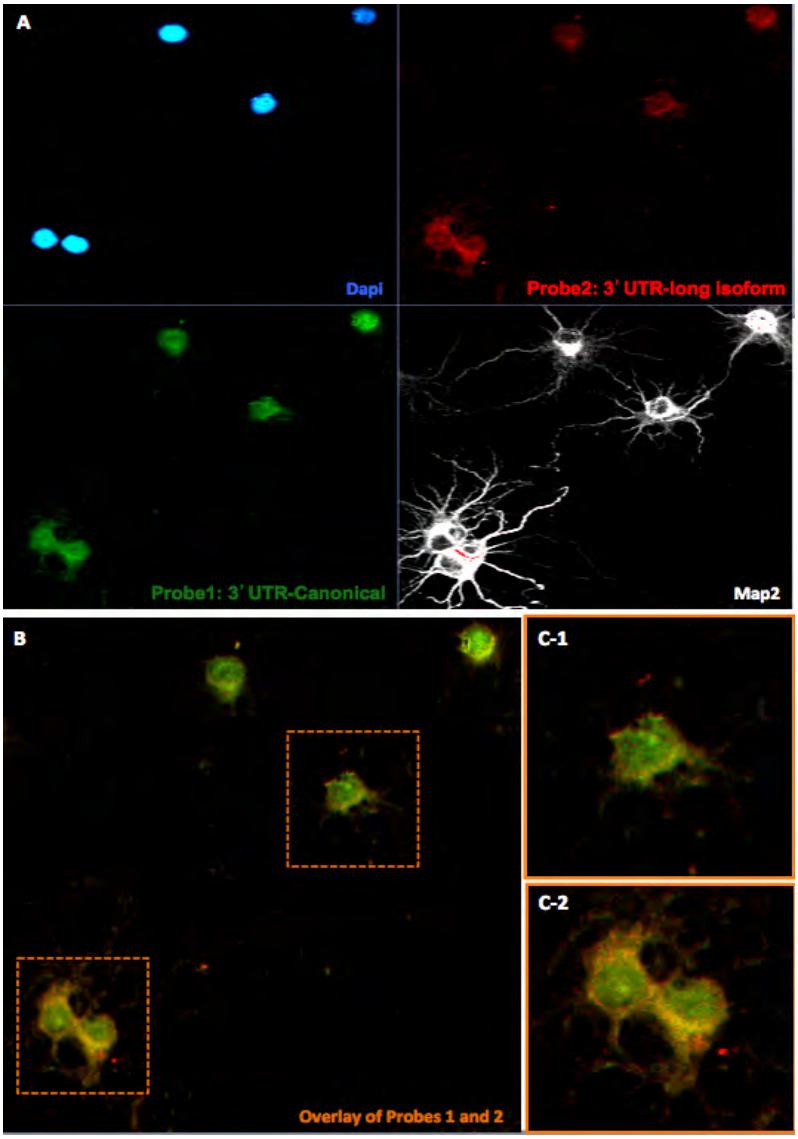


Figure 4.6: *In situ* hybridization reveals the existence of different 3'UTR length forms for a given transcript in a single cell and with different isoforms proportions between cells.

Fluorescent Microscopy evaluation of biotin-conjugated oligoprobes (against Long 3'UTR form) and DIG-conjugated oligoprobes (against canonical 3'UTR) on paraformaldehyde fixed 14-day cultured rat hippocampal neurons. The hybridized biotin-conjugated oligoprobes were detected with Alexa568 and the DIG-conjugated probes with Alexa 647. **(A),** 4 different images section representing DAPI nuclear stain, Map2 as a reference for neuronal cells and their dendritic architecture, Probe1 against the longer 3'UTR isoform of the Ubiquitin-fold modifier1 gene (*Ufm1*) and Probe2 against the canonical 3'UTR of *Ufm1* ; **(B),** Overlay of the Probe 1 and 2; **(C),** C1 and C2 Zoom-in on cells taken from panel **(B)**

Figure 4.7 provides a quantified example (by pixel intensity) contrasting a cell with two transcript isoforms (Figures 4.7-A1 and 4.7-B1) and a control cell (Figures 4.7-A2 and 4.7-B2), especially in relation to the nucleus (DAPI staining).

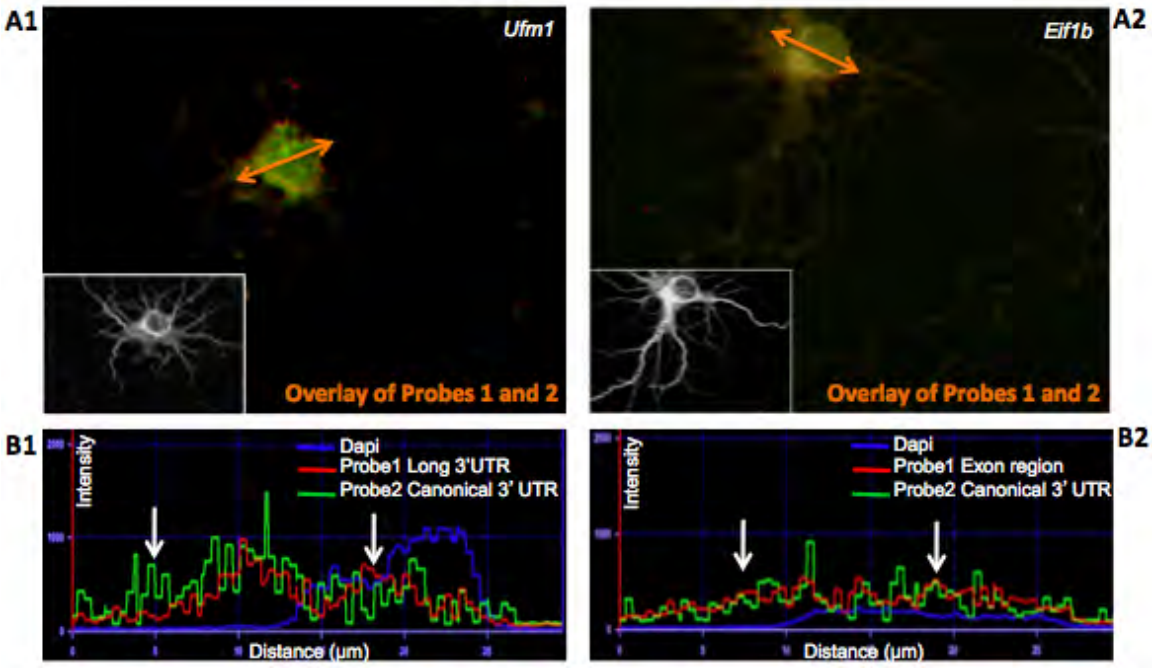


Figure 4.7: *In situ* hybridization highlights differences between control sample and sample having transcript with different isoforms for 3'UTR length.

Fluorescent Microscopy evaluation of biotin-conjugated oligoprobes (against either exonic region or Long 3'UTR form) and DIG-conjugated oligoprobes (against canonical 3'UTR) on paraformaldehyde fixed 14-day cultured rat hippocampal neurons. The hybridized biotin-conjugated oligoprobes were detected with Alexa568 and the DIG-conjugated probes with Alexa647. The bottom left panels correspond to MAP2 staining for neurons architecture.
(A-1), Cell representing the overlay of Probe1 against the longer 3'UTR isoform of the Ubiquitin-fold modifier1 gene (*Ufm1*) and Probe2 against the canonical 3'UTR of *Ufm1*
(A-2), Cell representing the overlay of Probe1 against the exonic region of Eukaryotic translation initiation factor 1b (*Eif1b)* transcript and Probe2 against the canonical 3'UTR of *Eif1b*
(B-1), Trace of pixel intensity throughout the cell soma in A-2. The white arrows highlight areas with different distribution of the transcript labeled with probes 1 and 2.
(B-2), Trace of pixel intensity throughout the cell soma in A-1. The white arrows highlight areas with similar distribution of the transcript labeled with probes 1 and 2.
The tracing was done on the area marked by the orange arrow (30µm distance) via ZEN software, Zeiss Inc.

As a preliminary investigation I quantified the *in situ* hybridization signal intensity of four candidate genes (Rab21, Foxp1, Prkc7 and Ufm1). The quantification was done on at least 6 cells per image using the same approach illustrated in Figure 4.7. These results confirmed the presence of different transcript isoforms and suggested a differential distribution of these within and between cells, except for the controls (*Eif1b* and *Coq3*), in which a much more consistent distribution was observed between the labeled canonical 3'UTR and the exonic region (Figure 4.7 B2). These preliminary investigations seem promising.

Within the current data set, I was not able to verify differential expression of the isoforms at the single cell level. This may be due to sample size or due to hybridization specificity of the probes. Clearly, new RNAseq data performed on more soma samples (total n = 10), together with more controlled *in situ* hybridization image processing and analysis, will provide a more robust support of our observations and will allow a better grasp of the reasons behind single-cell variability and their potential impact on cellular function and phenotype.

4.4 DISCUSSION

Uncovering the mechanisms involved in transcript regulation and the control of RNA dynamics of the cell are fundamental aspects that are necessary for a better understanding of single-cell biology. In quantitative single-cell studies, the low amount of nucleic acids recovered could be a critical limiting factor. It has not been clearly established if the variability in the outcome of these studies is due to real biological effects or experimental variations that might occur during the RNA amplification procedure.

In this study, I assayed, using microarray platforms, the mRNA level from single-cell diluted, aRNA-amplified samples, as well as from the original bulk mRNA source. The high correlation in gene expression recovered between the different samples verified that the aRNA amplification procedure is reliable, regardless of the small amount of starting material and number of amplification rounds. Thus, using this non-biased amplification procedure, I showed, through single-cell microarray data, a clear pattern of cell-to-cell variability between neuronal somas. This variability was investigated further through a preliminary analysis of RNAseq data from 3 rat hippocampal somas (in collaboration with Miler Lee, Ph.D.). This analysis, supported by additional *in situ* hybridization experiments, highlighted the presence of different isoforms of a given transcript between cells and within a single cell.

Isoform ratio variability can originate from multiple sources. The ones recovered in this study are most likely due to different polyadenylation sites as

the variability in transcript size reported here was mainly based on differential read depths beyond the canonical annotated 3'UTR region. Several previous investigations have revealed that a wide variability exists in gene length forms and showed the functional relevance of this variability. Recently, An and colleagues have shown that in hippocampal neurons, *BDNF* transcripts with short 3'UTR transcripts are specifically restricted to the cell soma while the long 3′UTR transcripts are specifically targeted to dendrites where they regulate dendritic spine morphology [17]. Also, long 3'UTR transcript isoforms could be involved in the translational silencing [21] of some genes, such as the AMPA receptor subunit 2 (*Gria2*, *GluR2*). This silencing could potentially be achieved through microRNA interactions. In fact, this idea was supported by the finding that some mRNA with short 3′UTR regions that do not contain microRNA target sites exhibit an increased level of expression in T-lymphocytes upon activation of the T-cell antigen receptor [15], whereas the long 3′UTR transcripts were associated with lower levels of expression.

Future RNAseq analysis on a larger number of soma samples, together with more advanced *in situ* hybridization experiments, should allow a better understanding of single-cell variability. Particularly, these larger datasets will demonstrate more accurately the spatial and temporal organization of transcript isoforms and will clarify if these are only part of a transitory state of the cell. Finally this diversity in 3′ UTR length may be the basis of regulation of mRNA turnover, translation and subcellular localization by cells, yielding different

phenotypes for the same protein product. Thus, alternative 3'UTR lengths within transcripts could be the signature of unique cell identities.

4.5 MATERIALS AND METHODS

Sample collection for transcriptome analysis

Hippocampi primary cultures from mouse E18 (C57BL/6 and BALB/c strains, Charles River Laboratories, Inc.) and rat E19 (Sprague-Dawley Charles River Laboratories, Inc.) were plated at 100,000 per ml in neurobasal medium (Invitrogen) with B-27 supplement (Sigma) on 12-mm round German Spiegelglas coverslips (Bellco Glass) and grown for 14 days [24]. Mouse and rat embryonic samples used for primary cultures were developmentally matched based on the protocol provided by Charles River Laboratories (http://www.criver.com/SiteCollectionDocuments/rm_rm_d_pregnant_rodent.pdf). These primary cultures allowed 3 single-cell collections from each species. For microarray analysis, 3 somas were isolated from each of these species: C57BL/6 and BALB/c mice strains and Sprague-Dawley rat. For RNAseq analysis, a total of 7 somas were collected (3 and 4 somas extracted from 2 different animal neurons culture dates) from Sprague-Dawley rat.

RNA isolation and microarrays

All samples were assessed through standard aRNA amplification methods, as described previously [25, 26]. After 2 rounds of amplification, a final aRNA amplification was performed with the Ambion Illumina TotalPrep RNA

Amplification kit with an incubation time of 14 h. The integrity of these amplified aRNAs was evaluated with an Agilent Technologies 2100 Bioanalyzer and RNA Nano LabChip. For Affymetrix Rat 230 2.0 and Mouse 430.2 analysis, 5µg of each aRNA was used.

In situ hybridization and imaging

DNA-oligomer probes, with a total of 25 bp length, designed to target the canonical 3'UTR region were biotin-labeled in their 5'end region (Sigma-Genosys®), and other DNA-oligomer probes, with also a total of 25 bp length, designed to target the corresponding long 3'UTR region were either conjugated with Digoxigenin (DIG) or Cy5 in their 5'end region (Integrated DNA Technology, IDT®), (Figure 4.8). Additionally, 2 control probes for eukaryotic translation initiation factor 1b, Eif1b and Coenzyme Q3 homolog methyltransferase, Coq3 gene were designed to target the exon region and were conjugated with Digoxigenin (DIG) in their 5'end region (IDT®). The identity of the transcripts assayed and their probe sequence are listed in Table 4.1. 14 day-old primary rat hippocampal neurons were fixed for 15 minutes in 4% paraformaldehyde, washed in 1X PBS and permeabilized with 0.2% Triton X-100 for 10 min at room temperature (RT). Cells were prehybridized at 36°C with 50% formamide, 1X Denhardt's solution, 4X SSC, 10 mM DTT, 0.1% CHAPS, 0.1% Tween-20, 500 µg/ml yeast tRNA and 500 µg/ml salmon sperm DNA. *In situ* hybridization was performed for 16 h at 42°C with 15 ng/µl probe in prehybridization buffer. After probe hybridization, rabbit anti-MAP2 primary antibody (1:1000) and mouse anti-

DIG (1:200) or mouse anti-Cy5 (1:50) were added to cells for 1 h at RT, followed by addition of secondary antibodies, Alexa 488 goat anti-rabbit antibody (1:750), Alexa 647 goat anti-mouse antibody (1:750) and Alexa 568 streptavidin-conjugated (1:750), for 1 h at RT. The co-staining for MAP2 was performed for two main reasons: First, MAP2 is known to be a marker for dendrites, second MAP2 is conserved in mammals and its expression is known to coincide with the maturation of neuronal morphology, and thus could be used as reference baseline for the maturity of both rat and mouse neurons fixed after 14 days in culture [27-29]. DAPI staining was performed before mounting the slides. The samples were visualized by confocal microscopy (Zeiss Inc., 710 meta, 40x objective with 12 Z-stack layers of 0.25 µm). The collected images were not processed at this point. The raw images presented in this chapter correspond to the overlay of the Z-stack sections.

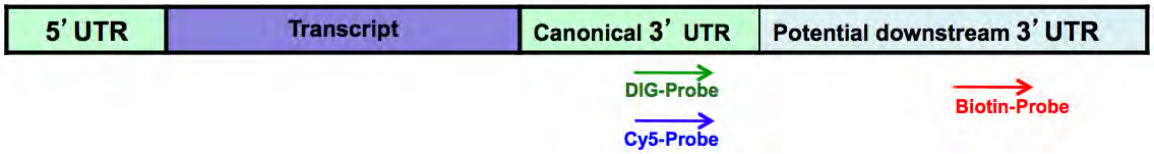


Figure 4.8: Schematic view of the regions targeted by the DNA oligo-probes designed to investigate transcript isoforms.

Table 4.1 List of transcripts investigated by *in situ* hybridization

Probe for 3'UTR	Sequence (5'-3')
AdCy5-Long	CCATGGGAAATCATAATTGTATCAT
CamK2a-Long	AAACACCAAGAATAAAACAGACGTC
CamK2d-Long	CTCACATCTTAGTGCAGAAGACAAA
Gabra4-Long	AGATTCTTTCTATAAGATGGGCGTT
Kcna1-Long	CTTTCTCAGCATGCGATTATAAAAT
Kcnd2-Long	TCAGTCACCACCAACACAGTATATT
MAP2-Long	CAATCTACATGTCCAGGCTAAAACT
Foxp1-Long	CCTGAGGTCAGAACTTAAAATGGT
Ufm1-Long	AAAGGCAGAATGGAAGAAATTTAAT
Prkcz-Long	TGAAGTGTGTCCATATTCAATAGGA
Arfrp1-Long	AATTGAGTAAAGATGGCCAGAAAG
Rab21-Long	TCCAATGAGGAAAGTATGTTTCCTA
Coq3-Long	TCCTCAGCGCTTTCTCATTT
Eif1b-Long	GGAAACATGTCAGGACGGAC
AdCy5-Canonical	ATATTTTCTTTTGCTCAGAGCCACT
CamK2a-Canonical	GTAGGGTGATAGGACAGGGAGAAT
CamK2d-Canonical	CATGGTAAGATGACGTGTCACATAT
Gabra4-Canonical	AATGTGACTGGAAAGAGAATACGAG
Kcna1-Canonical	GTTTCTCGGTGGTAGAAATAGTTGA
Kcnd2-Canonical	TGAAAAGGTAGCAAGACTTTTCACT
MAP2-Canonical	TTACTACAGTTGGGGGATGAGATTA
Foxp1-Canonical	GTTGGCTGTTGTCACTAAGGACAG
Ufm1-Canonical	ATTCCATAATTCAATGTTTGGTTTT
Prkcz-Canonical	GACACAAGAGATTGCTCTGTCTAGA
Arfrp1-Canonical	GTACGACACACTTCACCATCCATT
Rab21-Canonical	CAAAAACAAAAACAGTATGCTGTTT
Coq3-Exon	GCAAACTTCCCTTGCTCATC
Eif1b-Exon	TGGAGGTTCTGGATAGTGGA

Probes were designed to track transcripts with the long, downstream 3'UTR form (Labeled with Biotin, Red); and to track transcripts with the canonical 3'UTR (Labeled with either Cy5, Blue, or DIG, Green). Two controls (Coq3 and Eif1b) were designed as a baseline for tracking 2 different regions of a single transcript isoform: exon region and canonical 3'UTR.

Control Experiment

Adult female mouse brain cortex (C57BL/6, Charles River Laboratories, Inc.) was isolated and stored immediately at -80°C. Subsequently, the mRNA (15 µg) was isolated using TRIzol Reagent and a MicroFastTrack 2.0 Kit (Invitrogen). A sample of 5 µg was assessed on the Affymetrix Mouse 430.2 array. Aliquots from the same cortical mRNA were diluted to single-cell RNA levels (0.1, 1 and 10 pg) and independently amplified, as described above, for a total of 2 and 4 rounds and assessed on Affymetrix Mouse 430.2 arrays.

Rat brain hippocampus (Sprague-Dawley Charles River Laboratories, Inc.) was isolated and stored immediately at -80°C. Subsequently, the mRNA (15 µg) was isolated using TRIzol Reagent and a MicroFastTrack 2.0 Kit (Invitrogen).

As in the mouse control experiment, the samples prepared for RNAseq came form either the original bulk mRNA or from diluted samples of 0.1, 1 and 10 pg (total n = 7). These samples, except for the bulk sample, were aRNA-amplified twice before the RNAseq library construction.

Single-cell transcriptome paired-end sequencing (RNAseq)

The paired-end RNA library was compiled with an Illumina Paired-End DNA Sample Prep Kit (Illumina, San Diego, CA), following the manufacturer's instructions. In brief, double-stranded cDNA was synthesized from polyadenylated RNA and sheared. The approximately 200 base-pair fraction was isolated and amplified with ten cycles of PCR using the Illumina Genome

Analyser paired-end library protocol (Illumina). The resulting libraries were then sequenced on an Illumina Genome Analyzer II, following the manufacturer's instructions (Illumina Inc., San Diego, CA).

Computational Analysis of Single-Cell Transcriptome.

Array quantification

The expression intensities of the probes were summarized using the upper decile statistic with Affymetrix RMA 2.0 methods [14, 30]. All of the arrays were median-centered and scaled by the range of expression values between the 10^{th} and the 90^{th} percentile in each array.

Alignment of lllumina sequencing reads to rat genes

Specific read coverage for the 3 rat single cells sequenced via a paired-end was performed using Bowtie [31], version 0.9.8, using the default parameters for the rat genome v. 3.4 [32].

Paired-end reads were used to define high-confidence regions present in the transcriptome samples, while additional read coverage from unpaired single reads was used to augment the transcriptome maps to mitigate reduced sensitivity from the paired-end analysis on shorter features and lower-complexity sequences [31].

REFERENCES

1. Cai, L., N. Friedman, and X.S. Xie, *Stochastic protein expression in individual cells at the single molecule level.* Nature, 2006. **440**(7082): p. 358-62.
2. Levsky, J.M. and R.H. Singer, *Gene expression and the myth of the average cell.* Trends in cell biology, 2003. **13**(1): p. 4-6.
3. Gao, W., W. Zhang, and D.R. Meldrum, *RT-qPCR based quantitative analysis of gene expression in single bacterial cells.* Journal of microbiological methods, 2011. **85**(3): p. 221-7.
4. Bengtsson, M., et al., *Gene expression profiling in single cells from the pancreatic islets of Langerhans reveals lognormal distribution of mRNA levels.* Genome research, 2005. **15**(10): p. 1388-92.
5. Raj, A., et al., *Stochastic mRNA synthesis in mammalian cells.* PLoS Biol, 2006. **4**(10): p. e309.
6. Zenklusen, D., D.R. Larson, and R.H. Singer, *Single-RNA counting reveals alternative modes of gene expression in yeast.* Nature structural & molecular biology, 2008. **15**(12): p. 1263-71.
7. Gandhi, S.J., et al., *Transcription of functionally related constitutive genes is not coordinated.* Nature structural & molecular biology, 2011. **18**(1): p. 27-34.
8. Stahlberg, A. and M. Bengtsson, *Single-cell gene expression profiling using reverse transcription quantitative real-time PCR.* Methods, 2010. **50**(4): p. 282-8.

9. Raj, A. and A. van Oudenaarden, *Nature, nurture, or chance: stochastic gene expression and its consequences.* Cell, 2008. **135**(2): p. 216-26.

10. Raj, A., et al., *Variability in gene expression underlies incomplete penetrance.* Nature, 2010. **463**(7283): p. 913-8.

11. Bengtsson, M., et al., *Quantification of mRNA in single cells and modelling of RT-qPCR induced noise.* BMC molecular biology, 2008. **9**: p. 63.

12. Taniguchi, K., T. Kajiyama, and H. Kambara, *Quantitative analysis of gene expression in a single cell by qPCR.* Nature methods, 2009. **6**(7): p. 503-6.

13. Eberwine, J., et al., *Analysis of gene expression in single live neurons.* Proc Natl Acad Sci U S A, 1992. **89**(7): p. 3010-4.

14. Sul, J.Y., et al., *Transcriptome transfer produces a predictable cellular phenotype.* Proc Natl Acad Sci U S A, 2009. **106**(18): p. 7624-9.

15. Sandberg, R., et al., *Proliferating cells express mRNAs with shortened 3' untranslated regions and fewer microRNA target sites.* Science, 2008. **320**(5883): p. 1643-7.

16. Nguyen-Chi, M. and D. Morello, *[Aberrant regulation of mRNA 3' untranslated region in cancers and inflammation].* Medecine sciences : M/S, 2008. **24**(3): p. 290-6.

17. An, J.J., et al., *Distinct role of long 3' UTR BDNF mRNA in spine morphology and synaptic plasticity in hippocampal neurons.* Cell, 2008. **134**(1): p. 175-87.

18. Iijima, T., et al., *Hzf protein regulates dendritic localization and BDNF-induced translation of type 1 inositol 1,4,5-trisphosphate receptor mRNA.* Proc Natl Acad Sci U S A, 2005. **102**(47): p. 17190-5.

19. Driever, W. and C. Nusslein-Volhard, *A gradient of bicoid protein in Drosophila embryos.* Cell, 1988. **54**(1): p. 83-93.

20. Ghosh, T., et al., *MicroRNA-mediated up-regulation of an alternatively polyadenylated variant of the mouse cytoplasmic {beta}-actin gene.* Nucleic acids research, 2008. **36**(19): p. 6318-32.

21. Irier, H.A., et al., *Control of glutamate receptor 2 (GluR2) translational initiation by its alternative 3' untranslated regions.* Molecular pharmacology, 2009. **76**(6): p. 1145-9.

22. Guo, G., et al., *Resolution of cell fate decisions revealed by single-cell gene expression analysis from zygote to blastocyst.* Developmental cell, 2010. **18**(4): p. 675-85.

23. Woo, A.J., et al., *Identification of ZBP-89 as a novel GATA-1-associated transcription factor involved in megakaryocytic and erythroid development.* Molecular and cellular biology, 2008. **28**(8): p. 2675-89.

24. Buchhalter, J.R. and M.A. Dichter, *Electrophysiological comparison of pyramidal and stellate nonpyramidal neurons in dissociated cell culture of rat hippocampus.* Brain Res Bull, 1991. **26**(3): p. 333-338.

25. Crino, J.E.a.P., *Analysis of mRNA Populations from Single Live and Fixed Cells of the Central Nervous System.* Current Protocols in Neuroscience, 2001. **5**(5.3).

26. Eberwine, J., *Single-cell molecular biology.* Nature Neuroscience, 2001. **4**: p. 1155-1156.

27. Viereck, C., et al., *Phylogenetic conservation of brain microtubule-associated proteins MAP2 and tau.* Neuroscience, 1988. **26**(3): p. 893-904.

28. Tucker, R.P., L.I. Binder, and A.I. Matus, *Neuronal microtubule-associated proteins in the embryonic avian spinal cord.* The Journal of comparative neurology, 1988. **271**(1): p. 44-55.

29. Tucker, R.P., *The roles of microtubule-associated proteins in brain morphogenesis: a review.* Brain research. Brain research reviews, 1990. **15**(2): p. 101-20.

30. Irizarry, R.A., et al., *Exploration, normalization, and summaries of high density oligonucleotide array probe level data.* Biostatistics, 2003. **4**(2): p. 249-64.

31. Langmead, B., et al., *Ultrafast and memory-efficient alignment of short DNA sequences to the human genome.* Genome biology, 2009. **10**(3): p. R25.

32. Gibbs, R.A., et al., *Genome sequence of the Brown Norway rat yields insights into mammalian evolution.* Nature, 2004. **428**(6982): p. 493-521.

Chapter 5

SUMMARY OF RESULTS, CONCLUSIONS & IMPLICATIONS

5.1 SUMMARY OF RESULTS AND CONCLUSIONS

Throughout the course of the past chapters, by combining both computational and experimental biology, I have described hundreds of common, as well as different, mRNAs localized to rat and mouse dendrites. I showed that a significant difference exists in gene expression in the common subcellularly localized transcripts. I presented evidence that this localization might mediated by post-transcriptional regulatory events, most likely involving complex interactions between cis and trans elements, such as retained intron ID elements, embedded AU rich repeats in the 3'UTR and potential interactions with RBPs and microRNAs. These elements could be acting via common as well as species-specific mechanisms and potentially in a transcript-specific fashion.

First, in chapter 2, the data analysis of microarrays on a collection of isolated dendrites and cell soma from single mouse and rat neurons (in collaboration with Shreedhar Natarajan, Ph.D.) identified more than a thousand mRNAs localized in dendrites. We reported that 43.5% of the orthologous dendritic transcripts between the rat and mouse showed significantly different expression at an FDR correction of 0.1%. In the top 5% of the dendritically

expressed genes in rat and mouse, the overlap of common transcripts was smaller than expected, with only 19% common compared to 66% common transcripts in the top 5% of CNS and non-CNS tissues. Our analysis showed that the dendritic transcriptome of rats and mice are significantly more divergent than other homologous tissues. Furthermore, we found that the intron-retained cis elements, called ID elements, are significantly prevalent in rat and mouse dendrite-localized transcripts. Our findings support the functional implication of ID elements in transcript localization as reported in rat by Buckley *et al.* [1] and highlight an additional evolutionary aspect as the main functionality of these retrotransposons seems conserved between rat and mouse.

Second, in chapter 3, I comprehensively determined, via large scale *in situ* hybridization, the spatial pattern of RNA dendritic localization along with species differences. I conducted this study using a curated list of the 400 most representative dendrite-specific genes in rat and mouse taken from our previous microarrays analysis (Chapter 2). This study not only highlighted that subcellular localization of specific transcripts occurs in a species-specific fashion but also underscored differences in transcript coverage within the primary, secondary and tertiary dendritic branches. The functional classification of the dendritic transcripts highlighted their wide implications in common as well as species-specific cellular functions. This investigation uncovered *cis* and *trans* elements with possible implications in either transcript localization and/or gene expression regulation. Similar to our previous dendrite transcriptome array analysis, I found a high incidence of the intron-retained ID-elements in both rat and mouse.

Additionally, a search for sequence motifs showed a significantly high frequency of the 3'UTR AU-rich repeated elements involved in transcripts stability. Our study also provided a connection to diverse RNA binding proteins and microRNAs that could potentially be involved in mRNA targeting and post-transcriptional regulation via either common or species-specific mechanisms. These are likely to play a major role in the proper development and evolution of complex nervous systems. All data from this study has been compiled in a dendritic mRNA localization database (http://kim.bio.upenn.edu/insitu/public/), which could become a potential public resource and serve as a guide for a broad variety of future investigations that could impact a wide variety of biological fields.

The multi-level approaches I took in the major part of this study (Chapters 2 and 3) to assess the transcriptome in rat and mouse dendrites provided interesting and important new insights, all pointing toward the same main conclusion that subcellular localization of specific transcripts can occur in a species-specific fashion. This diversity in transcript subcellular localization detected in these two closely related species might play a role in their established differences in cognitive features [2-4]. This study supports our idea that evolution of behavior phenotypes might be linked to the evolution of subcellular localization of transcripts. Therefore, a greater understanding of the evolution of transcript regulation and localization may provide new insights into key genes that are likely to be relevant in the evolution and diversification of phenotypes. Another critical feature highlighted in our study (Chapter 2) is that transcriptome differences between species are more discernable at the single

cell and at subcellular dendrite level compared to the whole tissue level. This observation underscored the importance of single cell studies when investigating gene expression regulation, particularly in the brain, where a large degree of cellular heterogeneity exists.

Finally, in the context listed above, I devoted the last chapter (Chapter 4) to the inspection of single cell variability. I first verified the consistency of the aRNA amplification procedure regardless of the small amount of starting material and the numbers of rounds of amplifications. Then, using microarrays and RNAseq on isolated hippocampal neuron somas, I validated the biological variability of the transcriptome across single cell samples and investigated its basis (with the collaboration of Miler Lee, Ph.D.). Preliminary results suggested that individual cells might differentially use multiple 3'UTR isoforms. *In situ* hybridization experiments carried out on candidate genes using different labeled 3'UTR isoform probes supported this idea. This preliminary study supported the need for single cell analysis to allow a high-resolution clarification of the mechanisms of post-transcriptional regulation. Subsequent analysis on a larger sample size is expected to provide more accuracy on the spatial and temporal organization of transcripts isoforms. It will also further explore the contributions of varying 3′UTR length towards the establishment of a cell's unique identity.

5.2 IMPLICATIONS AND FUTURE DIRECTIONS

The results of our comparative analysis on transcript subcellular

localization in mouse and rat neurons suggest several interesting hypothesis with respect to mechanisms of dendritic RNA localization and potential implications in dendritic molecular physiology. Here, I will review some of the main future directions that I believe could be interesting for continued exploration.

5.2.1 Implications and future directions in neuronal molecular functions

The presence of mRNAs in dendrites offers a mechanism for synthesizing the appropriate proteins at the right place and time in response to local extracellular stimuli. This localized translation was shown through a wide variety of studies to be important for synaptogenesis and synaptic plasticity [5, 6]. Comparing and contrasting these studies between different species might allow a better understanding of how these mechanisms are regulated and the key players involved. Many questions could be investigated or arise from these comparative studies. Will an extracellular stimulus generate a similar response in rat and mouse orthologous transcripts in dendrites? If so, will this response be universal between different species? In chapters 2 and 3, I reported species differences in dendritic gene expression of members of multi-gene families. For instance, within the vesicular trafficking RAB family, RAB3 and RAB10 were mainly expressed in the mouse dendritic transcripts while RAB1, 8, 15 and 21 mainly expressed in the rat dendritic transcripts. Similarly the motor protein genes Kinesin and Myosin, where Kinesin 17 and Myosin 5a were highly expressed in mouse dendrites but low in rat, and inversely Kinesin 15 and

Myosin 5b were highly expressed in rat dendrites and low in mouse. A possible future experiment would be to directly compare the expression of theses genes in both species via quantitative PCR. One could also use siRNA silencing to investigate the effect of silencing some of these transcripts in one species versus the other. For example, one would expect that the silencing of Myosin 5a would have a more detrimental effect on mouse neurons function compared to rat and inversely when Myosin 5b are silenced, this would have a more detrimental effect on rat neurons function compared to mouse.

Additionally, understanding the level of conservation of the mechanisms of gene expression regulation could imply their level of functional relevance. Thus, the more a mechanism is conserved the more critical it could be for the cell proper functioning. For instance, this concept could be tested by investigating the knockout effect, in both mouse and rat, of some evolutionary conserved dendritic transcripts reported in this study such as the plasma membrane calcium-transporting ATPase 2 (Atp2b2) shown to be targeted similarly in both rodent (illustrated in Figure 3.8, with similar Dendrite/Soma ratio of 0.09 in rat and mouse). The answers to these questions could provide crucial information on the mechanisms of long-term synaptic plasticity.

Several studies have shown that an elevated calcium concentration in dendrites [7], induced either by neurotropins or synaptic stimulation, could lead to the stabilization of certain localized mRNAs and thereby enhance their translation and potentially impact neuronal plasticity [8, 9]. Based on the differences described in rat and mouse transcriptomes in our study, particularly for the

calcium channel transcripts Cacna1a and Cacna1g (t-test FDR <0.0005, Table 2.S1) one could wonder if a calcium increase, or any other synaptic stimulation reagent, could generate a similar outcome on rat and mouse neurons. Understanding these differences could shed light on key regulatory mechanisms in synaptic plasticity and more general brain functions.

Both our microarrays and *in situ* hybridization studies were carried out in neurons from areas in the brain known to be the very active in learning and memory (i.e. hippocampus and cortex). It would be very interesting to determine if similar species-divergence in patterns of localization are observed in areas of the brain less involved in memory formation such as the hypothalamus. For instance, if no species divergence is detected in the transcriptome of the hypothalamic region, one could speculate that species-specific subcellular localization is functionally meaningful particularly in the context of learning and memory formation and this could be a factor that differentiates cognitive feature in higher order brains. It should be noted that the transcriptome level does not necessarily reflect the proteome level, which reproduces more closely the physiological state of a sample. However, correlation as high as 0.75 has been reported between the proteome and transcriptome state [10] and the phenomenon of dendritic localization itself suggests a functional significance to levels of dendritic RNA.

The different *cis*-motifs (ID elements, ARE4, CpG) and *trans*-factors (RNA binding proteins, RBPs and microRNAs) reported in our study could act more or less simultaneously and in a cooperative fashion toward regulating mRNA

expression in dendrites. This might be the case for the miRNAs miR-9 and miR-134 that could be acting jointly with the RBPs ZFP36, ELAVL2 and SFRS1 to regulate the gene expression of the dendritically localized Sirtuin1 mRNA (Sirt1) (as shown in Tables 3.14 and 3.15) [11, 12]. These potential interactions could potentially occur thanks to the presence, within the Sirt1 transcripts, of complementary *cis*-motifs including the AU-rich elements (ARE4). Future investigations on transcripts like Sirt1 could be focused on understanding how these complex events are orchestrated and interact with each other. For instance, future studies could search for co-localization via *in situ* hybridization using simultaneous labeling for the transcripts and the RBPs of interest. In addition, the application of the PAIR technology [13] could be used to uncover all RBPs that may be interacting with each other. Also, directed mutagenesis or knock-out (KO) experiments toward a transcript known to localize in dendrites or toward a RBP gene, might have a different outcome depending on the species and could result in two different phenotypes. These potential species-specific results might uncover novel pleiotropic functions of genes previously unreported

Are dendritic mRNA populations always the same or are they actively changing depending on the growth stage or the activation state of the neuron? Is there a portion that is consistently localized to dendrites that serve as "housekeeping" genes in dendrites regardless of the cell state and the species (i.e. conserved at all stages and throughout the evolution of organisms)? If so, what is the identity and functionality of these genes? Comparative dendrite transcriptomic experiments, via RNAseq or even microarrays, on different

species and using neurons from several developmental stages (such as embryos, postnatal and adult) should clarify the answers to these questions.

I reported in the *in situ* hybridization study (Chapter 3), that several miRNAs could have either single or multiple gene targets. The synergetic effects of miRNA functions in the brain are poorly understood and represent an understudied area. A future investigation could be to address the complex miRNA-mRNA interactions that could play a role in neurodevelopment. We could hypothesize that certain miRNAs may function together to regulate the expression of a gene during neurogenesis or LTP. This might be the instance of Synaptotagmin 4 (Syt4), important for calcium exocytosis that seemed to be linked to 66 different miRNAs (listed in Table 3.15). In order to test this hypothesis, one could combine computational investigations to uncover the identity of the multiple miRNAs together with siRNA silencing experiments.

5.2.2 Implications and future directions in disease

Two major outcomes emerged from our investigations: The first one showed the divergence in subcellular localization in neurons of two closely related species. The second highlighted single cell variability in the transcriptome. Both results are critical factors to account for when investigating a disease.

Rat and mouse are the two major model organisms used when investigating a behavioral response, learning and memory, neurological diseases and responses to drugs. But is there a rationale behind the choice of one

organism versus the other? Is this choice guided by pure practicability as it is much more challenging to knockout a gene in a rat compared with a mouse? Our study urges the need to carefully choose the animal model depending on the aim of the study. Indeed this choice is critical for better understanding a disease, a behavior, or a given biological question. For instance a recent study on kinases involved in anti-viral and cancer chemotherapy [14] was performed in several tissues of rat and mouse and showed clear differences (2-10 folds) between rat and mouse enzymatic activities. This study highlighted the importance of cautiously selecting animal models for drug discovery and analysis of side effects.

Furthermore, as demonstrated in Chapter 4, single cell analysis could indicate important aspects of transcriptional regulation. Indeed, random fluctuations can cause genetically identical cells to vary resulting in a population of cells with the same genome to have unique proteomes. Small differences in protein abundance may confer a fitness advantage or disadvantage to the cell. Cell-to-cell variability could explain many biological processes such as cell variation in allelic penetrance, responsiveness to stimuli, disease outcome and drugs sensitivity [15, 16]. For instance cell variability has important roles in mammalian cells, such as affecting the outcome of stem-cell differentiation, the latency period of viruses, T-cell activation and the tolerance of cancer cells to chemotherapy [17]. Neurodegenerative diseases, such as Alzheimer Disease (AD), typically affect subpopulations of neurons often lying in close proximity to other neurons unaffected by the disease. Characterizing these diseased cells and identifying the factors that weaken them and make them susceptible to

damage is essential to uncovering the molecular events that underlie neurodegenerative diseases.

A specific feature that arose from our preliminary RNAseq analysis on single cells highlighted the presence of different 3'UTR isoform lengths for some transcripts that could have a functional effect on the cell's phenotype. In fact mRNA isoforms could have different and antagonistic cellular roles resulting in functional consequences from transcriptional variability. Along these lines, one could speculate that, with time and as single cells accumulate more and more variation in the isoform ratio of some transcripts, a new cellular phenotypic could emerge, resulting in the creation of a new disease phenotypic that was not detectable at earlier stages of life. Thus, susceptibility to a disease might simply be the result of a shift in these isoform ratios of a gene. This principle could potentially be involved in the manifestation of complex neurological disorders such as autism, schizophrenia and depression. A recent investigation linking variation in the 3'UTR of MECP2 and autism support our hypothesis stated above [18]. Other studies have associated different isoforms of BDNF transcripts with schizophrenia and epilepsy, providing additional support to our postulation [19, 20]. The use of single cell mRNA amplification techniques such as the aRNA procedure is fundamental in this line of investigation. Future single cell 'omics' investigations combining RNAseq, large scale *in vitro* and *in vivo in situ* hybridization experiments and advanced computational modeling (such as the use of machine learning), might reveal interesting results. Ultimately, variability within single cells might point toward a wide variety of astonishing discoveries.

5.2.3 Implications and future directions in brain's development

Both our microarrays and *in situ* hybridization studies were carried out in 14-day old neuron primary cultures known to correspond to fully mature neurons [21, 22]. However, the exact developmental stage of disassociated neurons and also their ex-vivo maturation is uncertain. This suggests that future investigations on the role of developmental factors in dendritic transcriptome based on comparing gene expression in developing neurons and brain tissues within and between species could be of interest. Several previous studies, performed within a species and based either on RT-PCR, microarrays or RNAseq, have shown developmental differences in the expression of thousands of genes, including many transcription factors [23-25]. Some genes were found selectively expressed at certain stages. For instance in the cerebellum, a helix-loop-helix transcription factor Mth1 is essential for embryogenesis of granule cells [26] and the sonic hedgehog pathway is more involved in postnatal granule cell proliferation [27]; in the cortex, neurogenesis-related genes, such as Sox4, Sox11, and zinc-finger proteins genes, were more highly expressed in mouse embryos (E18) than in postnatal (P7). However, the temporal changes in subcellular transcriptomes and their regulation are unknown. Future studies could be designed either based on different ages of primary neurons cultures such as culture days 7, 14 and 21 or based on different brain tissue age such as E18, P7, P14 and P21. Other factors that could be added in this comparative

study would be the inclusion of several types of brain tissues sources such as hippocampus, cortex, hypothalamus and cerebellum. The outcome of this large-scale analysis would allow us to develop a developmental picture of subcellular RNA processes and also examine their evolutionary differences.

5.2.4 Implications and future directions in brain's evolution

The brain is one the most distinguishing and intriguing organ in mammals. Understanding the evolution of the nervous system is paramount toward understanding brain functionality and diversity. Most investigations done on the brain of the rat or mouse are directed towards revealing mechanisms of brain function and dysfunction that relate to human mental disease or disorders. One potential problem with this bias is that we do not appear to understand, or recognize, the differences in rodent and human brains resulting from the evolutionary processes leading to the existing diverse phenotype in these species. Is it then appropriate to routinely extrapolate findings from laboratory rodent brains to humans?

In this view, more than two decades ago, a study highlighted that the use of animal models in investigating disorders of dopaminergic transmission in humans, such as Parkinson's disease, is erroneous as the morphological and electrophysiological properties of dopaminergic neurons are very different from rodents to monkeys and humans [28]. Another recent example of the potential problem of applying animal model studies to humans was the finding that

extended sleep deprivation is lethal to rats [29]. In contrast subsequent studies found different physiological responses and no lethal effect in response to sleep deprivation in humans and many other animals, including the closely related mouse (reviewed in [30]).

Comparative neuroscience studies have clearly enhanced tremendously our understanding of the basic structure, function and dysfunction of the nervous system and increased our knowledge of the evolution and function of the human brain. These studies started more than a century ago thanks to the work of Hall (1833) [31] and Cajal (1888) [32] and are even more powerful today due to the advance of high-throughput technologies and database generation.

Current investigations of the 3'UTR region of transcripts provide a good illustration of the importance of comparative evolutionary studies. Various studies have shown that the 3'UTR region is involved in coordinating the targeting to dendrites and gene expression regulation of several transcripts [33-38] in various organisms such as *Drosophila*, *Aplysia*, mouse or rat (see Table 1.1 for more details and references). Likewise, the importance of this region was highlighted in our rat and mouse *in situ* hybridization study (Chapter 2). As our investigation suggested common as well as different RBPs targeting the 3'UTR region of transcripts in rat and mouse, this outcome suggested that the functionality of the 3'UTR region might also be influenced by a species-specific effect. Indeed, a recent multi-species comparative analysis performed using the UTR database showed that, in most vertebrates, 3′UTRs are longer than their 5′ counterparts [39, 40]. Thus, even if no functional studies were carried out yet in this region in

primates or human, this evolutionary expansion of the 3′UTR region suggests that it might be involved in translational regulation in higher vertebrates, and that such control mechanisms could be significant in defining differences between species.

With the rapid increase in our knowledge of the complete genomes of many species, the power of bioinformatics and the technical advances in molecular and developmental neuroscience, multi-species comparative studies could lead to significant advances in our understanding of human neurological disorders as well as human brain evolution. The powerful outcomes of future comparative neurosciences studies will reinforce the argument that Dobzhansky stated three decades ago, that is “Nothing in biology makes sense except in the light of evolution” (Dobzhansky, 1973) [41].

REFERENCES

1. Buckley, P.T., et al., *Cytoplasmic Intron Sequence-Retaining Transcripts Can Be Dendritically Targeted via ID Element Retrotransposons.* Neuron, 2011. **69**(5): p. 877-84.
2. Routh, B.N., et al., *Anatomical and electrophysiological comparison of CA1 pyramidal neurons of the rat and mouse.* J Neurophysiol, 2009. **102**(4): p. 2288-302.
3. Frick, K.M., E.T. Stillner, and J. Berger-Sweeney, *Mice are not little rats: species differences in a one-day water maze task.* Neuroreport, 2000. **11**(16): p. 3461-5.
4. Snyder, J.S., et al., *Adult-born hippocampal neurons are more numerous, faster maturing, and more involved in behavior in rats than in mice.* J Neurosci, 2009. **29**(46): p. 14484-95.
5. Steward, O., *mRNA localization in neurons: a multipurpose mechanism?* Neuron, 1997. **18**(1): p. 9-12.
6. Moriya, M. and S. Tanaka, *Prominent expression of protein kinase C (gamma) mRNA in the dendrite-rich neuropil of mice cerebellum at the critical period for synaptogenesis.* Neuroreport, 1994. **5**(8): p. 929-32.
7. Berry, F.B. and I.R. Brown, *CaM I mRNA is localized to apical dendrites during postnatal development of neurons in the rat brain.* Journal of neuroscience research, 1996. **43**(5): p. 565-75.

8. Jacobson, A. and S.W. Peltz, *Interrelationships of the pathways of mRNA decay and translation in eukaryotic cells.* Annual review of biochemistry, 1996. **65**: p. 693-739.
9. Ross, J., *Control of messenger RNA stability in higher eukaryotes.* Trends in genetics : TIG, 1996. **12**(5): p. 171-5.
10. Futcher, B., et al., *A sampling of the yeast proteome.* Molecular and cellular biology, 1999. **19**(11): p. 7357-68.
11. Abdelmohsen, K., et al., *Phosphorylation of HuR by Chk2 regulates SIRT1 expression.* Molecular cell, 2007. **25**(4): p. 543-57.
12. Gao, J., et al., *A novel pathway regulates memory and plasticity via SIRT1 and miR-134.* Nature, 2010. **466**(7310): p. 1105-9.
13. Bell, T.J., et al., *PAIR technology: exon-specific RNA-binding protein isolation in live cells.* Methods in molecular biology, 2011. **683**: p. 473-86.
14. Mirzaee, S., S. Eriksson, and F. Albertioni, *Differences in cytosolic and mitochondrial 5'-nucleotidase and deoxynucleoside kinase activities in Sprague-Dawley rat and CD-1 mouse tissues: implication for the toxicity of nucleoside analogs in animal models.* Toxicology, 2010. **267**(1-3): p. 159-64.
15. Eberwine, J., *Single-cell molecular biology.* Nature Neuroscience, 2001. **4**: p. 1155-1156.
16. Kim, J. and J. Eberwine, *RNA: state memory and mediator of cellular phenotype.* Trends in cell biology, 2010. **20**(6): p. 311-8.

17. Kelz, M.B., et al., *Single-cell antisense RNA amplification and microarray analysis as a tool for studying neurological degeneration and restoration.* Science of aging knowledge environment : SAGE KE, 2002. **2002**(1): p. re1.

18. Coutinho, A.M., et al., *MECP2 coding sequence and 3'UTR variation in 172 unrelated autistic patients.* American journal of medical genetics. Part B, Neuropsychiatric genetics : the official publication of the International Society of Psychiatric Genetics, 2007. **144B**(4): p. 475-83.

19. Lu, B. and K. Martinowich, *Cell biology of BDNF and its relevance to schizophrenia.* Novartis Found Symp, 2008. **289**: p. 119-29; discussion 129-35, 193-5.

20. Tongiorgi, E., L. Domenici, and M. Simonato, *What is the biological significance of BDNF mRNA targeting in the dendrites? Clues from epilepsy and cortical development.* Molecular neurobiology, 2006. **33**(1): p. 17-32.

21. Tucker, R.P., *The roles of microtubule-associated proteins in brain morphogenesis: a review.* Brain research. Brain research reviews, 1990. **15**(2): p. 101-20.

22. Tucker, R.P., L.I. Binder, and A.I. Matus, *Neuronal microtubule-associated proteins in the embryonic avian spinal cord.* The Journal of comparative neurology, 1988. **271**(1): p. 44-55.

23. Matoba, R., et al., *Gene expression profiling of mouse postnatal cerebellar development.* Physiological genomics, 2000. **4**(2): p. 155-164.

24. Han, X., et al., *Transcriptome of embryonic and neonatal mouse cortex by high-throughput RNA sequencing.* Proc Natl Acad Sci U S A, 2009. **106**(31): p. 12741-6.

25. Zhang, Y., et al., *Gene expression profiling in developing human hippocampus.* Journal of neuroscience research, 2002. **70**(2): p. 200-8.

26. Ben-Arie, N., et al., *Math1 is essential for genesis of cerebellar granule neurons.* Nature, 1997. **390**(6656): p. 169-72.

27. Wechsler-Reya, R.J. and M.P. Scott, *Control of neuronal precursor proliferation in the cerebellum by Sonic Hedgehog.* Neuron, 1999. **22**(1): p. 103-14.

28. Kotter, R. and M. Feizelmeier, *Species-dependence and relationship of morphological and electrophysiological properties in nigral compacta neurons.* Progress in neurobiology, 1998. **54**(5): p. 619-32.

29. Rechtschaffen, A. and B.M. Bergmann, *Sleep deprivation in the rat: an update of the 1989 paper.* Sleep, 2002. **25**(1): p. 18-24.

30. Siegel, J.M., *Do all animals sleep?* Trends in neurosciences, 2008. **31**(4): p. 208-13.

31. Hall, M., *On the reflex function of the medulla oblongata and medulla spinalis.* Philosophical Transactions of the Royal Society of London, 1833. **123**: p. 635-665.

32. Garcia-Lopez, P., V. Garcia-Marin, and M. Freire, *The discovery of dendritic spines by Cajal in 1888 and its relevance in the present neuroscience.* Progress in neurobiology, 2007. **83**(2): p. 110-30.

33. An, J.J., et al., *Distinct role of long 3' UTR BDNF mRNA in spine morphology and synaptic plasticity in hippocampal neurons.* Cell, 2008. **134**(1): p. 175-87.

34. Yudin, D., et al., *Localized regulation of axonal RanGTPase controls retrograde injury signaling in peripheral nerve.* Neuron, 2008. **59**(2): p. 241-52.

35. de Moor, C.H., H. Meijer, and S. Lissenden, *Mechanisms of translational control by the 3' UTR in development and differentiation.* Seminars in cell & developmental biology, 2005. **16**(1): p. 49-58.

36. Sandberg, R., et al., *Proliferating cells express mRNAs with shortened 3' untranslated regions and fewer microRNA target sites.* Science, 2008. **320**(5883): p. 1643-7.

37. Irier, H.A., et al., *Control of glutamate receptor 2 (GluR2) translational initiation by its alternative 3' untranslated regions.* Molecular pharmacology, 2009. **76**(6): p. 1145-9.

38. Irier, H.A., et al., *Translational regulation of GluR2 mRNAs in rat hippocampus by alternative 3' untranslated regions.* Journal of neurochemistry, 2009. **109**(2): p. 584-94.

39. Mazumder, B., V. Seshadri, and P.L. Fox, *Translational control by the 3'-UTR: the ends specify the means.* Trends Biochem Sci, 2003. **28**(2): p. 91-8.

40. Pesole, G., et al., *UTRdb and UTRsite: specialized databases of sequences and functional elements of 5' and 3' untranslated regions of*

eukaryotic mRNAs. Update 2002. Nucleic acids research, 2002. **30**(1): p. 335-40.

41. Dobzhansky, T., *Nothing in Biology makes sense except in the light of evolution*, in *American Biology Teacher*1973. p. 125-129.