Distribution Agreement

In presenting this thesis or dissertation as a partial fulfillment of the requirements for an advanced degree from Emory University, I hereby grant to Emory University and its agents the non-exclusive license to archive, make accessible, and display my thesis or dissertation in whole or in part in all forms of media, now or hereafter known, including display on the world wide web. I understand that I may select some access restrictions as part of the online submission of this thesis or dissertation. I retain all ownership rights to the copyright of the thesis or dissertation. I also retain the right to use in future works (such as articles or books) all or part of this thesis or dissertation.

Signature:

_____ _____

ChangHui Pak Date

Mutation of the Conserved Polyadenosine RNA binding protein, *ZC3H14/dNab2*, Impairs Neural Function in *Drosophila* and Humans

By

ChangHui Pak
Doctor of Philosophy

Graduate Division of Biological and Biomedical Science
Genetics and Molecular Biology

Anita H. Corbett, Ph.D. Advisor	Jeremy M. Boss, Ph.D. Committee Member
Kenneth H. Moberg, Ph.D. Advisor	William G. Kelly, Ph.D. Committee Member
Gary J. Bassell, Ph.D. Committee Member	John C. Lucchesi, Ph.D. Committee Member

Accepted:

_____
Lisa A. Tedesco, Ph.D.
Dean of the James T. Laney School of Graduate Studies

_____
Date

Mutation of the Conserved Polyadenosine RNA binding protein, *ZC3H14/dNab2*, Impairs Neural Function in *Drosophila* and Humans

By

ChangHui Pak
B.A., University of Illinois at Urbana-Champaign, 2005

Advisors: Anita H. Corbett, Ph.D. and Kenneth H. Moberg, Ph.D.

An abstract of
A dissertation submitted to the Faculty of the
James T. Laney School of Graduate Studies of Emory University
in partial fulfillment of the requirements for the degree of
Doctor of Philosophy
in Graduate Division of Biological and Biomedical Science
Genetics and Molecular Biology
2011

Abstract

Mutation of the Conserved Polyadenosine RNA binding protein, *ZC3H14/dNab2*, Impairs Neural Function in *Drosophila* and Humans

By ChangHui Pak

Every cell contains the same genetic make up composed of DNA sequence. Yet tight regulation of gene expression occurs such that different tissues of a multicellular organism achieve gene expression appropriate for their functions in cell-type specific, spatial and temporal manners. While this regulation of gene expression can occur at the level of transcription and epigenetic modification, post-transcriptional regulatory mechanisms are also critical. Between transcription and translation, many processes are dedicated to ensure proper processing and maturation of transcripts and thus allow for proper regulation of translation both spatially and temporally. RNA-binding proteins play critical roles in achieving this regulation as highlighted by mutations that give rise to many human diseases. Interestingly, mutations in genes encoding RNA-binding proteins that are ubiquitously expressed and play important roles for overall RNA metabolism in all cell types result in tissue-specific phenotypes. Why certain tissues are more sensitive to defects in general RNA-binding protein functions is unclear. Studies addressing both the developmental and tissue-specific functional characterization of critical RNA-binding proteins will allow for better understanding of human disease biology.

We show for the first time that mutations in the human *ZC3H14* gene lead to non-syndromic autosomal recessive intellectual disability (NS-ARID) and use *Drosophila melanogaster* to model key aspects of the disease. Here we describe the first identification and characterization of dNab2, a *Drosophila* orthologue of ZC3H14/Nab2 class of Cys_3His (CCCH) tandem zinc finger (ZnF) polyadenosine RNA-binding protein. dNab2 is essential for development and required in neurons for normal locomotion and flight. Biochemical and genetic data indicate that dNab2 restricts bulk poly(A) tail length *in vivo*, suggesting that this function may underlie its role in development and disease. Furthermore, we define the role of dNab2 in controlling locomotor activity and memory formation in the *Drosophila* mushroom bodies (MBs), a highly specialized structure involved in higher cognitive functions and locomotion. Finally, using a genetic modifier screen, we identify putative dNab2 targets and/or interacting proteins that modulate dNab2-mediated neuronal function. These studies reveal a conserved requirement for *ZC3H14/dNab2* in the metazoan nervous system and allow for future studies on the molecular mechanisms underlying ZC3H14-associated human intellectual disability.

Mutation of the Conserved Polyadenosine RNA binding protein, *ZC3H14/dNab2*, Impairs Neural Function in *Drosophila* and Humans

By

ChangHui Pak
B.A., University of Illinois at Urbana-Champaign, 2005

Advisors: Anita H. Corbett, Ph.D. and Kenneth H. Moberg, Ph.D.

A dissertation submitted to the Faculty of the
James T. Laney School of Graduate Studies of Emory University
in partial fulfillment of the requirements for the degree of
Doctor of Philosophy
in Graduate Division of Biological and Biomedical Science
Genetics and Molecular Biology
2011

Acknowledgements

During my graduate training here at Emory, I was blessed with opportunities to meet and work with many great scientists across all disciplines. They basically shaped who I am now as a scientist. First and foremost, I want to thank my co-advisors, Anita and Ken. I believe it was fate that we three were able to work together on this rather risk-taking joint project. Why had Anita walk down the hall to Ken and started working on *Drosophila* couple months before I was looking for a new lab to join and why she published the paper on *ZC3H14* right at the time Kuss' group had found patients with mutations in *ZC3H14* are all too exact in timing to say that it was just all serendipity. Because this project was so new, it was very exciting as we made new discoveries. I have to thank Anita for initiating this project and for her courage to venture out and try completely new things. I also thank Ken for taking his risk and serving as a fantastic collaborator, spending his own resources and time to get this project up and running. I thank them both for taking me in under their wings and guiding me throughout the last three years. I remember these long meetings with them every week to discuss about results and experiments to do and walking out of Ken's office feeling overwhelmed. Those moments were actually the best times for me and motivated me to push the project forward. I could not have asked for better mentors. Anita and Ken were always on my side during committee meetings and always looked out for my interest. Both Anita and Ken trained me so well on how to think critically, execute experiments and communicate my research. I will truly benefit from their mentorship, which will impact thinking about how I should run my own lab and treat people.

Next I want to thank my thesis committee members. I especially thank Jerry for all his support from the very beginning of graduate school as he served as an excellent program director. Also he guided me through a very tough time when I had to transition into a new lab. I thank him for handling my situation very well and helping me find my fate – Anita and Ken. I also want to thank Bill and John for their input and encouragement during the committee meetings. They both are just the best committee members to have and I believe that's why they are so popular among students and sit in so many committees. Also, Gary has been a tremendous part of my graduate research as he contributed both to data on our project as well as being such a great mentor throughout. He would let me sit in his office to talk about data, models and finding postdoc positions. He is so supportive and encouraging. I often say the phrase, "I love Gary."

Including Gary, I also thank my collaborators here at Emory: Luciano, Christina, Yue and Sanyal. Our collaborators in Germany, Andreas Kuss and Masoud Garshasbi, also made tremendous contributions to further our understanding of dNab2/ZC3H14 and it has been a fun ride publishing our paper together.

Due to my joint status, I was able to interact with people from two labs. I thank members of the Corbett and Moberg laboratories. I especially thank Sara Leung, Milo and Seth (now a Mobergite) for helping me with experiments and contributing with useful ideas and comments. Since I was physically placed in Moberg lab, I feel more attached to the Mobergites (no offense, Corbettes). I thank the past members, Nate, Carolyn, Sarah, Jeremy and Sharon and the current members, Jacob, Melissa, Danny, John, Rick, and Brian for their support, every day jokes and lunch outings. There's also a tradition in the Moberg lab where upon joining the lab, people get engaged, married and get a dog. I followed that tradition and now I believe I started a new tradition – babies! I look forward to meeting their F2s in the future.

During the past six years in graduate school, I couldn't have done it without support from friends and family. I thank my GMB classmates, Pritty, Jennifer, Lauren, Lydia, Wendy, Adam and Chenying for their support and friendships. Without being in study groups with them, I don't think I could have passed the quals or classes. We are a unique class of people who work way too hard. I also especially thank Chenying and Nancy because they became the closest friends with me. They were around during every details of my life from birthdays to wedding to birth of Erin to every little humps and hurdles I passed throughout the six years. They were there with me when I was happy or sad. I truly enjoyed their friendship and love and hope we can continue our fellowship in the future no matter wherever we end up.

I dedicate this work to my parents in Korea and to my husband and daughter Erin. My parents have supported me from *in utero* until now and will be in the future. I thank their continuous support throughout my life and thank them for allowing me to be educated in a foreign country for so long. They always provide me with the best and want me to succeed. I believe that was a driving force that led me to success. I also thank my husband, Dong Seong. He stood by me throughout my graduate career, from taking the written qual to orals to many committee meetings to defense. He cried and laughed with me on a daily basis and always remained interested in what I do and what I have to say about my research. Science and religion is so much closer that we thought. Finally, I thank Erin for bringing me joy and meaning to what I do. I hope to become a great mom and a proud parent who enjoys and have passion for science. Maybe she will read this thesis some day and actually understand what it means. Last but not least, I give the Man upstairs, who is the source and breath of life, all the glory and honor.

Table of Contents

Figures

Tables

CHAPTER 1

Introduction

Introduction

Every cell contains the same genetic make up composed of DNA sequence. Yet, tight regulation of gene expression occurs such that different tissues of a multicellular organism achieve gene expression appropriate for their functions in cell-type specific and spatiotemporally regulated manners. While this regulation of gene expression can occur at the level of transcription and epigenetic modification, post-transcriptional regulatory mechanisms are also critical. Between transcription and translation, many processes are dedicated to properly process and mature transcripts and allow for proper translation into proteins at the right location within the cell and at the right time throughout development.

RNA-binding proteins play critical roles in achieving these processes as highlighted by mutations that give rise to many human diseases. Recently, we identified that mutations in the evolutionarily conserved Cys_3His (CCCH) tandem zinc finger (ZnF) polyadenosine RNA-binding protein, *ZC3H14*, lead to non-syndromic autosomal recessive intellectual disability (NS-ARID) (Chapter 2). Including this disease, many human mutations in RNA-binding proteins that show ubiquitous expression and play important roles for overall RNA metabolism in all cell types result in tissue-specific phenotypes. Why certain tissues are more sensitive to defects in general RNA-binding protein functions is unclear. In order to address this fundamental biological question, I undertook studies addressing both the developmental and tissue-specific functional characterization of the *Drosophila melanogaster* orthologue termed dNab2 in this dissertation. Our findings allow for better understanding of human disease biology and offer potential therapeutic insights.

I. Post-transcriptional regulation of gene expression in eukaryotes

Eukaryotic gene expression is highly regulated by many different post-transcriptional regulatory mechanisms in order to make proper protein products. Messenger RNAs (mRNAs) are first synthesized as pre-mRNAs in the nucleus that are subject to multiple regulatory steps to achieve proper processing, including 5'-end capping, splicing, 3'-end cleavage and polyadenylation (Figure 1.1). Numerous human diseases have been linked to defects in mRNA processing (Figure 1.1'). Once processed, mRNAs are exported to the cytoplasm for translation into proteins. As opposed to this simplistic description of post-transcriptional mRNA processing events, many recent studies have begun to appreciate the complexities, regulation and cross-talks among different regulatory steps that are traditionally thought to be separate processes (1, 2).

At the core of post-transcriptional gene regulation is a cohort of RNA-binding proteins (RBPs), that coat the mRNAs to form messenger ribonucleoprotein (mRNP) complexes, which guide these transcripts through the precise and intricate processing steps and finally, translation into proteins. These mRNPs, which act as adaptors, contact multiple proteins and machineries that mediate different aspects of mRNA metabolism. It is the unique combination of these factors that associate with mRNAs and the dynamic exchange of the different mRNPs along the given transcript that dictate the ultimate fate of the mRNA in the cytoplasm. By examining the different types of mRNPs a single mRNA associates with during its life cycle, one can appreciate the multiple interactions an mRNA has to go through in order to achieve precise spatial and temporal regulation of its gene expression. However, given the complexity of mRNA processing, nuclear export and translational regulatory mechanisms, not all the players have been identified or their

precise molecular interactions with their target mRNAs and/or interacting proteins. In light of these concerns, the focus of this dissertation lies on investigating how a particular mRNP component, dNab2/ZC3H14, affects processing of mRNAs in the context of the development of a muticellular organism and how mutations in *ZC3H14* lead to NS-ARID in humans.

Dynamic composition and diversity of mRNPs

Messenger RNAs associate with a wide range of RBPs to form distinct mRNPs that mediate and regulate pre-mRNA processing as well as transport, localization, translation and stability of the mature transcripts, in both the nucleus and cytoplasm. An important concept in mRNP composition is how the combination of factors dynamically changes within a given mRNP as the mRNA proceeds through its various stages of life (3). Moreover, depending on the cellular environment, dynamic interactions of mRNPs can be regulated, adding yet another layer of complexity (4, 5).

Coincident with transcription, many of the mRNPs are bound to pre-mRNAs to mediate pre-mRNA splicing, nuclear export and subsequent steps. Many of these mRNPs have overlapping functions, as evident by the coupling of transcription with capping, splicing and polyadenylation, as well as coupling of the different mRNA processing steps, such as splicing and nuclear export. mRNPs also function both in the nucleus and the cytoplasm, as many of the mRNPs shuttle between both subcellular compartments, as they travel with their mRNA. Due to these findings, trying to distinguish mRNP components based on certain mRNA processing steps is no longer meaningful.

To date, there are numerous RBPs that decorate the mRNAs and serve as components of mRNPs, including proteins with different RNA-binding motifs, such as RNA recognition motif (RRMs), RGG boxes, KH domains, zinc fingers and double-stranded RNA binding domains (dsRBDs) (6-8). Different mRNPs can target different positions along the transcript in both sequence-dependent and sequence-independent manners. Some RBPs bind to the 7-methylguanosine cap at the 5'-end of mRNAs while others can bind to the poly(A) tails at the 3'-end of mRNAs (9, 10). Certain mRNPs, such as the exon junction complex (EJC), bind to specific positions of the mRNA, the exon-exon junctions, independent of sequence (11). Most commonly, mRNPs target specific structures or sequences in the untranslated regions (UTRs) at the 5'- and 3'-end of messages. In this chapter, I will focus on a specific set of RBPs, the poly(A)-binding proteins, and their key roles in regulation of polyadenylation and translation.

II. Control of poly(A) tail length as a key step in post-transcriptional regulation of gene expression

One key post-transcriptional mechanism that modulates gene expression is the control of poly(A) tail length. Poly(A) tails of mRNAs play critical roles in regulating gene expression including nuclear export, mRNA stability and translation. In the nucleus, poly(A) tails are first synthesized and added at the 3'UTR of pre-mRNAs by the polyadenylation complex, consisting of the poly(A) polymerase (PAP), cleavage and polyadenylation specificity factor (CPSF), and poly(A) binding protein, nuclear 1 (PABPN1) (12). Once synthesized to a certain length (around 200-250 adenosines in higher eukaryotes), polyadenylated mRNAs are then exported out of the nucleus into the cytoplasm where the length of poly(A) tails are further regulated. Cytoplasmic polyadenylation has been extensively studied in *Xenopus* oocytes during early development (13) and emerging studies reveal that critical regulatory events take place in both vertebrate and invertebrate neurons, associated with synaptic plasticity and memory consolidation (14, 15). Despite this critical regulatory role in neurons, prior to this dissertation work, no polyadenylation defect had been linked directly to a brain disorder.

Nuclear polyadenylation – making the tails

Upon transcription of protein coding genes by RNA polymerase II (Pol II), several protein complexes are recruited to the carboxy-terminal domain (CTD) of the polymerase in order to co-transcriptionally process RNA transcripts (16). The protein complexes are subsequently loaded onto the nascent transcript for the addition of 5'-cap structure, removal of the introns by the splicesosome, and addition of 3'-poly(A) tails

(12, 17, 18). Addition of poly(A) tails involves two steps: cleavage of pre-mRNAs at the 3'-end and subsequent synthesis of poly(A) tails. Specific sequence elements (AAUAAA) that are present near the cleavage site allow the recruitment of a multi-subunit cleavage and polyadenylation complex (12). Average lengths of poly(A)s in budding yeast *S. cerevisiae* and human are about 70 and 250 adenosyl residues, respectively. In all eukaryotic species, poly(A) tails are recognized by poly(A)-binding proteins (Pabs). In humans, the nuclear Pab, PABPN1, is thought to stimulate the processivity of PAP along with CPSF (19, 20) and then terminate processive elongation once the poly(A) tail has reached ~250 nts (21). Once mRNAs are exported out of the nucleus, it is thought that nuclear Pabs are displaced by their cytoplasmic counterparts, PABCs, although the mechanistic details of this molecular displacement are unclear.

Studies in *S. cerevisiae* show a direct link between nuclear polyadenylation and mRNA export, as specific mutations in poly(A) polymerase, Pap1, and the nuclear Pab, Nab2 (Nuclear polyadenylated RNA-binding protein), result in export defects with accumulation of poly(A) RNAs in the nucleus (22-27). In addition, disruption of the nuclear Pabs results in defective poly(A) tail lengths, arguing for the important roles of Pabs in maintaining proper length of poly(A) tails. For example, mutants in the *Drosophila* nuclear Pab, *Pabp2*, show shortening of specific transcripts (28) and the budding yeast *Nab2* mutants lead to hyperadenylated transcripts (25). In mammalian systems, knock-down experiments in primary muscle cells showed that partial loss of PABPN1 results in shortening of bulk poly(A) tails (29). However, it is not yet clear how the length of poly(A) tails is determined through cooperation of these factors.

Cytoplasmic polyadenylation – the role in early development and neuronal function

During early metazoan development, mRNAs in the oocyte are thought to be dormant and repressed from translation until oocyte maturation or fertilization occurs. Several of these repressed mRNAs have short poly(A) tails (~20-40 nts) and only when the poly(A) tails are extended does translation occur. The biochemical mechanisms underlying the regulation of translation through polyadenylation is best described from studies in the *Xenopus* oocytes through the discovery of a sequence-specific RNA-binding protein called cyotoplasmic polyadenylation element-binding protein (CPEB) (13). CPEB recognizes the 3'UTR-residing cytoplasmic polyadenylation element (CPE; with a consensus sequence of UUUUUAU) and is responsible for both translational activation and translational repression through association of multiple factors (15).

Following processing and polyadenylation, mRNAs are exported out of the nucleus into the cytoplasm, where specific CPE-containing mRNAs interact with CPEB and additional factors that remove the poly(A) tails. This CPEB-containing multi-subunit complex includes symplekin, which acts as a platform for the recruitment of factors; poly(A) ribonuclease (PARN), which is a deadenylating enzyme; and germ-line development factor 2 (Gld2), which is a poly(A) polymerase (30-32). Since the biochemical activity of PARN outcompetes Gld2, poly(A) tails are trimmed and remain short. Following a progesterone-based signaling cascade, Aurora A kinase is activated leading to phosphorylation of CPEB at Ser174 (33). Once phosphorylated, PARN is expelled from the mRNP complex, allowing Gld2 to elongate the poly(A) tails (Figure 1.2A).

Translational control can also be achieved by CPEB-Maskin interaction-mediated mechanisms. Maskin is a CPEB-interacting protein that associates with eukaryotic translation initiation factor 4E (eIF4E). Maskin acts as a competitor for eIF4G, which normal binds to eIF4E for indirect recruitment of 40S ribosomal subunit to the 5' end of mRNAs to initiate translation. Hence, binding of Maskin inhibits translation by preventing the recruitment of 40S ribosomal subunit to the mRNA. When poly(A) tails remain short, Maskin interacts with CPEB and eIF4E, thereby inhibiting translation. However, upon phosphorylation of CPEB and elongation of the poly(A) tail, cytoplasmic Pab is recruited to the poly(A) tail which, in turn, binds eIF4G, in doing so, displacing Maskin from eIF4E and facilitating initiation of translation (Figure 1.2B).

Recent evidence reveals that similar polyadenylation-mediated regulatory mechanisms controlling early development are also present in the central nervous system of both invertebrate and vertebrate animals (34-37). In the brain, neurons can differentiate between synaptic inputs that have been stimulated just once and those that were stimulated multiple times. This phenomenon where synapses undergo biochemical and morphological changes in response to stimulation is termed synaptic plasticity. Synaptic plasticity, which is the underlying basis for learning and memory, is highly dependant on localized protein synthesis (38). One mechanism for regulating synaptic protein synthesis is mediated by polyadenylation as evidenced by studies showing CPEB-mediated polyadenylation and translation of αCamKII (the α subunit of calcium-calmodulin-dependent protein kinase II) mRNA (39-41). CPEB is now thought to control mRNA translation in neurons of several metazoans (34, 42, 43) and emerging studies in *Drosophila* indicate that both orb2 and DmGLD2, *Drosophila* orthologues of CPEB and

Gld2, respectively, are both required for long-term memory (36, 37). These studies highlight the significance of regulation of translation through polyadenylation as a means to control higher order brain function.

Engaging the translational machinery in the cytoplasm

The 5'- and 3'-ends of mRNAs are sites for dynamic exchange of mRNPs that couple mRNA processing, namely capping and polyadenylation, to protein synthesis in the cytoplasm. Cap binding complex, CBC20/80, which is initially loaded onto the 5'-end of mRNAs in the nucleus, interacts with the translation initiation factor 4G (eIF4G) which then recruits the small ribosomal subunit and eventually facilitates formation of 80S complex, competent for protein synthesis (44). After the initial round of translation, another major remodeling of mRNPs occurs, which allows the switch between CBC20/80 with the major cytoplasmic cap-binding protein eIF4E. This is also the point where the binding of Nuclear poly(A) binding protein (PABPN1) to the poly(A) tails at the 3'-end of transcripts is replaced by its cytoplasmic counterpart, PABCs (45, 46). The network of interactions between the 5' cap, eIF4E, eIF4G, PABCs and the poly(A) tails allows the formation of a loop that is efficiently translated by polyribosomes, protected from mRNA degradation and serves as a platform for translational control by other mRNPs at the 3'UTRs (see Figure 1.1).

Localized mRNPs – location matters

Numerous mRNPs are important for maintaining both the quiescent state of mRNAs and localization to specific subcellular regions for protein synthesis. As one can

imagine, depending on which proteins are bound, mRNAs can have multiple fates. This is especially the case for highly specialized cells, such as oocytes and neurons. For instance, in *Drosophila*, oskar mRNA is specifically localized to the posterior pole of the oocyte for proper germline and abdomen formation in the future embryo (47, 48). The mRNP that constitutes the translationally silenced oskar mRNA during transport includes Cup, which is a 3' UTR-bound 4E inhibitory protein (14). In addition to Cup, deposition of EJC components, *Mago nashi* and *Tsunagi*, is required for oskar cytoplamsic localization (49, 50).

In mammalian neurons, ~400 different mRNAs have been identified that localize asymmetrically in the dendrites (51). This highly regulated subcellular localization of translationally dormant mRNAs is mediated by large RNP particles that house multiple mRNAs and mRNPs (51). One particular example is the localization of β-actin mRNA into axons and growth cones of neurons (52). The molecular interaction required for this transport of RNA particle is mediated by Zipcode binding protein (ZBP1), which binds to the zipcode sequences located in the 3' UTR of β-actin mRNAs (53). Moreover, fragile X mental retardation protein (FMRP), which acts as a translational repressor (54), is also present in RNP particles that associate with microtubules for mRNA transport in dendrites (55).

Poly(A)-binding proteins (Pabs) – the traditional RRM-containing Pabs

Poly(A)-binding proteins (Pabs) are critical for binding to poly(A) tails of mRNAs to control polyadenylation, mRNA export, translation and mRNA stability. Nuclear Pabs facilitate the synthesis of poly(A) tail and regulate its length and maturation

whereas cytoplasmic Pabs facilitate the formation of a 'closed loop' structure that allows binding of additional factors in regulating translation, recycling of ribosomes and mRNA stability. In humans, there are four cytoplasmic Pabs (PABPC1, PABPC3, iPABP and PABPC5) (10). Prior to this dissertation, one nuclear Pab, PABPN1 had been identified (10). As seen in Figure 1.3, conventional Pabs contain RNA-recognition motifs (RRMs), which comprise a globular domain of four-stranded anti-parallel β sheet backed by two α helices (56). Pabs interact with poly(A) RNA via these RRMs (56).

New class of Pabs in town – tandem zinc finger polyadenosine RNA-binding proteins

While these conventional RRM-containing Pabs have been extensively studied, a new class of Pabs that recognizes poly(A) RNA through tandem zinc-finger (ZnF) motifs has recently been described (57). The founding member of this family is the *S. cerevisiae* Nuclear abundant poly(A)-Binding protein 2 (Nab2), which recognizes poly(A) RNA through a tandem Cys_3His (CCCH) ZnF motif (57, 58), unlike the conventional Pabs (Figure1.4). In budding yeast, Nab2 plays the role of nuclear Pab. Nab2 is essential for cell viability and required for proper mRNA processing and export (25, 59, 60). Specific mutations in Nab2 result in increased bulk poly(A) tail length and accumulation of bulk poly(A) RNA in the nucleus (25). The N-terminal PWI-like domain of Nab2 interacts with Mlp proteins at the nuclear pore which suggested the model that Nab2 facilitates efficient poly(A) RNA export (61, 62). Based on the findings that Nab2 shuttles between the nucleus and the cytoplasm in a transcription dependent manner (60), binds specifically to poly(A) RNA and regulates poly(A) tail length and nuclear export of

mRNA transcripts, the current model is that Nab2 functions to properly process and escort transcripts out of the nucleus.

Recently, a human counterpart of Nab2, Cys_3His Zinc Finger Protein #14 (ZC3H14), was identified (63). ZC3H14 is homologous to Nab2 at both the N-terminal PWI-like domain and C-terminal tandem CCCH ZnF domain (Figure1.4). ZC3H14 binds to poly(A) RNA *in vitro* via the ZnF domain (57). The *ZC3H14* locus encodes several alternatively spliced transcripts (63) (Figure 1.5). Longer isoforms 1-3, which shares the most similarity with Nab2, are found in the nucleus at steady-state and co-localize to SC35-positive nuclear speckles, which are sites of RNA metabolism (63). The shorter isoform, isoform 4, localizes to the cytoplasm at steady-state and is brain and testes enriched suggesting possible tissue-specific roles in modulating gene expression (63). However, no studies have examined the functional role of this novel class of ZnF Pabs in multicellular organisms or their contribution to development, despite the fact that candidate Nab2 orthologues are present in all higher eukaryotes including humans, mice, zebrafish and most importantly for this dissertation, the fruit fly *Drosophila melanogaster*.

III. RNA-binding proteins implicated in human disease

RNA-binding proteins (RBPs) are the core components of mRNP complexes that guide mRNAs through maturation and translation into proteins. Since RBPs are involved in multiple aspects of RNA metabolism, altering expression or function of these proteins can impact different genes and pathways, likely resulting in complex phenotypes. In humans, more than 500 RBPs have been identified (64). To date, there are many human diseases related to aberrant RBP function either through a loss-of-function or toxic gain-of-function mechanism (65). Interestingly, diseases that arise from mutations in RBPs that show ubiquitous expression and play important roles for overall RNA metabolism in all cell types, result in tissue-specific phenotypes. Why certain tissues are more sensitive to defects in specific RNA-binding protein functions is unclear. To describe this phenomenon further, this section focuses on three specific human genetic disorders.

Fragile X syndrome

Fragile X syndrome (FXS) is the most commonly inherited form of intellectual disability. It is an X-linked disorder affecting about 1 in 4000 males and 1 in 6000 females. In addition to moderate to severe cognitive impairment, patients display an array of autism-associated behaviors (66). Common physical features associated with FXS include elongated faces, prominent ears and macroorchidism (67). The gene responsible for FXS is fragile X mental retardation 1 gene (*FMR1*), in which most commonly an expansion in a noncoding CGG-trinucleotide repeat in the 5' UTR and in turn, loss of gene expression, gives rise to the disease. Interestingly, greater than 200 CGG repeats give rise to FXS, whereas individuals with 60-200 CGGs, often referred to

as a premutation, carry an independent neurological disorder, fragile-X-associated tremor/ataxia syndrome (FXTAS).

FMR1 encodes an mRNA-binding protein FMRP, which is thought to act as a regulator of translation. FMRP contains two hnRNP K-protein homology domains (KH) and an RGG box that mediate binding to higher order structure RNAs (68-70). FMRP binds to ~4% of the mRNA present in the mammalian brain (71), highlighting its significant function in the brain. In neurons, FMRP is present in cytoplasmic RNPs associated with polyribosomes (72) and is implicated in repressing translation and allowing transport of target mRNAs to distal locations, where localized protein synthesis takes place (55). Target mRNAs of FMRP that play critical roles in brain function have been identified, whose translation is negatively affected in FXS models, including MAP1B/Futsch, α-CamKII, chickadee/profiling, pickpocket and PSD-95 (73-77). Both mouse and fly models of FXS serve as excellent tools to understand the molecular pathogenesis of FXS. Yet, the underlying fundamental biological question as to how loss of a ubiquitously expressed protein, FMRP, leads to specific neurological phenotypes and syndromic clinical features still remains unknown.

Spinal Muscular Atrophy

Loss or mutations of the survival motor neuron 1 gene (*SMN1*) lead to reduced SMN protein levels, which cause selective dysfunction of motor neurons and atrophy of the muscle (78-80). This disease is often referred to as spinal muscular atrophy (SMA). SMA is an autosomal recessive disease and is a common genetic cause of infant death (81). In humans, two genes *SMN1* and *SMN2* encode SMN proteins. Loss or mutation of

SMN1 and retention of *SMN2* is the hallmark of the disease (80). *SMN1* and *SMN2* are almost identical except for the functional difference caused by a C-T change in exon 7, which changes the amount of exon 7 that is incorporated into the final *SMN* transcript. *SMN2* lacks this exon, which disrupts the SMN oligomerization and leads to unstable protein, which is then rapidly degraded. Since normally the *SMN2* gene produces considerably less protein than *SMN1*, it makes sense that the severity of the disease is determined by the residual amount of SMN2 (82).

Based on animal model studies, two hypotheses have been formed. First, the best-characterized role for SMN proteins is in the functional assembly of snRNPs that are crucial for recognition of splice sites and catalytic removal of introns from pre-mRNAs (83-85). Due to this role of SMN, disrupting the formation of snRNPs likely affects the splicing of many different genes that are needed for motor neuron circuitry (83-85). The second hypothesis relies on the observation that SMN functions in a unique axonal complex with hnRNPQ/R and ZBPs to affect transport of β-actin mRNA and other mRNAs that are yet to be determined (86-88). It is yet unclear which model is correct or whether the two models are linked. However, as with Fragile X syndrome, one challenge that the field needs to address is how a mutation in a ubiquitously expressed gene causes a specific neurological disorder.

Oculopharyngeal muscular dystrophy (OPMD)

Oculopharyngeal muscular dystrophy is a late-onset, autosomal dominant disorder caused by an abnormal expansion of a (GCG)n triucleotides in the coding region of the *PABPN1* gene (89). Normally, the PABPN1 N-terminus contain 10 alanines encoded by $(GCG)_6$

which is expanded to 12-17 alanines in disease-causing mutant proteins. OPMD is mainly characterized by progressive eyelid drooping (ptosis) and difficulties in swallowing (90). In addition, progressive muscle weakness in the proximal limb, facial and other muscles are apparent (91, 92). One pathological hallmark of this disease is the presence of aggregates in the nuclei of skeletal muscle fibers (93, 94).

PABPN1 is a ubiquitously expressed polyadeonsine RNA binding protein that is important for regulation of poly(A) tail length control (95). Why the mutant PABPN1 causes OPMD is unknown. Moreover, despite the ubiquitous expression of the protein, why skeletal muscles are primarily affected in patients is also unknown. However, recent study identified a role for PABPN1 in myoblast proliferation as well as differentiation in primary muscle cells of mice (29), suggesting PABPN1 plays tissue-specific functions that are critical for skeletal muscles. Similar studies dissecting the functional and tissue-specific roles of critical post-transcriptional regulatory RBPs are needed to understand how mutations in these proteins lead to tissue-specific disease.

The conundrum

One interesting observation is that disrupting the function of ubiquitously expressed proteins result in selective phenotypes in certain tissues and not others. In particular, alanine expansion within the 5'-end of the nuclear Pab, PABPN1, results in OPMD, which selectively impacts the muscles of the eyelid and pharynx in addition to skeletal muscles. In addition, mutations in the general splicing factor SMN1, affect only lower motor neurons. Moreover, the main clinical feature of FXS patients is impaired higher-cognitive function. While patients with these diseases carry the same mutations in

all tissues, specific tissues seem to be more sensitive to the misregulated state of these proteins. Characterization of tissue-specific roles and identification of targets of these disease-causing RBPs will provide clues to this conundrum and furthermore, suggest possible treatment for these disorders.

IV. *Drosophila melanogaster* as a model system

Drosophila melanogaster is an attractive model organism due to its rapid generation time and tools available to dissect genetic pathways, model human disease and serve as biological platforms for therapeutic development. With its genome fully sequenced (96), there are many methods available to carry out molecular genetic research (further discussed in Appendix). In this section, modeling human disease with *Drosophila* as a model organism with an emphasis on intellectual disability will be described. In addition, a discussion on the *Drosophila* mushroom bodies, as a place to study higher order brain function, is to follow.

Drosophila as a model for human developmental brain disorders

Modeling human genetic disorders in *Drosophila* has served as a valuable tool in understanding the molecular pathologies underlying disease. More specifically, the development of fly models of human intellectual disability (97, 98), previously referred to as mental retardation, exemplifies the power of fly genetics in dissecting our understanding of developmental brain disorders.

Intellectual disability (ID)

Intellectual disability (ID) is characterized by limited intellectual capacities reflected by an IQ below 70 and major constraints in adaptive behavior (99). Typically both of these characteristics are apparent before the age of 18 years. Therapeutic options for the treatment of ID are extremely limited and its comparatively high prevalence of about 2% renders this disorder a major socioeconomic burden (99). Approximately 50%

of severe ID cases are thought to be the result of chromosomal abnormalities, rearrangements, submicroscopic deletions or duplications, and point mutations (100). In most cases, the phenotypes caused by these mutations follow a recessive mode of inheritance. Of these, more than 91 genes present on the X-chromosome have been linked to ID while only six autosomal recessive ID (ARID) genes have been identified; this bias is likely due to the fact that X-chromosomal genes affect every male carrier and are thus more easily detected in populations (99, 101).

Fly models of ID

Analysis of OMIM (Online Mendelian Inheritance in Man; http://www.ncbi.nlm.nih.gov/omim/) database reveals over 282 human genetic diseases that present ID as a clinical feature (102). Of those human ID genes, 76% have at least one orthologue in *Drosophila*, indicating that ID-associated molecules are remarkably conserved (102). The best-characterized ID gene in *Drosophila* is the *Drosophila fragile X mental retardation 1* (*dFmr1*), a fly counterpart of the human gene responsible for Fragile X syndrome (FXS) (103). As mentioned in section III, FXS is the most commonly inherited intellectual disability.

Several null and hypomorphic alleles of *dFmr1* have been created and characterized (73, 104-106). In most cases, *dFmr1* null flies are viable with no gross abnormalities throughout development; however, some groups have reported that they are semi-lethal (107, 108). Morphological studies of *dFmr1* mutants revealed defects in axons or dendrites of specific neurons. Neurons in both the peripheral nervous system (PNS) and central nervous system (CNS) display over-elaboration and extensive

arborization of dendritic morphologies in *dFmr1* mutants (73, 104, 106, 107). These phenotypes are consistent with the finding that loss of FMRP in mice leads to abnormally elaborated dendritic-spine development (109). *dFmr1* mutant adult brains also exhibit a midline-crossing defect in the β-lobe of the mushroom bodies (MBs) (110, 111). In certain neuronal subtypes, synaptic transmission is compromised due to loss of dFMRP (73).

Behavioral and cognitive phenotypes complement the morphological defects seen in *dFmr1* mutants. These flies lack normal circadian rhythm and interest in courtship (104, 105, 107). Courtship activity is a well-described social behavior in flies that can be used to model social impairment. *dFmr1* null flies show deficits in both short-term memory in courtship conditioning assays (111) and long-term memory in olfactory-based assays (112).

Using both fly and mouse models of FXS, pharmacological rescue studies have led to disease interventions that are currently in clinical trials. Studies were initiated based on the 'mGluR theory' which suggests that many FXS symptoms result from the inability of FMRP to act as a negative regulator of translation, resulting in excessive translation of target mRNAs that are responsive to activation of Group I metabotropic glutamate receptors (mGluRs) (113, 114). Strikingly, *dFmr1* mutant flies that were fed with mGluR antagonists during development showed rescue of defects in naïve courtship, immediate-recall memory, short-term memory and MB crossover (111). These defects, except axonal crossing over in MBs, were rescued with treatment during adulthood (111), suggesting that post-developmental adult brains of other species might respond to drug treatments. Based on these observations along with other studies in mouse models,

mGluR antagonists including lithium, drugs fenobam and STX107 have been or are currently in clinical trials in hopes to treat FXS patients as well as other ID-related disorders (115).

Drosophila mushroom bodies – center for higher cognitive function

Drosophila can be used to model cognitive functions in humans. Because flies manifest a wide range of experience-dependent behaviors, they can be used to identify and characterize genetic mutants that are defective in behavior plasticity (116, 117). In analysis of these genes, molecular mechanisms underlying learning and memory have been found to be highly conserved between vertebrates and invertebrates (118). In addition, flies can be subjected to various learning paradigms, such as olfactory discrimination, operant visual conditioning, spatial learning and courtship conditioning (119), in order to study how manipulation of genes and pharmacological agents can be used to modulate learning and memory.

Drosophila mushroom bodies (MBs) are highly specialized structures in the CNS that are responsible for learning and memory (120). Due to their roles in learning and memory, MBs are often cited as functional analogs of the vertebrate hippocampus (119). Recently, studies have shown that MBs are also implicated in control of sleep (121, 122) and more controversially, locomotor behavior (123, 124). Anatomically, they are a paired structure with thick bundles formed by the long, densely packed parallel fibers of the intrinsic MB neurons, the Kenyon cells (around 2500 in *Drosophila*) (Figure 1.6). MBs receive multimodal sensory signal via the dendritic calyx in the posterior brain and send

out axonal projections to the anterior brain where they bifurcate to form the vertical (α and α') and medial (β, β' and γ) lobes (Figure 1.6) (119).

Using the different learning paradigms mentioned above, learning functions of the MBs were identified. For example, in an olfactory discrimination learning paradigm, one of two odorants is paired with a positive or negative stimulus, generally a sugar solution or electric shock. Depending on the paired stimulus, reactions to the odorants can be learned and remembered (125, 126). Genetic mutants that alter the structure MBs or flies with chemically (HU) ablated MBs fail to learn in this paradigm (127, 128). In addition, studies have shown that MBs are involved in courtship conditioning-based memory (129-131). In a courtship suppression assay, when a naïve male is exposed to a mated female, he attempts to court but finds his advances unsuccessful as the female is already mated. Over time, the male learns that courtship will not be productive (132, 133). This unsuccessful courtship lasts a period of time and if this now-conditioned male is presented with a new virgin female, whom he is supposed to court successfully, he 'remembers' the previous learning experience and the time spent courting drastically reduces to that of naïve males. Flies with defective MB structures show a memory defect (129-131). Due to these studies showing high functional similarity of the Drosophila MBs with the vertebrate hippocampus, we utilized the MBs as a system to dissect the molecular and cellular functions of *dNab2/ZC3H14* in Chapter 3.

V. Scope and significance of the dissertation

The aim of my dissertation focuses on answering a fundamental biological question: How does misregulation of RNA-binding proteins critical for RNA metabolism lead to tissue-specific phenotypes manifested by human diseases? To answer this question, I used *Drosophila melanogaster* as a model system to understand the functional and developmental roles of a previously uncharacterized fly orthologue of the Cys_3His (CCCH) tandem ZnF polyadenosine RNA-binding protein Nab2/ZC3H14. Studying both the developmental and functional roles of Nab2/ZC3H14 class of proteins in a multicellular context provides a window into understanding the biological relevance of mRNA processing in tissue-specific functions and furthermore, the molecular mechanisms underlying disease.

Prior to my work, the role of Nab2/ZC3H14 has been exclusively studied in the unicellular model, budding yeast, *S. cerevisiae*. Previous studies found that Nab2 is essential for cell viability and that specific mutations lead to hyperaccumulated poly(A) RNA in the nucleus and increased bulk poly(A) tail length (25, 59) (60). Based on these findings along with other studies (26, 61), it has been postulated that Nab2 binds to poly(A) RNAs and functions to properly process and escort transcripts out of the nucleus. More recently, the identification of human Nab2, ZC3H14, has led to the idea that Nab2 class of proteins are highly conserved across species, including its steady-state subcellular localization and poly(A) RNA-binding properties (57, 63). In addition, ZC3H14 has been shown to be alternatively spliced, generating multiple isoforms that are enriched in tissue-specific manner (63), suggesting possible tissue-specific roles of ZC3H14 in modulating gene expression. However, no studies have been done to address

this point. Moreover, no studies have been done with Nab2/ZC3H14 outside *S. cerevisiae* and immortalized cell lines.

Given the previous findings, I hypothesized that accurate processing of mRNAs is fundamental to gene expression important for development. By creating loss-of-function mutations in the fly *Nab2/ZC3H14* gene, *dNab2*, we characterized both the organismal and molecular phenotypes of *dNab2* mutants in Chapter 2. *dNab2* is essential as loss of *dNab2* results in organismal lethality in majority of flies; however, a portion of flies survive to adulthood with myriad of defective phenotypes including held-out wings, thoracic bristle malformation, flightlessness, poor locomotor activity and short lifespan. Using transgenic rescue constructs, we further showed that *dNab2* is both necessary and sufficient in all neurons to promote normal behavior, including flight and locomotion. Biochemical and genetic evidence suggest that *dNab2* is a polyadenosine RNA-binding protein that plays key regulatory role in maintaining poly(A) tail length and more specifically, acts to antagonize the nuclear Pab, *Pabp2* in *Drosophila*. In the course of studying these mutants, we identified a novel link to human neurological disorder, whereby mutations in *ZC3H14* gene cause non-syndromic autosomal recessive intellectual disability (NS-ARID). This disease link gave us insight into identifying the neuronal role of dNab2 and triggered further studies in understanding the role of *dNab2* in the fly mushroom body neurons, as discussed in Chapter 3.

Based on the human data that mutations in *ZC3H14* cause NS-ARID, I focused further efforts on studying the role of *dNab2* in a specific subset of neurons, the fly mushroom bodies (MBs), which are the functional analog to the vertebrate hippocampus and are important for learning and memory. In Chapter 3, data is presented suggesting

that *dNab2* is necessary for promoting normal locomotor activity in the MB neurons. Furthermore, we show that depletion of dNab2 in these cell types result in flies that are defective in short-term memory, indicating that dNab2 is important for memory formation. The precise molecular mechanisms underlying this process remain to be discovered.

In order to understand how *dNab2* controls poly(A) tail length and promote normal behavior, we undertook both candidate-based and discovery-based genetic screens in an attempt to identify *dNab2* interactors. As shown in data presented in Chapter 4, we found genetic interactions between multiple genes that are known to be important for brain function with *dNab2*. In characterizing these interactions, we found an interesting link between *dNab2* and *dFmr1*, the *Drosophila* Fragile X mental retardation gene, whereby both *dNab2* and *dFmr1* might act in the same or parallel pathways to control normal behavior in neurons.

In summary, this dissertation identifies and characterizes the role of a novel *Drosophila* CCCH tandem ZnF polyadenosine RNA-binding protein, dNab2 during multicellular organism development. dNab2 binds to polyadenosine RNA and controls poly(A) tail length. dNab2 function is critical for multiple tissue-specific developmental processes as loss of dNab2 results in multiple organismal phenotypes. With the recent identification linking ZC3H14 to human disease affecting neurological function, dNab2's role was shown to be both necessary and sufficient in all neurons to promote normal behavior. More specifically, characterization of dNab2's function in the *Drosophila* mushroom bodies suggest the important role for dNab2 in controlling both locomotion and memory formation. Furthermore, novel genetic interactions between dNab2 and

candidate genes provide a platform for further studies understanding the molecular mechanisms underlying human disease. Taken together, these findings reveal a critically important role for a class of CCCH tandem ZnF-mediated polyadenosine RNA-binding proteins in organism development and provide further insight into post-transcriptional regulation of gene expression in human health.

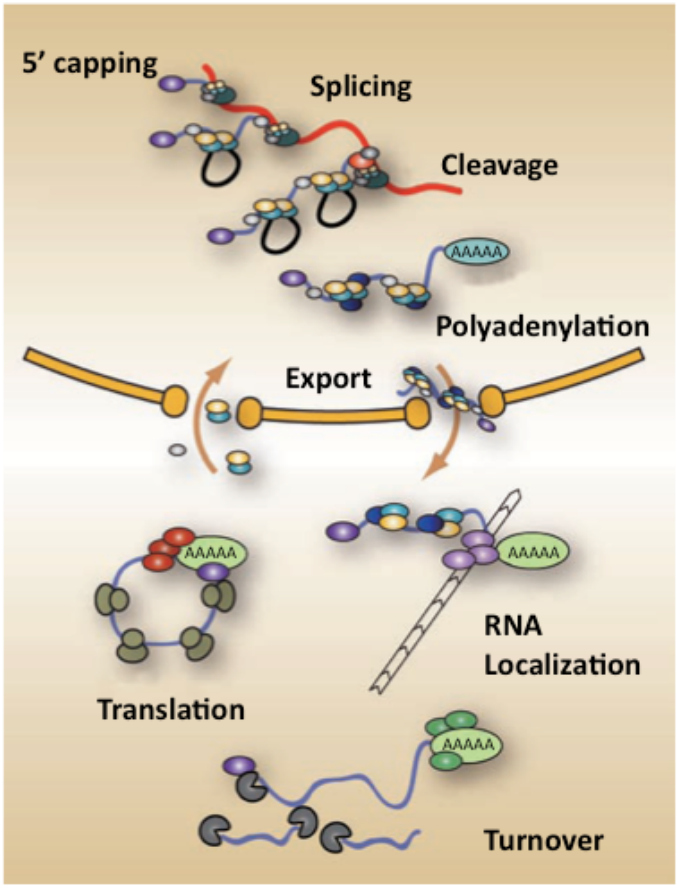


Figure 1.1. Post-transcriptional regulatory mechanisms involving proper processing and expression of mRNAs. In the nucleus, transcripts are capped, spliced, cleaved and polyadenylated co- and post-transcriptionally. Properly processed mRNAs are then allowed for export into the cytoplasm where they are subject to multiple fates, including subcellular localization, translation and degradation. All of these processes are mediated by various RNA-binding proteins that form messenger ribonucleoprotein complexes (mRNPs) with transcripts, which help guide and determine the fate of mRNAs. (Adapted from McKee and Silver, 2007)

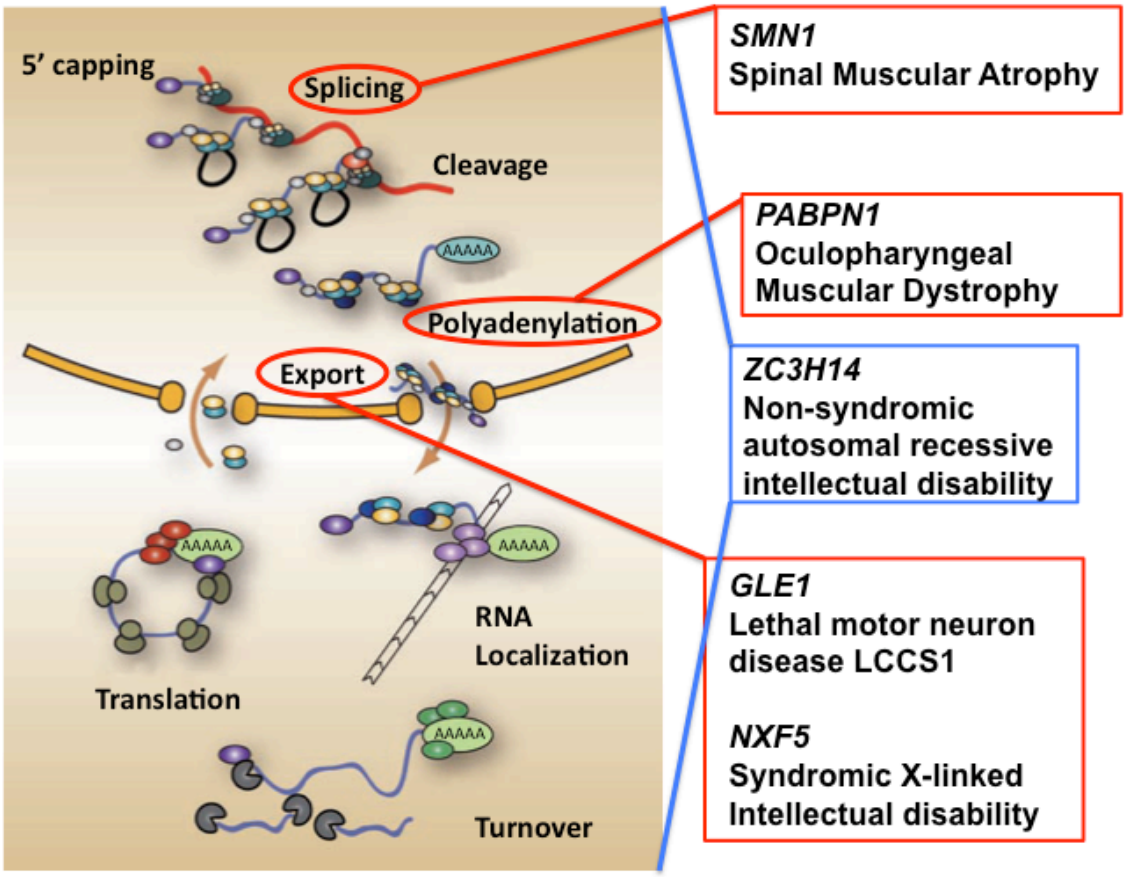


Figure 1.1'. Human diseases associated with mutations in mRNA processing factors. Highlighted in colored boxes indicate the genes mutated in human diseases and names of diseases associated with those mutations. Circles represent the mRNA processing steps at which these genes are known to function. ZC3H14-mediated non-syndromic autosomal recessive intellectual disability (NS-ARID) was identified during the course of this dissertation. (Adapted from McKee and Silver, 2007)

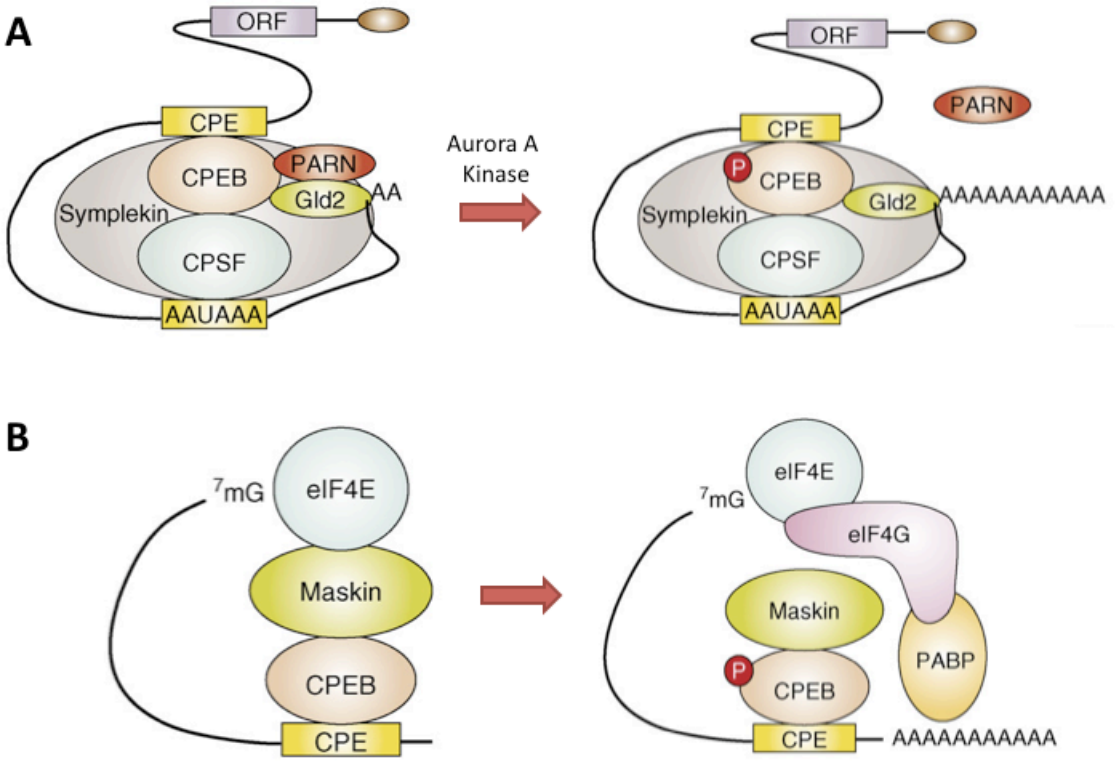


Figure 1.2. CPEB-mediated regulation of polyadenylation and translation. (**A**) CPEB-regulated polyadenylation. In the cytoplasm, polyadenylated CPE-containing RNAs are bound to CPEB in addition to the multi-subunit complex containing CPSF, which recognizes the AAUAAA polyadenylation hexanucleotide; PARN, a deadenylating enzyme; Gld2, a poly(A) polymerase; and symplekin, a scaffolding protein. Because PARN activity is more robust than Gld2, poly(A) tails are trimmed and remain short. Following a progesterone-based signaling cascade, Aurora A kinase is activated and phosphorylates CPEB at Ser174, which causes expulsion of PARN from the complex and allows Gld2 to elongate the poly(A) tails. (**B**) Translational control by CPEB-Maskin. CPEB binds to both CPE-containing RNAs and Maskin. Maskin interacts with cap-binding protein eIF4E and inhibits translation via inhibiting the binding of eIF4G to eIF4E. Following CPEB phosphorylation and polyadenylation (see Figure **A**),

cytoplasmic Poly(A)-binding protein (PABP) binds to the poly(A) tails and eIF4G. Upon doing so, this interaction mediates displacement of Maskin from eIF4E and allows initiation of translation to occur. (Adapted from Richter, 2007)

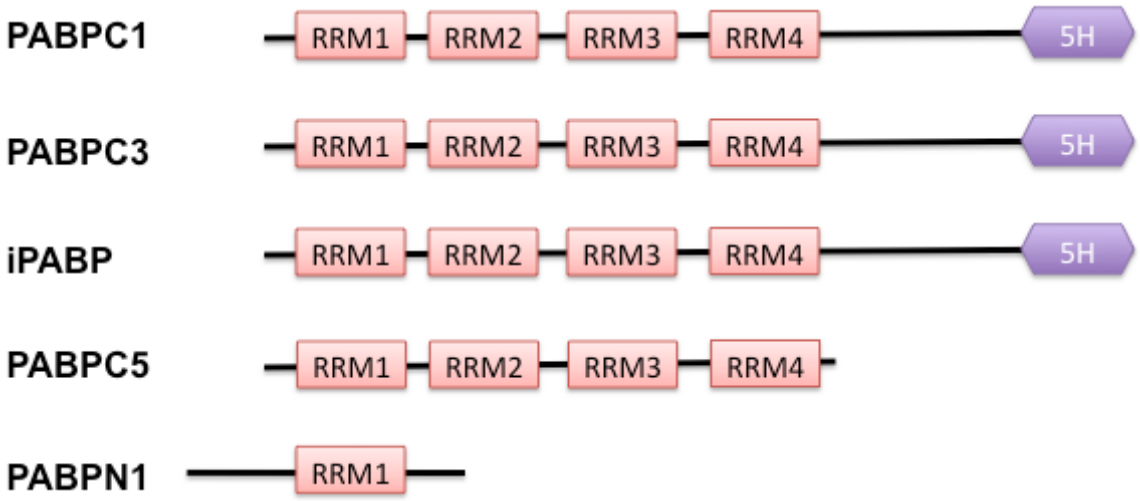


Figure 1.3. Domains of human poly(A)-binding proteins (Pabs). Cytoplasmic Pabs, PABPC1, PABPC3, iPABP and PABPC5 all share four RNA-recognition motifs (RRMs) whereas the nuclear Pab, PABPN1, contains a single RRM. Purple hexagons containing 5H represent the five conserved helices at the C-terminus. (Adapted from Mangus *et al*., 2003)

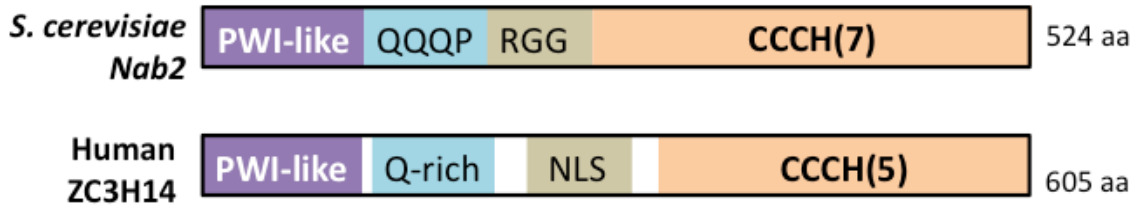


Figure 1.4. Domain structures of the budding yeast *S. cerevisiae* Nab2 and human ZC3H14. The conserved N-terminal PWI-like fold, Q-rich, RGG/predicted nuclear localization signal (NLS) and C-terminal tandem Cys_3His zinc finger (ZnF) RNA-binding motif (CCCH) domains are indicated. Budding yeast Nab2 contains seven tandem ZnFs whereas human ZC3H14 contains five. Crystal structure of the budding yeast PWI-like fold has been solved (62) and this domain mediates interaction of Nab2 to the nuclear pore and facilitates nuclear export (61). C-terminal CCCH domain of Nab2 has been extensively characterized for its recognition of polyadenosine RNA (58).

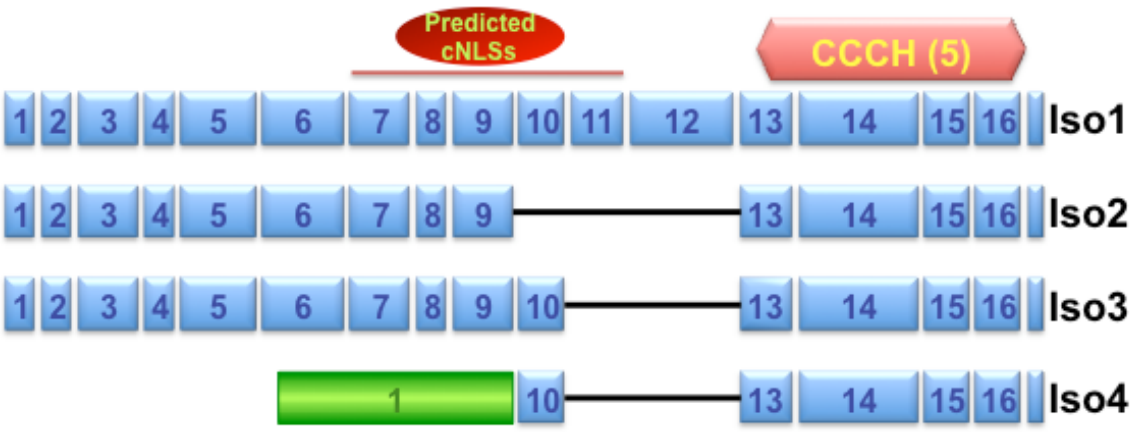


Figure 1.5. Alternative splicing of *ZC3H14* generates multiple isoforms. Diagram showing exons predicted for the *ZC3H14* splice variants encoding ZC3H14 isoform 1 (Iso1), isoform 2 (Iso2), isoform 3 (Iso3), and isoform 4 (Iso4). Approximate locations of predicted classical NLS motifs (cNLSs) as well as the tandem Cys_3His zinc finger domain (CCCH) are indicated. (Adapted from Leung *et al*., 2009)

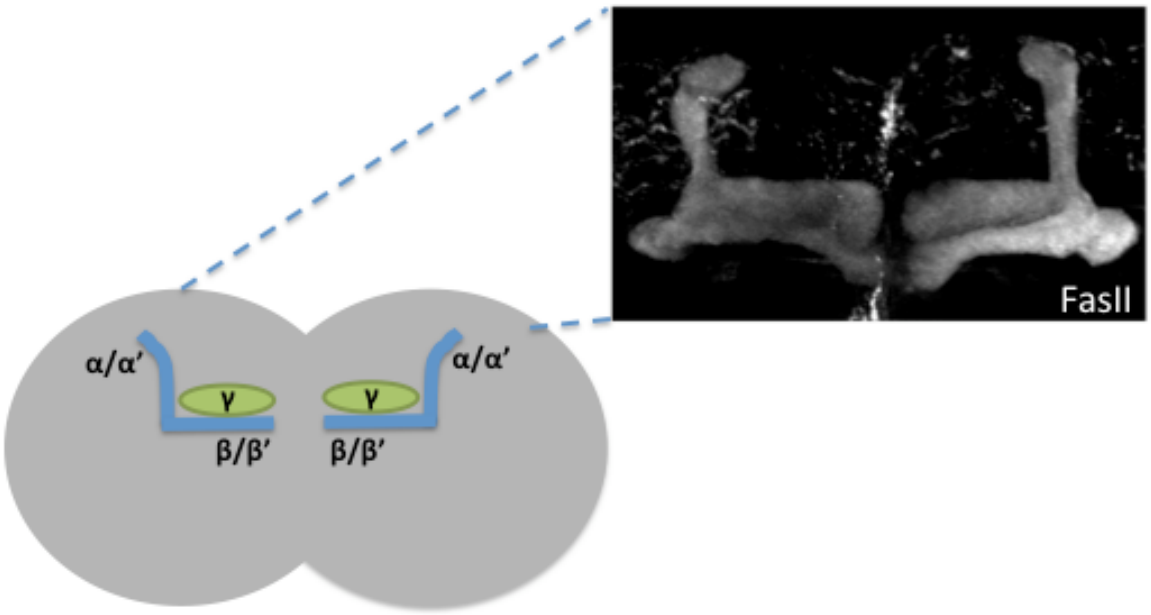


Figure 1.6. Structure of the *Drosophila* adult mushroom bodies (MBs). Each MB consists of approximately 2500 Kenyon cell fibers that primarily receive sensory information via the dendritic calyx in the posterior part of the brain (not shown) (120). Illustrated in cartoon images are axonal projections to the anterior part of the brain where they bifurcate to form the vertical (α/α') and medial (β/β' and γ) lobes. These lobes can be visualized by staining adult brains with anti-Fasciculin II (FasII) antibody. Shown on the right is a confocal image of control adult brain stained with FasII, marking all the lobes of the MB.

References

1. Komili S & Silver PA (2008) Coupling and coordination in gene expression processes: a systems biology view. *Nat Rev Genet* 9(1):38-48.
2. Kelly SM & Corbett AH (2009) Messenger RNA export from the nucleus: a series of molecular wardrobe changes. *Traffic* 10(9):1199-1208.
3. McKee AE & Silver PA (2007) Systems perspectives on mRNA processing. *Cell Res* 17(7):581-590.
4. Keene JD & Tenenbaum SA (2002) Eukaryotic mRNPs may represent posttranscriptional operons. *Mol Cell* 9(6):1161-1167.
5. Hieronymus H & Silver PA (2004) A systems view of mRNP biology. *Genes Dev* 18(23):2845-2860.
6. Maris C, Dominguez C, & Allain FH (2005) The RNA recognition motif, a plastic RNA-binding platform to regulate post-transcriptional gene expression. *FEBS J* 272(9):2118-2131.
7. Messias AC & Sattler M (2004) Structural basis of single-stranded RNA recognition. *Acc Chem Res* 37(5):279-287.
8. Stefl R, Skrisovska L, & Allain FH (2005) RNA sequence- and shape-dependent recognition by proteins in the ribonucleoprotein particle. *EMBO Rep* 6(1):33-38.
9. Fechter P & Brownlee GG (2005) Recognition of mRNA cap structures by viral and cellular proteins. *J Gen Virol* 86(Pt 5):1239-1249.
10. Mangus DA, Evans MC, & Jacobson A (2003) Poly(A)-binding proteins: multifunctional scaffolds for the post-transcriptional control of gene expression. *Genome Biol* 4(7):223.

11. Tange TO, Nott A, & Moore MJ (2004) The ever-increasing complexities of the exon junction complex. *Curr Opin Cell Biol* 16(3):279-284.

12. Mandel CR, Bai Y, & Tong L (2008) Protein factors in pre-mRNA 3'-end processing. *Cell Mol Life Sci* 65(7-8):1099-1122.

13. Richter JD (1999) Cytoplasmic polyadenylation in development and beyond. *Microbiol Mol Biol Rev* 63(2):446-456.

14. Richter JD & Sonenberg N (2005) Regulation of cap-dependent translation by eIF4E inhibitory proteins. *Nature* 433(7025):477-480.

15. Richter JD (2007) CPEB: a life in translation. *Trends Biochem Sci* 32(6):279-285.

16. Phatnani HP & Greenleaf AL (2006) Phosphorylation and functions of the RNA polymerase II CTD. *Genes Dev* 20(21):2922-2936.

17. Fong YW & Zhou Q (2001) Stimulatory effect of splicing factors on transcriptional elongation. *Nature* 414(6866):929-933.

18. Shuman S (1997) Origins of mRNA identity: capping enzymes bind to the phosphorylated C-terminal domain of RNA polymerase II. *Proc Natl Acad Sci U S A* 94(24):12758-12760.

19. Wahle E (1991) A novel poly(A)-binding protein acts as a specificity factor in the second phase of messenger RNA polyadenylation. *Cell* 66(4):759-768.

20. Bienroth S, Keller W, & Wahle E (1993) Assembly of a processive messenger RNA polyadenylation complex. *EMBO J* 12(2):585-594.

21. Kuhn U, *et al*. (2009) Poly(A) tail length is controlled by the nuclear poly(A)-binding protein regulating the interaction between poly(A) polymerase and the

cleavage and polyadenylation specificity factor. *J Biol Chem* 284(34):22803-22814.

22. Hammell CM, *et al*. (2002) Coupling of termination, 3' processing, and mRNA export. *Mol Cell Biol* 22(18):6441-6457.

23. Brodsky AS & Silver PA (2000) Pre-mRNA processing factors are required for nuclear export. *RNA* 6(12):1737-1749.

24. Libri D, *et al*. (2002) Interactions between mRNA export commitment, 3'-end quality control, and nuclear degradation. *Mol Cell Biol* 22(23):8254-8266.

25. Hector RE, *et al*. (2002) Dual requirement for yeast hnRNP Nab2p in mRNA poly(A) tail length control and nuclear export. *EMBO J* 21(7):1800-1810.

26. Marfatia KA, Crafton EB, Green DM, & Corbett AH (2003) Domain analysis of the Saccharomyces cerevisiae heterogeneous nuclear ribonucleoprotein, Nab2p. Dissecting the requirements for Nab2p-facilitated poly(A) RNA export. *J Biol Chem* 278(9):6731-6740.

27. Viphakone N, Voisinet-Hakil F, & Minvielle-Sebastia L (2008) Molecular dissection of mRNA poly(A) tail length control in yeast. *Nucleic Acids Res* 36(7):2418-2433.

28. Juge F, Zaessinger S, Temme C, Wahle E, & Simonelig M (2002) Control of poly(A) polymerase level is essential to cytoplasmic polyadenylation and early development in Drosophila. *EMBO J* 21(23):6603-6613.

29. Apponi LH, *et al*. (2010) Loss of nuclear poly(A)-binding protein 1 causes defects in myogenesis and mRNA biogenesis. *Hum Mol Genet* 19(6):1058-1065.

30. Barnard DC, Ryan K, Manley JL, & Richter JD (2004) Symplekin and xGLD-2 are required for CPEB-mediated cytoplasmic polyadenylation. *Cell* 119(5):641-651.

31. Rouhana L, *et al*. (2005) Vertebrate GLD2 poly(A) polymerases in the germline and the brain. *RNA* 11(7):1117-1130.

32. Rouhana L & Wickens M (2007) Autoregulation of GLD-2 cytoplasmic poly(A) polymerase. *RNA* 13(2):188-199.

33. Sarkissian M, Mendez R, & Richter JD (2004) Progesterone and insulin stimulation of CPEB-dependent polyadenylation is regulated by Aurora A and glycogen synthase kinase-3. *Genes Dev* 18(1):48-61.

34. Si K, *et al*. (2003) A neuronal isoform of CPEB regulates local protein synthesis and stabilizes synapse-specific long-term facilitation in aplysia. *Cell* 115(7):893-904.

35. Alarcon JM, *et al*. (2004) Selective modulation of some forms of schaffer collateral-CA1 synaptic plasticity in mice with a disruption of the CPEB-1 gene. *Learn Mem* 11(3):318-327.

36. Kwak JE, *et al*. (2008) GLD2 poly(A) polymerase is required for long-term memory. *Proc Natl Acad Sci U S A* 105(38):14644-14649.

37. Keleman K, Kruttner S, Alenius M, & Dickson BJ (2007) Function of the Drosophila CPEB protein Orb2 in long-term courtship memory. *Nat Neurosci* 10(12):1587-1593.

38. Kang H & Schuman EM (1996) A requirement for local protein synthesis in neurotrophin-induced hippocampal synaptic plasticity. *Science* 273(5280):1402-1406.

39. Wu L, *et al*. (1998) CPEB-mediated cytoplasmic polyadenylation and the regulation of experience-dependent translation of alpha-CaMKII mRNA at synapses. *Neuron* 21(5):1129-1139.

40. Huang YS, Jung MY, Sarkissian M, & Richter JD (2002) N-methyl-D-aspartate receptor signaling results in Aurora kinase-catalyzed CPEB phosphorylation and alpha CaMKII mRNA polyadenylation at synapses. *EMBO J* 21(9):2139-2148.

41. Shin CY, Kundel M, & Wells DG (2004) Rapid, activity-induced increase in tissue plasminogen activator is mediated by metabotropic glutamate receptor-dependent mRNA translation. *J Neurosci* 24(42):9425-9433.

42. Liu J & Schwartz JH (2003) The cytoplasmic polyadenylation element binding protein and polyadenylation of messenger RNA in Aplysia neurons. *Brain Res* 959(1):68-76.

43. Liu LO, Li G, McCall MA, & Cooper NG (2000) Photoreceptor regulated expression of Ca(2+)/calmodulin-dependent protein kinase II in the mouse retina. *Brain Res Mol Brain Res* 82(1-2):150-166.

44. Lejeune F, Ranganathan AC, & Maquat LE (2004) eIF4G is required for the pioneer round of translation in mammalian cells. *Nat Struct Mol Biol* 11(10):992-1000.

45. Amrani N, Ghosh S, Mangus DA, & Jacobson A (2008) Translation factors promote the formation of two states of the closed-loop mRNP. *Nature* 453(7199):1276-1280.

46. Wells SE, Hillner PE, Vale RD, & Sachs AB (1998) Circularization of mRNA by eukaryotic translation initiation factors. *Mol Cell* 2(1):135-140.

47. Kim-Ha J, Smith JL, & Macdonald PM (1991) oskar mRNA is localized to the posterior pole of the Drosophila oocyte. *Cell* 66(1):23-35.

48. Jenny A, *et al*. (2006) A translation-independent role of oskar RNA in early Drosophila oogenesis. *Development* 133(15):2827-2833.

49. Mohr SE, Dillon ST, & Boswell RE (2001) The RNA-binding protein Tsunagi interacts with Mago Nashi to establish polarity and localize oskar mRNA during Drosophila oogenesis. *Genes Dev* 15(21):2886-2899.

50. Hachet O & Ephrussi A (2001) Drosophila Y14 shuttles to the posterior of the oocyte and is required for oskar mRNA transport. *Curr Biol* 11(21):1666-1674.

51. St Johnston D (2005) Moving messages: the intracellular localization of mRNAs. *Nat Rev Mol Cell Biol* 6(5):363-375.

52. Zhang HL, Singer RH, & Bassell GJ (1999) Neurotrophin regulation of beta-actin mRNA and protein localization within growth cones. *J Cell Biol* 147(1):59-70.

53. Zhang HL, *et al*. (2001) Neurotrophin-induced transport of a beta-actin mRNP complex increases beta-actin levels and stimulates growth cone motility. *Neuron* 31(2):261-275.

54. Bassell GJ & Warren ST (2008) Fragile X syndrome: loss of local mRNA regulation alters synaptic development and function. *Neuron* 60(2):201-214.

55. Wang H, *et al*. (2008) Dynamic association of the fragile X mental retardation protein as a messenger ribonucleoprotein between microtubules and polyribosomes. *Mol Biol Cell* 19(1):105-114.

56. Deo RC, Bonanno JB, Sonenberg N, & Burley SK (1999) Recognition of polyadenylate RNA by the poly(A)-binding protein. *Cell* 98(6):835-845.

57. Kelly SM, *et al*. (2007) Recognition of polyadenosine RNA by zinc finger proteins. *Proc Natl Acad Sci U S A* 104(30):12306-12311.

58. Kelly SM, *et al*. (2010) Recognition of Polyadenosine RNA by the Zinc Finger Domain of Nuclear Poly(A) RNA-binding Protein 2 (Nab2) Is Required for Correct mRNA 3'-End Formation. *J Biol Chem* 285(34):26022-26032.

59. Anderson JT, Wilson SM, Datar KV, & Swanson MS (1993) NAB2: a yeast nuclear polyadenylated RNA-binding protein essential for cell viability. *Mol Cell Biol* 13(5):2730-2741.

60. Green DM, *et al*. (2002) Nab2p is required for poly(A) RNA export in Saccharomyces cerevisiae and is regulated by arginine methylation via Hmt1p. *J Biol Chem* 277(10):7752-7760.

61. Fasken MB, Stewart M, & Corbett AH (2008) Functional significance of the interaction between the mRNA-binding protein, Nab2, and the nuclear pore-associated protein, Mlp1, in mRNA export. *J Biol Chem* 283(40):27130-27143.

62. Grant RP, *et al*. (2008) Structure of the N-terminal Mlp1-binding domain of the Saccharomyces cerevisiae mRNA-binding protein, Nab2. *J Mol Biol* 376(4):1048-1059.

63. Leung SW, *et al*. (2009) Splice variants of the human ZC3H14 gene generate multiple isoforms of a zinc finger polyadenosine RNA binding protein. *Gene* 439(1-2):71-78.

64. Anantharaman V, Koonin EV, & Aravind L (2002) Comparative genomics and evolution of proteins involved in RNA metabolism. *Nucleic Acids Res* 30(7):1427-1464.

65. Lukong KE, Chang KW, Khandjian EW, & Richard S (2008) RNA-binding proteins in human genetic disease. *Trends Genet* 24(8):416-425.

66. Hagerman PJ (2008) The fragile X prevalence paradox. *J Med Genet* 45(8):498-499.

67. Hagerman RJ, *et al*. (2008) Treatment of fragile X-associated tremor ataxia syndrome (FXTAS) and related neurological problems. *Clin Interv Aging* 3(2):251-262.

68. Siomi H, Siomi MC, Nussbaum RL, & Dreyfuss G (1993) The protein product of the fragile X gene, FMR1, has characteristics of an RNA-binding protein. *Cell* 74(2):291-298.

69. Schaeffer C, *et al*. (2001) The fragile X mental retardation protein binds specifically to its mRNA via a purine quartet motif. *EMBO J* 20(17):4803-4813.

70. Darnell JC, *et al*. (2001) Fragile X mental retardation protein targets G quartet mRNAs important for neuronal function. *Cell* 107(4):489-499.

71. Ashley CT, Jr., Wilkinson KD, Reines D, & Warren ST (1993) FMR1 protein: conserved RNP family domains and selective RNA binding. *Science* 262(5133):563-566.

72. Feng Y, *et al*. (1997) FMRP associates with polyribosomes as an mRNP, and the I304N mutation of severe fragile X syndrome abolishes this association. *Mol Cell* 1(1):109-118.

73. Zhang YQ, *et al*. (2001) Drosophila fragile X-related gene regulates the MAP1B homolog Futsch to control synaptic structure and function. *Cell* 107(5):591-603.

74. Zalfa F, *et al*. (2003) The fragile X syndrome protein FMRP associates with BC1 RNA and regulates the translation of specific mRNAs at synapses. *Cell* 112(3):317-327.

75. Xu K, *et al*. (2004) The fragile X-related gene affects the crawling behavior of Drosophila larvae by regulating the mRNA level of the DEG/ENaC protein pickpocket1. *Curr Biol* 14(12):1025-1034.

76. Reeve SP, *et al*. (2005) The Drosophila fragile X mental retardation protein controls actin dynamics by directly regulating profilin in the brain. *Curr Biol* 15(12):1156-1163.

77. Hou L, *et al*. (2006) Dynamic translational and proteasomal regulation of fragile X mental retardation protein controls mGluR-dependent long-term depression. *Neuron* 51(4):441-454.

78. Crawford TO & Pardo CA (1996) The neurobiology of childhood spinal muscular atrophy. *Neurobiol Dis* 3(2):97-110.

79. McAndrew PE, *et al*. (1997) Identification of proximal spinal muscular atrophy carriers and patients by analysis of SMNT and SMNC gene copy number. *Am J Hum Genet* 60(6):1411-1422.

80. Lefebvre S, *et al*. (1995) Identification and characterization of a spinal muscular atrophy-determining gene. *Cell* 80(1):155-165.

81. Roberts DF, Chavez J, & Court SD (1970) The genetic component in child mortality. *Arch Dis Child* 45(239):33-38.

82. Lefebvre S, *et al*. (1997) Correlation between severity and SMN protein level in spinal muscular atrophy. *Nat Genet* 16(3):265-269.

83. Eggert C, Chari A, Laggerbauer B, & Fischer U (2006) Spinal muscular atrophy: the RNP connection. *Trends Mol Med* 12(3):113-121.

84. Pellizzoni L (2007) Chaperoning ribonucleoprotein biogenesis in health and disease. *EMBO Rep* 8(4):340-345.

85. Liu Q, Fischer U, Wang F, & Dreyfuss G (1997) The spinal muscular atrophy disease gene product, SMN, and its associated protein SIP1 are in a complex with spliceosomal snRNP proteins. *Cell* 90(6):1013-1021.

86. Rossoll W, *et al*. (2002) Specific interaction of Smn, the spinal muscular atrophy determining gene product, with hnRNP-R and gry-rbp/hnRNP-Q: a role for Smn in RNA processing in motor axons? *Hum Mol Genet* 11(1):93-105.

87. Rossoll W, *et al*. (2003) Smn, the spinal muscular atrophy-determining gene product, modulates axon growth and localization of beta-actin mRNA in growth cones of motoneurons. *J Cell Biol* 163(4):801-812.

88. Gu W, Pan F, Zhang H, Bassell GJ, & Singer RH (2002) A predominantly nuclear protein affecting cytoplasmic localization of beta-actin mRNA in fibroblasts and neurons. *J Cell Biol* 156(1):41-51.

89. Brais B, *et al*. (1998) Short GCG expansions in the PABP2 gene cause oculopharyngeal muscular dystrophy. *Nat Genet* 18(2):164-167.

90. Victor M, Hayes R, & Adams RD (1962) Oculopharyngeal muscular dystrophy. A familial disease of late life characterized by dysphagia and progressive ptosis of the evelids. *N Engl J Med* 267:1267-1272.

91. Ruegg S, *et al*. (2005) Oculopharyngeal muscular dystrophy - an under-diagnosed disorder? *Swiss Med Wkly* 135(39-40):574-586.

92. Van Der Sluijs BM, Hoefsloot LH, Padberg GW, Van Der Maarel SM, & Van Engelen BG (2003) Oculopharyngeal muscular dystrophy with limb girdle weakness as major complaint. *J Neurol* 250(11):1307-1312.

93. Tome FM & Fardeau M (1980) Nuclear inclusions in oculopharyngeal dystrophy. *Acta Neuropathol* 49(1):85-87.

94. Tome FM, Chateau D, Helbling-Leclerc A, & Fardeau M (1997) Morphological changes in muscle fibers in oculopharyngeal muscular dystrophy. *Neuromuscul Disord* 7 Suppl 1:S63-69.

95. Kuhn U & Wahle E (2004) Structure and function of poly(A) binding proteins. *Biochim Biophys Acta* 1678(2-3):67-84.

96. Adams MD, *et al*. (2000) The genome sequence of Drosophila melanogaster. *Science* 287(5461):2185-2195.

97. Restifo LL (2005) Mental retardation genes in drosophila: New approaches to understanding and treating developmental brain disorders. *Ment Retard Dev Disabil Res Rev* 11(4):286-294.

98. Bolduc FV & Tully T (2009) Fruit flies and intellectual disability. *Fly (Austin)* 3(1):91-104.

99. Ropers HH (2010) Genetics of early onset cognitive impairment. *Annu Rev Genomics Hum Genet* 11:161-187.

100. Chelly J, Khelfaoui M, Francis F, Cherif B, & Bienvenu T (2006) Genetics and pathophysiology of mental retardation. *Eur J Hum Genet* 14(6):701-713.

101. Gecz J, Shoubridge C, & Corbett M (2009) The genetic landscape of intellectual disability arising from chromosome X. *Trends Genet* 25(7):308-316.

102. Inlow JK & Restifo LL (2004) Molecular and comparative genetics of mental retardation. *Genetics* 166(2):835-881.

103. Wan L, Dockendorff TC, Jongens TA, & Dreyfuss G (2000) Characterization of dFMR1, a Drosophila melanogaster homolog of the fragile X mental retardation protein. *Mol Cell Biol* 20(22):8536-8547.

104. Dockendorff TC, *et al*. (2002) Drosophila lacking dfmr1 activity show defects in circadian output and fail to maintain courtship interest. *Neuron* 34(6):973-984.

105. Inoue S, *et al*. (2002) A role for the Drosophila fragile X-related gene in circadian output. *Curr Biol* 12(15):1331-1335.

106. Lee A, *et al*. (2003) Control of dendritic development by the Drosophila fragile X-related gene involves the small GTPase Rac1. *Development* 130(22):5543-5552.

107. Morales J, *et al*. (2002) Drosophila fragile X protein, DFXR, regulates neuronal morphology and function in the brain. *Neuron* 34(6):961-972.

108. Chang S, *et al*. (2008) Identification of small molecules rescuing fragile X syndrome phenotypes in Drosophila. *Nat Chem Biol* 4(4):256-263.

109. Comery TA, *et al*. (1997) Abnormal dendritic spines in fragile X knockout mice: maturation and pruning deficits. *Proc Natl Acad Sci U S A* 94(10):5401-5404.

110. Michel CI, Kraft R, & Restifo LL (2004) Defective neuronal development in the mushroom bodies of Drosophila fragile X mental retardation 1 mutants. *J Neurosci* 24(25):5798-5809.

111. McBride SM, *et al*. (2005) Pharmacological rescue of synaptic plasticity, courtship behavior, and mushroom body defects in a Drosophila model of fragile X syndrome. *Neuron* 45(5):753-764.

112. Bolduc FV, Bell K, Cox H, Broadie KS, & Tully T (2008) Excess protein synthesis in Drosophila fragile X mutants impairs long-term memory. *Nat Neurosci* 11(10):1143-1145.

113. Huber KM, Gallagher SM, Warren ST, & Bear MF (2002) Altered synaptic plasticity in a mouse model of fragile X mental retardation. *Proc Natl Acad Sci U S A* 99(11):7746-7750.

114. Bear MF, Huber KM, & Warren ST (2004) The mGluR theory of fragile X mental retardation. *Trends Neurosci* 27(7):370-377.

115. Bhogal B & Jongens TA (2010) Fragile X syndrome and model organisms: identifying potential routes of therapeutic intervention. *Dis Model Mech* 3(11-12):693-700.

116. Dudai Y, Jan YN, Byers D, Quinn WG, & Benzer S (1976) dunce, a mutant of Drosophila deficient in learning. *Proc Natl Acad Sci U S A* 73(5):1684-1688.

117. Greenspan RJ (1995) Flies, genes, learning, and memory. *Neuron* 15(4):747-750.

118. Mayford M & Kandel ER (1999) Genetic approaches to memory storage. *Trends Genet* 15(11):463-470.

119. Zars T (2000) Behavioral functions of the insect mushroom bodies. *Curr Opin Neurobiol* 10(6):790-795.

120. Heisenberg M (2003) Mushroom body memoir: from maps to models. *Nat Rev Neurosci* 4(4):266-275.

121. Guo F, Yi W, Zhou M, & Guo A (2011) Go signaling in mushroom bodies regulates sleep in Drosophila. *Sleep* 34(3):273-281.

122. Pitman JL, McGill JJ, Keegan KP, & Allada R (2006) A dynamic role for the mushroom bodies in promoting sleep in Drosophila. *Nature* 441(7094):753-756.

123. Serway CN, Kaufman RR, Strauss R, & de Belle JS (2009) Mushroom bodies enhance initial motor activity in Drosophila. *J Neurogenet* 23(1-2):173-184.

124. Martin JR, Ernst R, & Heisenberg M (1998) Mushroom bodies suppress locomotor activity in Drosophila melanogaster. *Learn Mem* 5(1-2):179-191.

125. Bitterman ME, Menzel R, Fietz A, & Schafer S (1983) Classical conditioning of proboscis extension in honeybees (Apis mellifera). *J Comp Psychol* 97(2):107-119.

126. Tully T & Quinn WG (1985) Classical conditioning and retention in normal and mutant Drosophila melanogaster. *J Comp Physiol A* 157(2):263-277.

127. Heisenberg M, Borst A, Wagner S, & Byers D (1985) Drosophila mushroom body mutants are deficient in olfactory learning. *J Neurogenet* 2(1):1-30.

128. de Belle JS & Heisenberg M (1994) Associative odor learning in Drosophila abolished by chemical ablation of mushroom bodies. *Science* 263(5147):692-695.

129. Joiner MA & Griffith LC (2000) Visual input regulates circuit configuration in courtship conditioning of Drosophila melanogaster. *Learn Mem* 7(1):32-42.

130. Joiner MA & Griffith LC (1999) Mapping of the anatomical circuit of CaM kinase-dependent courtship conditioning in Drosophila. *Learn Mem* 6(2):177-192.

131. McBride SM, *et al*. (1999) Mushroom body ablation impairs short-term memory and long-term memory of courtship conditioning in Drosophila melanogaster. *Neuron* 24(4):967-977.

132. Hall JC (1994) The mating of a fly. *Science* 264(5166):1702-1714.

133. Siegel RW, Hall JC, Gailey DA, & Kyriacou CP (1984) Genetic elements of courtship in Drosophila: mosaics and learning mutants. *Behav Genet* 14(5):383-410.

CHAPTER 2

Mutation of the Conserved Polyadenosine RNA Binding Protein, ZC3H14/dNab2, Impairs Neural Function in *Drosophila* and Humans

This chapter is adapted from the following published paper:
Mutation of the conserved polyadenosine RNA binding protein, ZC3H14/dNab2, impairs neural function in *Drosophila* and humans *PNAS 2011 ; published ahead of print July 6, 2011, doi:10.1073/pnas.1107103108*

ChangHui Pak[a,b,c,1], Masoud Garshasbi[d,e,1], Kimia Kahrizi[d], Christina Gross[a], Luciano H. Apponi[f], John J. Noto[a,b], Seth M. Kelly[a], Sara W. Leung[b], Andreas Tzschach[d], Farkhondeh Behjati[e], Seyedeh Sedigheh Abedini[e], Marzieh Mohseni[e], Lars R. Jensen[d], Hao Hu[d], Brenda Huang[c], Sara N. Stahley[a,g], Guanglu Liu[f], Kathryn R.Williams[a,g], Sharon. K Burdick[a], Yue Feng[f], Subhabrata Sanyal[a], Gary J. Bassell[a], Hans Hilger Ropers[d], Hossein Najmabadi[e], Anita H. Corbett[b,2], Kenneth H. Moberg[a,2], Andreas W. Kuss[d,2,3]

Departments of [a] Cell Biology, [b]Biochemistry, [f]Pharmacology and Graduate Programs in [c]Genetics and Molecular Biology, [g]Biochemistry, Cell and Developmental Biology, Emory University School of Medicine, Atlanta, USA; [d]Max Planck Institute for Molecular Genetics, Berlin, Germany; and [e]Genetics Research Center, University of Social Welfare and Rehabilitation Sciences, Tehran, Iran

[1]These authors made equal contribution; [2]co-corresponding authors; [3]current address: Department of Human Genetics, Interfaculty Institute of Genetics and Functional Genomics, Ernst Moritz Arndt University, Greifswald, Germany.

Author contributions

CP generated the *dNab2* alleles and conducted the majority of fly experiments.

MG performed linkage analysis, mutation screening, mutation characterization, expression analysis and cell line immunostaining.

LHA generated the dNab2 antibody and performed bulk poly(A) tail length assays.

CG conducted FISH in vertebrate cultured neurons and tissues.

GL performed immunostaining on vertebrate brain sections.

SWL conducted western blotting of patient samples.

SMK created GST-dNab2 ZnF construct.

SKB contributed to the cloning of transgenes.

BH, JJN and KRW performed fly behavioral assays.

SNS conducted immunostaining on fly larvae.

SS, YF and GJB contributed to experimental design and analysis.

KK, AT, SSA, MM and HN recruited patients and/or contributed to their clinical characterization.

FB, LRJ and HH performed patient chromosome analyses.

HN and HHR initiated the human ARID project.

KHM and AHC initiated the fly project, and together with AWK, coordinated the combined project, designed the studies, interpreted the results and together with CP, MG, and HHR and drafted the manuscript.

Abstract

Here we report a new human intellectual disability disease locus on chromosome 14q31.3 corresponding to mutation of the *ZC3H14* gene, which encodes a conserved polyadenosine RNA-binding protein. We identify *ZC3H14* mRNA transcripts in the human central nervous system, and find that rodent ZC3H14 protein is expressed in hippocampal neurons and co-localizes with poly(A) RNA in neuronal cell bodies. A *Drosophila melanogaster* model of this disease created by mutation of the gene encoding the ZC3H14 orthologue, dNab2, which also binds polyadenosine RNA, reveals that dNab2 is essential for development and required in neurons for normal locomotion and flight. Biochemical and genetic data indicate that dNab2 restricts bulk poly(A) tail length *in vivo*, suggesting that this function may underlie its role in development and disease. These studies reveal a conserved requirement for *ZC3H14/dNab2* in the metazoan nervous system and identify the first poly(A) RNA-binding protein associated with a human brain disorder.

Introduction

Unraveling the complex networks underlying brain function is a challenging problem for both basic and medical science. One way to understand brain function is to identify and characterize genes that when mutated impair normal human intellectual development. Intellectual disability (ID), previously referred to as mental retardation, is characterized by limited intellectual capacities reflected by an IQ below 70 and major constraints in adaptive behavior (1). Therapeutic options for the treatment of ID are extremely limited and its comparatively high prevalence of about 2% renders this disorder a major socioeconomic burden (1).

During the course of a large-scale systematic study to identify autosomal recessive ID (ARID) causing genetic defects in large Iranian families with intellectually disabled children born from blood-related parents (2, 3), we identified a locus for unspecific or non-syndromic ARID (NS-ARID) on chromosome 14q31.3 corresponding to mutation of the *ZC3H14* gene in two independent families. *ZC3H14* encodes an evolutionarily conserved Cys_3His tandem zinc finger polyadenosine RNA-binding protein (4, 5). The founding member of this protein family, *S. cerevisiae* Nab2, is essential for viability and required for proper 3'end formation and poly(A) RNA export from the nucleus (6, 7). Although multiple tissue-specific splice variants of human ZC3H14 have been described (5), their function in multicellular organisms has not been examined.

To better understand ZC3H14/Nab2 function in metazoans, we exploited *D. melanogaster* as a model for the developmental consequences of *ZC3H14* loss in humans. Loss of the putative *Drosophila* ZC3H14 orthologue, dNab2, disrupts normal

development and impairs neural function. Using tissue-specific depletion, we identify a pan-neuronal requirement for *dNab2* in normal behavior. Biochemical and genetic analyses indicate that *dNab2* restricts bulk mRNA poly(A) tail length *in vivo* and suggest that this conserved function may underlie the effect of *dNab2* loss on development and behavior. Taken together, these studies reveal a conserved requirement for ZC3H14/dNab2 in the metazoan nervous system and identify the first poly(A) RNA-binding protein associated with a human brain disorder.

Results

The *ZC3H14* gene is mutated in ARID patients. To identify molecular causes of NS-ARID, we performed a large-scale autozygosity mapping and linkage analysis in a cohort of more than 200 consanguineous Iranian families (3). This analysis identified a novel NS-ARID locus on chromosome 14q31.3-q32.12 in a family with three affected males (Figure 2.1*A*; Table S2.1). The linkage interval had the maximum attainable LOD score of 2.7 (Figure S2.1*B*) and no other autosomal linkage intervals were observed according to the "one-LOD-down" rule (8). A second significant interval was identified on chromosome Xp22.11-p11.4 (LOD=1.2) but this interval contained no sequence changes in protein coding regions (Figure S2.1*B*,*D*,*E*). Sequencing of all protein coding regions within the 14q31.3-q32.12 locus (Figure S2.1*F*,*G*) identified a homozygous nonsense mutation (R154X) in exon 6 of the *ZC3H14* gene (Figure 2.1*B*; Figure S2.2*A*), which co-segregated with the disease. This mutation was absent in 1864 chromosomes from healthy individuals, including 1184 from ethnically matched controls, 310 from German controls and 370 from the 1000-genome pilot projects 1 and 2 (9). Moreover, screening of the entire gene in a subset of 330 chromosomes from the Iranian controls and the 370 chromosomes from the 1000-genome project detected no deleterious mutations.

The *ZC3H14* gene encodes a poly(A) RNA-binding protein with similarity to *S. cerevisiae* Nab2 (4, 5). As shown in Figure 2.1B, ZC3H14 is alternatively spliced to encode four ZC3H14 protein isoforms (5). The R154X mutation is predicted to disrupt the ubiquitously expressed longer isoforms 1-3, but not the shorter brain and testes enriched isoform 4. Immunoblot analysis using an anti-ZC3H14 antibody raised against the N-terminal PWI-like domain of isoforms 1-3, which exclusively recognizes isoforms

1-3 (5), confirmed that R154X patient-derived lymphoblasts lack ZC3H14 isoforms 1-3 (Figure 2.1*C*). Parallel staining of R154X patient fibroblasts with a commercial ZC3H14 antibody that recognizes all four ZC3H14 isoforms (Abcam) revealed no detectable nuclear ZC3H14 (isoforms 1-3) while the cytoplasmic pool of protein, corresponding to isoform 4 (5), was still present (Figure 2.1*D*, *bottom panel*). Subsequent sequencing of *ZC3H14* in a second family showing NS-ARID, and maximum attainable LOD-score (2.5) to the same chromosome 14 linkage interval, revealed a 25 base pair (bp) deletion located 16 bps downstream of the 3'-end boundary of the annotated common exon 16 of *ZC3H14* (Table S2.1; Figure 2.1*B*; Figure S2.1*A,C,H*; Figure S2.2*B-D*). This mutation co-segregated with the patient phenotype and was not found to be homozygous in 831 control individuals.

The ZC3H14 protein is expressed in the central nervous system and co-localizes with poly(A) mRNA in hippocampal neurons. We confirmed that *ZC3H14* is expressed in the brain as *ZC3H14* transcripts were readily detected in adult and fetal human brain samples by RT-PCR (Figure 2.2*A*). Immunostaining of sections of adult mouse brain revealed that ZC3H14 protein is enriched in hippocampal neurons relative to glia (Figure 2.2*B*). Furthermore, poly(A) RNA-specific fluorescence *in situ* hybridization (FISH) in combination with anti-ZC3H14 immunostaining using the anti-PWI-like domain antibody demonstrated that ZC3H14 and poly(A) RNA co-localize in nuclear speckles in both the pyramidal layer of mouse CA1 (Figure 2.2*C*) and cultured rat hippocampal neurons (Figure 2.2*D*).

dNab2 is a putative *D. melanogaster* orthologue of ZC3H14 and a member of an evolutionarily conserved class of zinc finger (ZnF) polyadenosine RNA-binding proteins. We next exploited *Drosophila* as a system to understand tissue-specific roles and requirements for *ZC3H14* in metazoans. Based on sequence similarity and domain conservation, we identified the uncharacterized gene *CG5720* (Flybase.org) as the putative *Drosophila* ZC3H14/Nab2 orthologue (dNab2) (Figure 2.3*A*). The conserved dNab2 C-terminal tandem Cys_3His zinc finger domain (Figure 2.3*A*), which mediates polyadenosine RNA-binding in other species (4, 10), showed preferential binding to polyadenosine RNA *in vitro* (Figure 2.3*B*). The intracellular localization of dNab2 mirrored the localization of both Nab2 (7) and ZC3H14 (isoforms 1-3) (5) as immunostaining for dNab2 revealed nuclear expression throughout development in all tissues examined including the nervous system (Figure S2.3*A,B*).

***dNab2* is essential for normal development.** To determine whether *dNab2* contributes to development or function of the nervous system, we created *dNab2* alleles by imprecise excision of a P-element (*P{EPgy2}EY08422*) located upstream of the *dNab2* gene (Figure 2.3*C*). Five excisions of *EY08422* (*ex1-ex5*) were recovered with genomic deletions ranging from 0.9 kilobases (kb) to 1.5 kb that extend into the *dNab2* gene (Figure 2.3*C*); all of these alleles failed to express *dNab2* mRNA and protein (Figure 2.3*D*; Figure S2.3*C*). The $dNab2^{ex3}$ null allele was used for all subsequent experiments. Through mid-pupal development, $dNab2^{ex3}$ homozygous mutants showed no evidence of morphological or behavioral defects, reduced viability, or developmental delay. However, a majority of the $dNab2^{ex3}$ homozygotes died during late pupal phase and displayed eclosion defects (Figure 2.3*E*); the few remaining $dNab2^{ex3}$ mutants (~5%) emerged completely but

exhibited a shortened lifespan (~1.5 weeks) and morphological defects reminiscent of other mutants in RNA-binding proteins, including those required for neuronal function (11, 12). These phenotypes include 'wings-held out' in which flies fail to fold their wings together over the dorsal surface of the thorax and abdomen (Figure 2.3*G*) and disorganization and bending of thoracic bristles (Figure 2.3*J,K*). All of these phenotypes were also present in animals carrying *dNab2*^{ex3} *in trans* to genomic deletions (deficiencies) that completely remove the *dNab2* gene (Figure 2.3*F,H,L,M*; Figure S2.3*D-K*) but were absent in control flies that are homozygous for a precise excision (*p-ex*) of the *EY08422* element (Figure 2.3*I,N,O*). Embryos lacking germline contribution of *dNab2* died early in embryogenesis, indicating that *dNab2* is required both for embryonic viability and development.

***dNab2* is required in neurons for normal *Drosophila* behavior.** In addition to morphological defects, *dNab2*^{ex3} mutant flies or those carrying *dNab2*^{ex3} *in trans* to an uncovering deficiency displayed severely compromised flight behavior and poor locomotor activity (Figure 2.4*A*; Figure S2.4*A,B*). However, zygotic loss of *dNab2* caused no changes in expression patterns of the neuronal marker Elav or the pre-synaptic active zone marker Nc82 that were evident at the level of the whole brain (Figure S2.3*L-O*). Similarly, loss of *dNab2* had no detectable effect on gross synaptic structure, based on analysis of synaptic bouton number and organization at the larval muscle 6/7 neuromuscular junction (Figure S2.3*P*).

To examine tissue-specific requirements for *dNab2*, we used an inverted repeat (IR) *dNab2* RNA-interference transgene, which reduced dNab2 protein levels *in vivo* (Figure S2.4*E*), to deplete dNab2 from specific tissues. Pan-neuronal knockdown of

dNab2 (*Elav-Gal4>Dcr2, IR*) caused flight defects indicating that *dNab2* is required in neurons for normal flight behavior (Figure S2.4*A*). Furthermore, pan-neuronal knockdown of *dNab2* strongly recapitulated the locomotor defect of the *dNab2* genomic null allele (Figure 2.4*A*), as well as an age-dependent decline in locomotor ability similar to that of $dNab2^{ex3}$ flies. (Figure S2.4*C, C'*). A modest decline in locomotor activity was also observed in flies depleted of *dNab2* specifically in motor neurons (*OK6-Gal4>IR*) (Figure S2.4*D*) suggesting that *dNab2* acts within multiple types of neurons to support normal locomotor behavior. In contrast, pan-muscle *dNab2* knockdown flies generated by using two independent muscle-specific *Gal4* drivers (*Mef2-Gal4* or *Mhc-Gal4*) showed a negative geotaxis response indistinguishable from controls (Figure 2.4*A*; Figure S2.4*F*) revealing that *dNab2* may be specifically required in neurons for normal locomotor activity.

To confirm the neuronal requirement for *dNab2*, we tested whether a transgene expressing wild type *dNab2* (*UAS-dNab2-Flag)* only in neurons could rescue the $dNab2^{ex3}$ mutant phenotype. The *UAS-dNab2* transgene alone mildly rescued eclosion rates and locomotor defects among *dNab2* mutants in the absence of a *Gal4* driver (rescue control) (Figure 2.4*B,C*). Consistent with this rescue, a low level of leaky expression from the transgene was confirmed by immunoblotting. Importantly, restoring *dNab2* expression pan-neuronally (*Elav-Gal4*) in *dNab2* null flies completely rescued the eclosion and locomotor defects (Figure 2.4*B,C*). In control experiments, ubiquitous overexpression of *dNab2* was lethal to wild type flies but those overexpressing *dNab2* from a pan-neuronal driver (*Elav-Gal4*) were viable and performed normally in locomotor assays (Figure S2.4*G,G'*) indicating that rescue by neuronal expression of

dNab2 was not due to enhanced performance. Thus, expression of *dNab2* only in neurons appears to be sufficient to rescue viability and behavior in animals otherwise lacking *dNab2*.

***dNab2* is required for proper control of poly(A) tail length.** To begin to understand the molecular role of *dNab2* in the nervous system, we utilized a genetic-modifier approach to screen a small collection of alleles of select genes for their ability to modify a rough-eye phenotype produced by overexpressing wild type *dNab2* in the differentiating neurons of the eye (Figure 2.5*A*, *panels i-ii*). While most alleles tested showed little or no effect, we observed robust and fully penetrant genetic interactions between *dNab2* and two components of the polyadenylation machinery (Table S2.2): the poly(A) polymerase *hiiragi* (*hrg*) (13) and *Pabp2*, the *Drosophila* orthologue of the nuclear poly(A)-binding protein PABPN1/PABP2 that promotes polyadenylation of mRNAs (14). Heterozygosity for an *hrg* loss-of-function allele, hrg^{10} (15), strongly enhanced the *dNab2*-driven adult eye phenotype, but had no dominant effect on eye morphology in a wild type background. Similarly, the $Pabp2^{55}$ loss-of-function allele dominantly enhanced the *dNab2* overexpression phenotype, resulting in a smaller, more disorganized, and blackened eye (Figure 2.5*A*, *panel iii*). Reciprocally, overexpression of *Pabp2* using the *EP2264* allele (16) significantly suppressed the *dNab2*-driven rough-eye phenotype (Figure 2.5*A*, *panel iv*), while expression of a control *UAS-eGFP* transgene had no effect (Figure 2.5*A*, *panel v*). This qualitative modification of the *dNab2*-driven rough, small eye phenotype by *Pabp2* alleles was confirmed by two-dimensional quantification of eye size (Figure 2.5*B*).

To determine whether the genetic interactions between *dNab2* and *Pabp2* in the adult eye reflect a role for *dNab2* in poly(A) tail length control, bulk RNA poly(A) tail length was measured and quantified in adult heads (Figure 2.5*C,D*). As described for *S. cerevisiae nab2* mutants (6, 10), poly(A) tail length was increased in *dNab2*^{ex3} mutant heads (Figure 2.5*C*, *lane 2*) relative to *p-ex* controls (Figure 2.5*C*, *lane 1*). Similar data were obtained from analysis of RNA isolated from whole flies (Figure 2.5*E*). Reciprocally, overexpression of *dNab2* in the eye shortened bulk poly(A) tail length relative to *p-ex* control, and consistent with the genetic modification data, this molecular effect was rescued by co-overexpression of *Pabp2* (Figure 2.5*C,D*). These biochemical effects thus parallel the genetic relationship between *dNab2* and *Pabp2*, and argue strongly that these proteins have antagonistic effects on bulk poly(A) tail length in neuron-enriched adult tissues such as the eye.

In contrast to the effect of *dNab2* on polyadenylation, no obvious change in poly(A) RNA localization was apparent in clones of *dNab2*^{ex3} mutant larval wing disc cells (-/-) generated in the background of wild type control cells (+/+) (Figure 2.5*F*, *top panels*). As a control, cells mutant for the mRNA export receptor *sbr* (17) showed nuclear accumulation of poly(A) RNA (Figure 2.5*F*, *bottom panels, see arrowheads*). Thus, loss of *dNab2* appears to dysregulate poly(A) tail length and cellular function without perturbing bulk poly(A) RNA export from the nucleus.

Discussion

In an effort to better understand molecular and cellular processes that underlie normal brain function, we have sought to identify mutations that lead to ID in the human population. Here we identify mutations in the human *ZC3H14* gene in human patients with NS-ARID and create a tractable genetic model that recapitulates key phenotypic elements of the human disease. We show that the *Drosophila* ZC3H14 orthologue dNab2 regulates RNA poly(A) tail length, and that loss of dNab2 leads to extended RNA poly(A) tails. The effect of *dNab2* on RNA poly(A) tail length coupled with neuronal-specific behavioral phenotypes seen in *dNab2* mutant flies and *ZC3H14*-associated non-syndromic ARID patients provides evidence that *dNab2*-mediated control of RNA poly(A) tail length is required for normal neuronal function.

Loss of *dNab2* in *Drosophila* or *NAB2* in budding yeast causes an increase in bulk RNA poly(A) tail length (6), yet the mechanism by which these hyperadenylated mRNAs contribute to neuronal dysfunction or possibly to human disease is not established. Likely consequences of hyperadenylated mRNAs could include altered transcript stability, titration of critical poly(A) RNA-binding proteins and/or bypass of cytoplasmic polyadenylation necessary for activity-dependent translation of neuronal mRNAs. Individually or in combination, these defects could disrupt spatiotemporal control of gene expression needed for development of the nervous system and higher order brain function. Thus, we speculate that ZC3H14/dNab2 could play critical roles in neurons such as ensuring that transcripts are properly targeted to sites of localized translation. This hypothesis is consistent with a report that budding yeast Nab2 aids in targeting transcripts to the bud site (18). Alternatively, ZC3H14 and dNab2 may regulate a set of

mRNAs that play key roles in neurons, such that ZC3H14/dNab2 loss disproportionately affects these cells. These mechanisms could explain why mutation of ubiquitously expressed post-transcriptional regulatory factors such as dNab2 and ZC3H14 leads to neuronal defects in flies and, more critically, to NS-ARID in humans.

Although the *NAB2* and *dNab2* genes are essential (7), loss of the corresponding forms of the ZC3H14 protein (isoforms 1-3) in humans appears to selectively impair brain function as patients display non-syndromic intellectual disability. At this early stage of investigation, it is unclear whether ZC3H14 is simply not essential in humans or whether the remaining cytoplasmic isoform of the protein, isoform 4, suffices in all tissues except the brain. Alternatively, we cannot rule out the possibility that a protein that is functionally redundant with ZC3H14 exists; however, the human genome does not encode any apparent sequence orthologues of ZC3H14. As isoform 4 is only expressed in mammals (5), future studies exploiting mammalian model systems will be required to address the functional requirements for specific isoforms of ZC3H14 as they relate to human intellectual disability.

The identification of *ZC3H14* mutations in non-syndromic ARID places ZC3H14/dNab2 among several other RNA-binding proteins implicated in human diseases that impact neural function (19). However, this study identifies the first poly(A) RNA-binding protein altered in a human brain disorder and provides insight into the molecular basis of intellectual disability and brain function.

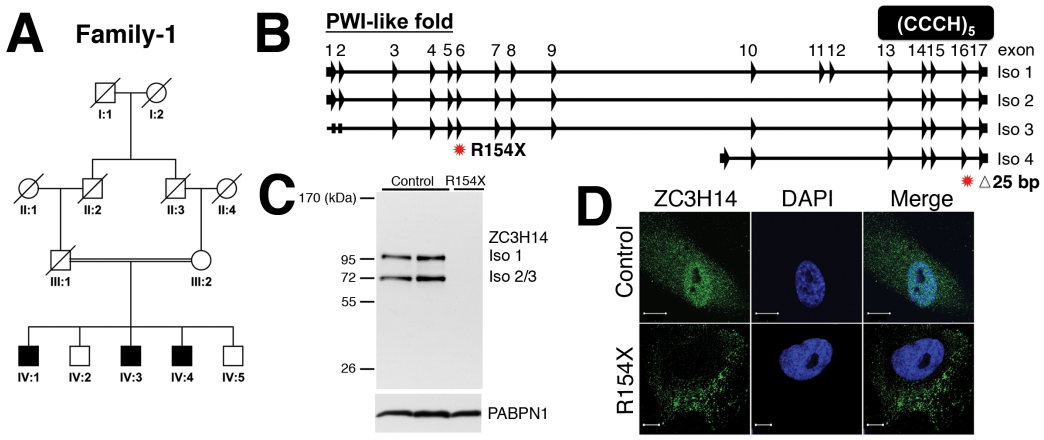


Figure 2.1. *ZC3H14* is mutated in NS-ARID patients. (*A*) Pedigree of Family-1. (*B*) Schematic of four *ZC3H14* splice variants indicating exons encoding the N-terminal PWI-like domain and C-terminal Cys_3His zinc finger RNA-binding motif (CCCH) domain. Positions of patient mutations are indicated by red stars. (*C*) Anti-ZC3H14 immunoblot of two control lymphoblast lines and one derived from a Family-1 R154X patient. The ZC3H14 antibody (5) recognizes ZC3H14 isoforms 1 and 2/3. Anti-PABPN1 is shown as a loading control (20). (*D*) Immunofluorescent detection of ZC3H14 (green) in control or patient (R154X) fibroblasts using a commercial ZC3H14 antibody (Abcam) directed against the common zinc finger domain. DAPI (blue) marks nuclei. Scale bar=10 μm.

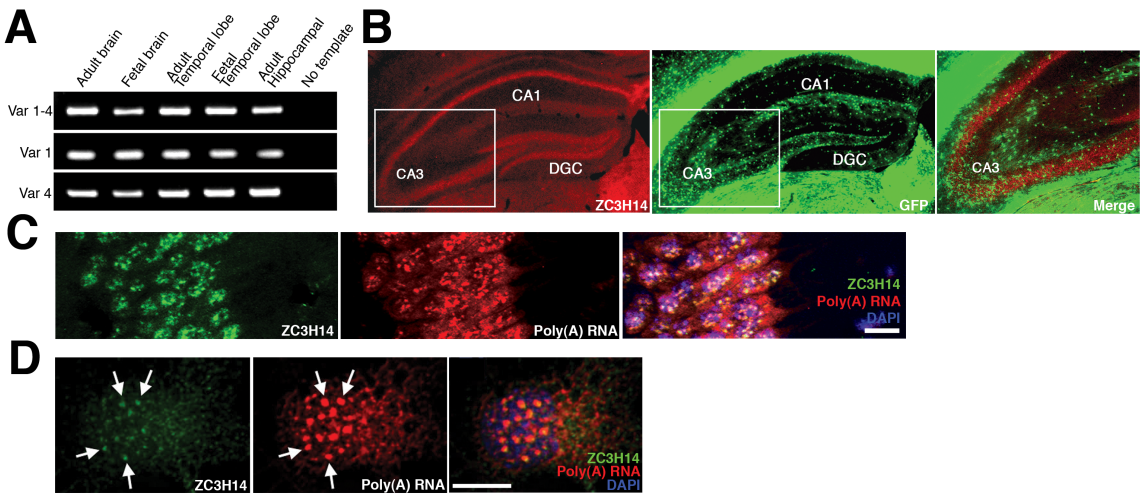


Figure 2.2. ZC3H14 is expressed in vertebrate hippocampal neurons. (*A*) RT-PCR analysis of ZC3H14 splice variants: variants 1-4 (top panel), variant 1 (middle panel), or variant 4 (bottom panel) from indicated tissues. For panels *B*, *C*, and *D*, ZC3H14 was detected with the N-terminal antibody that recognizes ZC3H14 isoforms 1 and 2/3 (5). (*B*) Immunofluorescent detection of ZC3H14 protein (red) in a mouse hippocampal section expressing oligodendroglia-GFP (green) (21). CA1 and CA3 (cornu ammonis fields 1 and 3) and DGC (dentate gyrus granular cells) regions of the hippocampus are indicated. White box indicates the zoom-in region in the merge. (*C*) Adolescent mouse brain sections probed with an oligo-dT FISH probe to detect poly(A) RNA (red) and co-stained for ZC3H14 (green). ZC3H14 appears in poly(A) RNA-positive nuclear speckles in hippocampal pyramidal neurons. DAPI (blue) marks nuclei. Scale bars=20 μm. (*D*) Poly(A) RNA FISH (red) and indirect immunofluorescence in cultured rat embryonic hippocampal neurons reveals co-localization of ZC3H14 protein (green) with poly(A) RNA speckles in the nucleus (blue). Scale bar=5 μm.

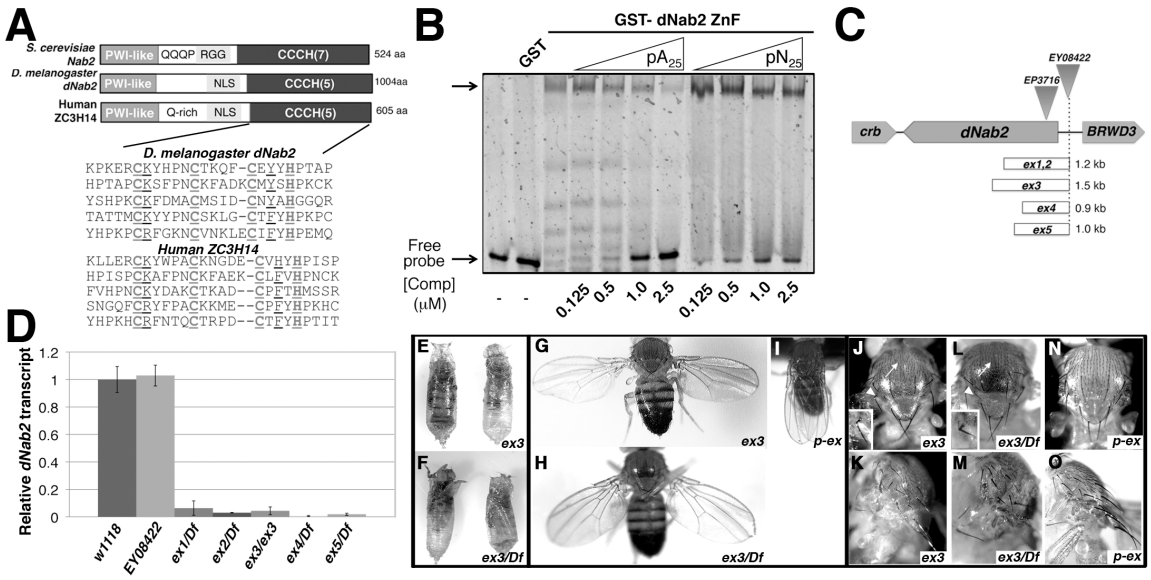
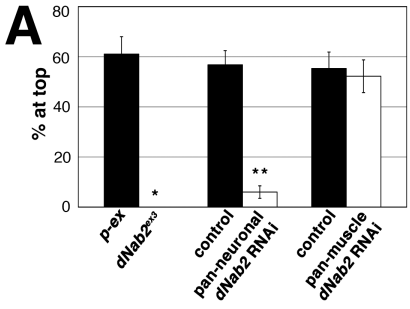


Figure 2.3. dNab2 is a putative *D. melanogaster* orthologue of ZC3H14. (*A*) Domain alignment of *S. cerevisiae* Nab2, *Drosophila* dNab2 and human ZC3H14. The conserved N-terminal PWI-like fold, Q-rich, RGG/predicted nuclear localization signal (NLS) and C-terminal tandem Cys_3His zinc finger RNA-binding motif (CCCH) domains are indicated (5, 10). Amino acid alignment of the five Cys_3His tandem zinc fingers from fly dNab2 and human ZC3H14 show conserved spacing and intervening basic and aromatic residues (underlined) that are required for RNA binding in *S. cerevisiae* Nab2 (10). (*B*) RNA-binding properties of dNab2 analyzed by RNA electrophoretic mobility shift assay. GST-dNab2 ZnF (zinc fingers 1-5), but not GST, binds to polyadenosine 25-mer (pA_{25}) RNA. Top arrow indicates a shift. Unlabeled pA_{25}, but not randomized polyN 25-mer (pN_{25}) RNA competitor oligonucleotide, competes efficiently for binding to dNab2 ZnF. (*C*) Schematic of the *dNab2* locus indicating the location of the *EP3716* and *EY08422* elements (inverted triangles) and the five imprecise excision alleles (*ex1-ex5*). (*D*) qRT-PCR analysis of *dNab2* transcript levels in adults flies. All genotypes were analyzed in triplicate and normalized to *dNab2* transcript levels in w^{1118} control animals (set to 1.0).

β-tub is an internal control. Error bars=s.d. (*E-O*) Light microscopic images of adult flies of indicated genotypes. (*E-F*) The majority of *dNab2*^{ex3} and *dNab2*^{ex3}/*Df* (Df(3R)Exel8178) animals die at pharate adult stage, often as partially emerged adults. (*G-H*) The remainder emerge with a 'wings held-out' phenotype that is (*I*) absent in controls (*p-ex*). (*J-O*) Front and side views of the thorax showing thoracic bristles. *dNab2*^{ex3} (*J-K*) and *dNab2*^{ex3}/*Df* (*L-M*) mutants show bent major thoracic bristles (arrowhead in (*J, L*) and enlarged in the inset) and disorganized minor thoracic bristles (arrow in (*J, L*)), compared to *p-ex* controls (*N-O*).



B

Genotype	% eclosed (of expected)
dNab2^{ex3}/dNab2^{ex3}	~3%
UAS-dNab2-Flag;;dNab2^{ex3}/dNab2^{ex3} (rescue control)	39%
Elav-Gal4>UAS-dNab2-Flag;; dNab2^{ex3}/dNab2^{ex3} (pan-neuronal rescue)	93%

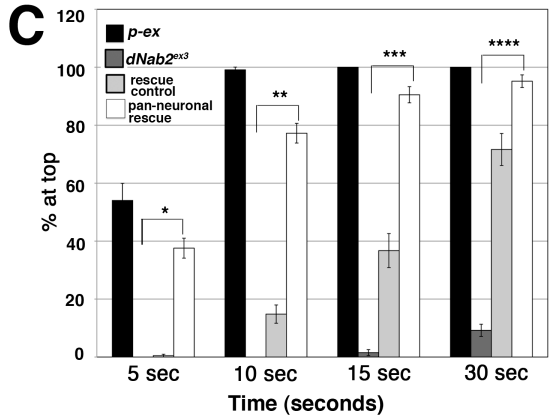


Figure 2.4. A neuronal-specific requirement for *dNab2* in normal behavior. (*A*) Locomotor phenotypes of genomic alleles and tissue-specific RNAi of *dNab2*. Data are presented as the average % of flies that reach the top of a cylinder after 5 seconds across all trials. Groups of ten 5-day old flies were tested for at least ten independent trials per genotype (*p=8.36x10^{-9} and **p=8.38x10^{-9} in a two-tailed *t*-test). Error bars=s.e.m. (*B-C*) Pan-neuronal expression of dNab2 rescues both eclosion and locomotor defects. (*B*) Table summarizing percentage of flies eclosed (of expected) for indicated genotypes. (*C*) For the locomotor assay, data are presented as the average % of flies that reach the top of a cylinder after indicated time points across all trials. Groups of ten 2-day old flies were

tested for at least ten independent trials per genotype (*p=5.12×10^{-13}, **p=6.59×10^{-13}, ***p=1.97×10^{-7}, ****p=0.002 in a two-tailed *t*-test). Error bars=s.e.m.

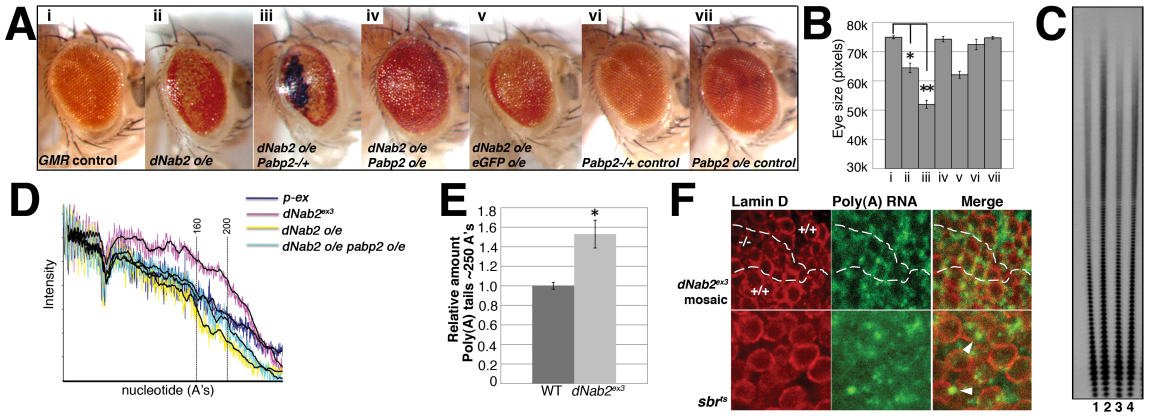


Figure 2.5. *dNab2* regulates poly(A) tail length. Light microscopic images of adult fly eye (*A*) and quantification of eye size (pixels) (*B*) for the following genotypes: (i) *GMR-Gal4*/+ (*GMR* control), (ii) *GMR-Gal4*/+;$dNab2^{EP3716}$/+ (*dNab2 o/e*), (iii) *GMR-Gal4*/$Pabp2^{55}$;$dNab2^{EP3716}$/+ (*dNab2 o/e Pabp2-/+*), (iv) *GMR-Gal4*/$Pabp2^{EP2264}$;$dNab2^{EP3716}$/+ (*dNab2 o/e Pabp2 o/e*), (v) *GMR-Gal4*/+;$dNab2^{EP3716}$/*UAS-eGFP* (*dNab2 o/e eGFP o/e*), (vi) *GMR-Gal4*/$Pabp2^{55}$ (*Pabp2-/+ control*), (vii) *GMR-Gal4*/$Pabp2^{EP2264}$ (*Pabp2 o/e control*). o/e=overexpression. * p=0.0006, **p=0.0014 in a two-tailed *t*-test. n=5 per genotype. Error bars=s.e.m. (*C*) Bulk poly(A) tail length measurements in heads of 1) control (*p-ex*) 2) $dNab2^{ex3}$ 3) *GMR-Gal4*/+;$dNab2^{EP3716}$/+ 4)*GMR-Gal4*/$Pabp2^{EP2264}$;$dNab2^{EP3716}$/+. (*D*) Densitometric quantification of poly(A) tracts (Image J) from (*C*) showing poly(A) tail length profiles of the indicated genotypes (highlighted in colored lines). (*E*) Bulk poly(A) tail length from whole flies as analyzed by densitometric quantification of poly(A) tracts (Image J) of ~250 nucleotides (nt) normalized to poly(A) tracts of ~100 nt for WT (w^{1118}) and $dNab2^{ex3}$. w^{1118} control was set to 1.0. (*p<0.04, n=3, two-tailed *t*-test). Error bars=s.d. (*F*) Poly(A) RNA localization was analyzed in wing disc cells subjected to FISH to visualize poly(A) RNA (green) and co-stained with anti-Lamin D (red) to visualize the nuclear

periphery. Top panels show a *dNab2*^{ex3} mosaic larval wing disc; dotted lines denote boundaries between wild type ('+/+') and *dNab2*^{ex3} mutant clones ('-/-'). Bottom panels show a wing disc homozygous for the *sbr*^{ts} allele (17) shifted to 33°C. Arrowheads indicate nuclei accumulating poly(A) RNA.

Supplemental Tables

Table S2.1. Clinical features of patients. Clinical features of affected patients from two families are described.

Patient	Sex	Age at examination	Intellectual disability[1]	Height	Occipitofrontal circumference
Family-1 (M233)					
IV:1	Male	21 y	Mild-Moderate	180cm (50^{th} centile)	57.5cm (97^{th} centile)
IV:3	Male	26 y	Mild-Moderate	170 cm (25^{th} centile)	56.5 cm (50^{th} centile)
IV:4	Male	28 y	Mild-Moderate	171 cm (25^{th} centile)	60 cm (>97^{th} centile)
Family-2 (M168)					
IV:1	Male	13 y	Severe	147 cm (10^{th} centile)	53.5 cm (50^{th} centile)
IV:2	Male	15 y	Severe	163 cm (10^{th} centile)	54 cm (50^{th} centile)
IV:4	Male	17 y	Severe	186 cm (75^{th} centile)	54.5 cm (50^{th} centile)

[1]Degree of intellectual disability: mild IQ 50–75, moderate IQ 35–50, severe IQ<35

Table S2.2. Modifiers of *dNab2* eye-overexpression phenotype. Genes indicated were tested for their ability to modify a rough-eye phenotype produced by overexpression of wild type *dNab2* in differentiating neurons of the developing eye (*GMR-Gal4/+;dNab2*^{EP3716}/+). All alleles tested show no dominant effect on eye morphology in a wild type background.

Gene	Allele	Result[a]
abstrakt	00620	-
	f01698	-
bruno-2	EY18918	-
	MB00431	-
hiiragi	10	*Enhanced*
Hrb87F	KG02089	-
Hrb98DE	ZCL0558	-
mushroom-body expressed	04093	*Suppressed*
NTF2-related export protein 1	f04855	-
	DG05102	-
Pabp2	55	*Enhanced*
	EP2264	*Suppressed*
pumilio	3	-
	bem	-
small bristles	1	-
	magellan	-

[a] (-) indicates 'no interaction'

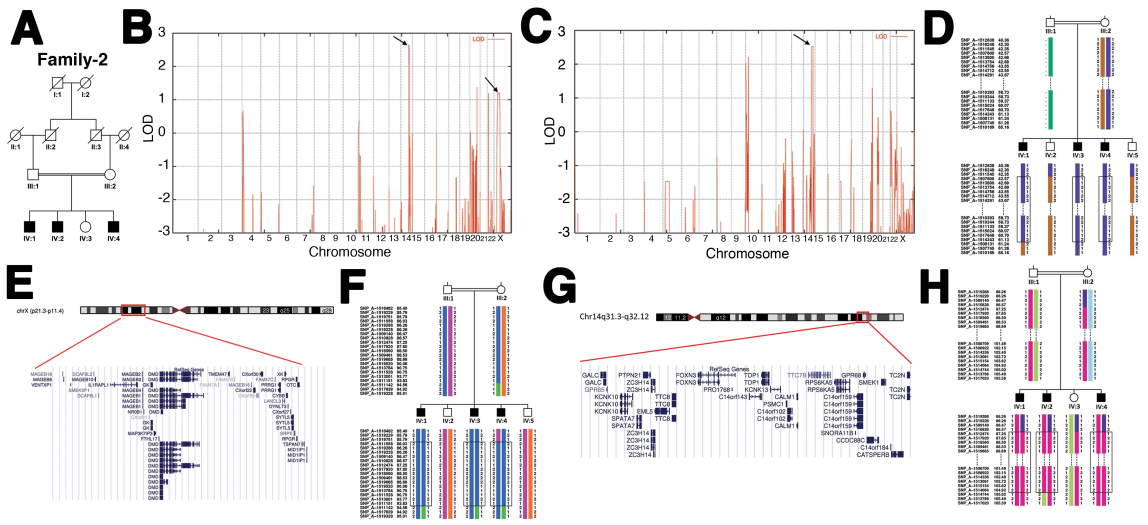


Figure S2.1. Pedigree and genetic analysis of affected ARID patients. (*A*) Pedigree of Family-2. Genome-wide multipoint parametric linkage analysis identifies two candidate ARID regions at 14q31.3-q32.12 and Xp22.11-p11.4 (indicated by arrows) for Family-1 with maximum LOD scores of 2.7 and 1.2 respectively (*B*), and one region (arrow) with a maximum LOD score of 2.5 at 14q31.3-q32.12 (arrow) in Family-2 (*C*). (*D*) Haplotype of the linkage interval on chromosome Xp22.11-p11.4 in Family-1. Markers from both borders of the interval on chromosome X are shown, which is defined by SNP_A-1511545 and SNP_A-1508131 (boxed). (*E*) Gene names within the intervals on chromosome X sequenced in Family-1 are indicated. (*F*) Haplotype of the linkage interval on chromosome 14q31.3-q32.12 in Family-1. The most inclusive region of homozygosity on chromosome 14 is located between SNP_A-1519751 and SNP_A-1511142 markers (boxed). (*G*) Gene names within the intervals on chromosome 14 sequenced in Family-1 are indicated. (*H*) Haplotype of the single linkage interval with maximum LOD score of 2.5 in Family-2. Markers from both borders of the interval on

chromosome 14 are shown. The most inclusive region of homozygosity is located between SNP_A-1510828 and SNP_A-1514744 (boxed).

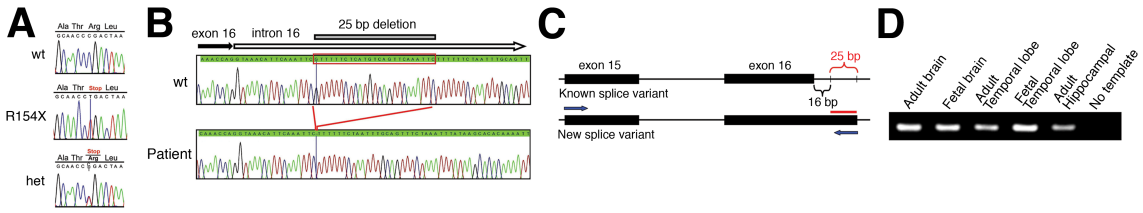


Figure S2.2. *ZC3H14* is mutated in non-syndromic ARID patients and the *ZC3H14* mRNA is expressed in the human brain. (*A* and *B*) DNA sequence traces for identifying lesions found in Family-1 (*A*) and Family-2 (*B*). (*C*) Schematic of the *ZC3H14* exon15/16 region and location of RT-PCR primers (blue arrows). (*D*) RT-PCR analysis to detect inclusion of Family-2 deleted sequences (red bar in (*C*)) in *ZC3H14* cDNAs prepared from indicated tissues.

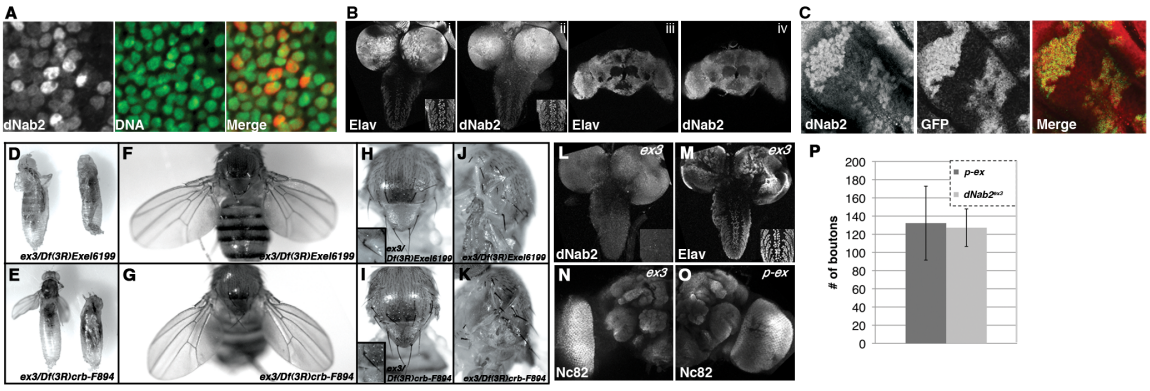


Figure S2.3. dNab2 localization, expression and organismal phenotypes. (*A*) Confocal image of dNab2-overexpressing larval wing imaginal disc cells (*Engrailed-Gal4>dNab2EP*^{3716}) co-stained to detect dNab2 protein (red) and DNA (green). (*B*) Confocal images of larval (i-ii) or adult (iii-iv) brains stained as indicated with anti-Elav (i and iii) or polyclonal anti-dNab2 (ii and iv). Insets in (i) and (ii) show expression of dNab2 and Elav in motor neuron nuclei in the ventral nerve cord (VNC). (*C*) Confocal image of a wing disc bearing *dNab2*^{ex3} homozygous mutant clones (GFP-negative; absence of green) and stained for dNab2 protein (red). (*D-K*) Light microscopic images of adult flies of *dNab2*^{ex3} in trans to deficiencies. (*D*,*E*) The majority of *dNab2*^{ex3} in trans to deficiencies die at pharate adult stage, often as partially emerged adults. (*F*,*G*) The remainder emerge as adults with a 'wings held-out' phenotype. (*H-K*) Front and side views of the thorax showing major and minor thoracic bristles. *dNab2*^{ex3} in trans to deficiencies show bent thoracic bristles (enlarged in the inset) as well as disorganization of minor thoracic bristles. (*L, M*) Confocal images of *dNab2*^{ex3} larval brains showing loss of dNab2 protein in the brain lobes and VNC (see inset), with no effect on expression of Elav or overall brain structure. (*N, O*) Confocal images of *dNab2*^{ex3} (*N*) and *p-ex* (*O*) adult brains stained with Nc82. *dNab2*^{ex3} adult brains show no effect on expression of synaptic marker Nc82 or overall brain structure. (*P*) Quantification of the average number of

Synaptotagmin-positive synaptic boutons per larval neuromuscular junction at muscle 6/7 in *p-ex* (dark gray) and *dNab2*^{ex3} (light gray) larvae (p=0.748, n=10 for each genotype, two-tailed t-test). Error bars=s.d.

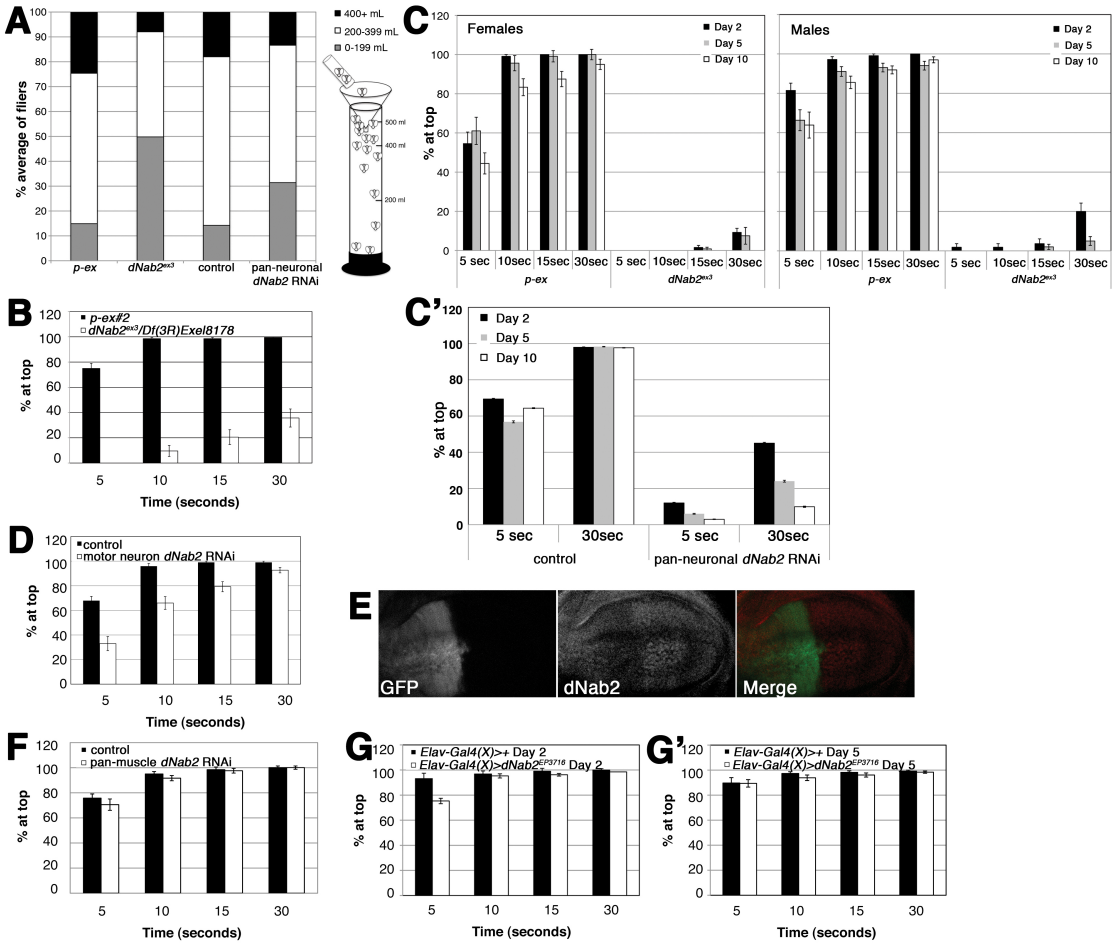


Figure S2.4. dNab2 is required for normal flight and locomotor behavior. (*A*) A flight assay was performed on *p-ex*, *dNab2*^{ex3}, control (*Elav-Gal4>dcr2*) and pan-neuronal *dNab2* RNAi (*Elav-Gal4>dcr2, IR*) flies using the apparatus depicted to the right (see *Materials and Methods*). For all genotypes, the stacked bar graph indicates the average % of flies landing in each vertical section of the 500-mL graduated cylinder: 1-199 mL (dark gray), 200-399 mL (white) and 400+ mL (black). Groups of twenty 5-day old female flies were tested in at least ten independent trials for each genotype. The 2x2 tables comparing *p-ex* vs. *dNab2*^{ex3} and control vs. pan-neuronal *dNab2* RNAi (high fliers >200-mL mark compared to low fliers <200-mL mark) yielded two-sided p-values of <0.0001 calculated by Fisher's exact test. (*B*) Results of the negative geotaxis assay for

control *p-ex#2* (a second independent precise excision allele) and $dNab2^{ex3}$ in trans to deficiency ($dNab2^{ex3}$/*Df(3R)Exel8178*) at Day 5 are shown. (*C*) A negative geotaxis assay measuring climbing ability was performed on control adult (*p-ex*) and mutant adult ($dNab2^{ex3}$) flies at Day 2, Day 5 and Day 10 independently for both females and males. The climbing defect is independent of sex but is enhanced with age. (*C'*) Plot of age-dependent decline in climbing ability of pan-neuronal *dNab2* knockdown flies at Day 2, 5 and 10 for both 5-second and 30-second time points. Repeated measures ANOVA showed significant differences by genotype ($p<0.0001$) and age ($p=0.01$), as well as a nominally significant interaction between age and genotype ($p=0.04$). (*D*) Results of the negative geotaxis assay for motor neuron *dNab2* RNAi flies (*OK6-Gal4>IR*) compared to controls (*OK6-Gal4>+*) at Day 5 are shown. (*E*) Wing imaginal disc cells knocked down for *dNab2* are positively marked by GFP expression (green) (*Engrailed-Gal4>UAS-GFP, IR*). Antibody staining for dNab2 (red) reveals reduced protein level in the GFP-positive cells compared to neighboring control GFP-negative cells. (*F*) Results of the negative geotaxis assay for pan-muscle *dNab2* RNAi flies (*Mhc-Gal4>IR*) compared to controls (*Mhc-Gal4>+*) at Day 2 are shown. (*G, G'*) Results of the negative geotaxis assay measuring climbing ability for neuronal cell type-specific dNab2 overexpression flies (*Elav-Gal4(X)>*$dNab2^{EP3716}$) compared to control (*Elav-Gal4(X)>+*) at Day 2 and Day 5. (*B, C* and *E-F'*) Data are presented as the average percentage of flies reaching the top of a 25-mL graduated cylinder after 5, 10, 15 and 30 seconds across all trials. Groups of ten flies were tested in at least ten independent trials for each genotype. Error bars=s.e.m.

Experimental procedures

Subjects: Sample collection and clinical evaluation were carried out as previously described (1) with the informed, written consent of the parents.

Genetic analyses: All affected members, parents and healthy siblings from both families were genotyped using the Affymetrix GeneChip® Human Mapping 10K Arrays.

Statistical significance and multiple testing corrections: To verify the relationship between individuals, the data were subjected to standard quality control routines, including graphical representation of relationship errors (GRR) (2) and gender check by calculating the number of heterozygous markers on the X chromosome for each genotype. Additionally, mendelian inconsistencies and unlikely genotypes were detected by the PedCheck (3) and Merlin (4) programs, respectively, and excluded prior to linkage analysis.

Linkage disequilibrium: To generate input files with the appropriate format for the linkage analysis programs, we utilized ALOHOMORA software (5). The two-point and multipoint linkage analyses, using the Genehunter (6), Allegro (7) and Merlin softwares, were performed assuming a fully penetrant autosomal recessive trait with a disease frequency of 0.001 and no phenocopies. Non-parametric linkage analysis was performed using Merlin and Genchunter softwares.

Haplotype analyses: Haplotypes were constructed using the Merlin and Allegro programs and visualized by the Haplopainter software (8).

PCR: Standard PCR methods were employed for all mutation mapping and sequencing. Sequences of primers employed for mapping human mutations on 14q31.3-q32.12 are

available upon request. Primers employed for fly deletion mapping are also available upon request.

RT-PCR/Quantitative real-time RT-PCR (qRT-PCR): Total RNA from both human and *Drosophila* samples was extracted using TRIzol reagent (Invitrogen). Total RNA samples from human adult and fetal brain (Biocat, Cat#: R1234035-50-BC, Cat#: R1244035-50-BC), fetal and adult temporal lobe (Biocat, Cat#: R1234078-50-BC and Cat#: R1244078-50-BC) and hippocampus (Biocat, Cat#: R1234052-10-BC) were purchased. cDNAs were generated by using the SuperScript III Reverse Transcriptase (Invitrogen). Quantitative Real-time PCR was performed according to standard protocols with LightCycler 480 (9). Primer sequences are available upon request.

Protein expression, purification and generation of antibody: A DNA fragment containing the entire dNab2 open reading frame was cloned into pGEX-4T (Amersham Biosciences) to create a GST-tagged fusion protein encoding full-length dNab2 for antibody generation and a DNA fragment encoding amino acids 839-1004 was cloned into pGEX-4T to create a GST-tagged fusion protein encoding the zinc finger domain for *in vitro* binding assay. GST-tagged proteins were expressed in *E. coli* and purified as described (10).

Tissue preparation and neuronal cell culture: Brain tissue from C57/B6 mice (postnatal day 21) was prepared as described previously (11) and frozen tissue was sectioned at 12 μm. Embryonic day-18 rat primary hippocampal neurons were cultured as described previously (12) and processed for immunostaining and FISH at three days *in vitro*.

Immunoblotting and immunostaining: Immunoblotting was performed as described

previously (13). Antibodies used were rabbit anti-ZC3H14 (1:10,000) (13); rabbit anti-PABPN1 (1:5000) (14). Immunostaining was performed as described previously for human cell lines (13), mounted mouse tissue section (15), and *Drosophila* tissues (16, 17). Antibodies used were mouse anti-ZC3H14 (Abcam, 1:500); chicken anti-GFP (Aves Lab, 1:200); rabbit anti-dNab2 (1:250 for discs, 1:10 for brain); rat anti-Elav (Developmental Studies Hybridoma Bank, 1:25); mouse anti-Nc82 (Developmental Studies Hybridoma Bank, 1:2); mouse anti-Synaptotagmin (Developmental Studies Hybridoma Bank, 1:200). DAPI and Draq-5 (Biostatus) were used to mark nuclear DNA.

Microscopy: *Drosophila* eye images were photographed with a Leica DFC500 charge-coupled device digital camera. Fluorescent images were acquired with a Zeiss LSM510 confocal laser-scanning microscope and Olympus IX-51 fluorescent microscope. Postacquisition image processing was done using Adobe Photoshop.

RNA electrophoretic mobility shift assay: All RNA oligonucleotides were obtained from Dharmacon/Thermo Scientific, Inc. (Lafayette, CO). Assays were carried out as previously described (10).

Fluorescence *in situ* hybridization: Fluorescence *in situ* hybridization (FISH) for *Drosophila* tissues was performed as described previously (18). Co-staining for nuclear periphery was performed using rabbit anti-lamin 836 antibody (1:3000) (Paul A. Fisher) coupled with Cy5 anti-rabbit (Jackson ImmunoResearch). FISH for hippocampal neurons and mounted tissue sections was performed as described previously (19). To co-localize ZC3H14 with poly(A) RNA, rabbit anti-ZC3H14 (1:1000 for cell culture, 1:200 for tissue) (13) coupled with Cy2 anti-rabbit (Jackson ImmunoResearch) was used. DAPI was used to mark nuclear DNA.

Determination of bulk poly(A) tail length: Bulk poly(A) tails were analyzed using a standard assay (20) as described previously (14).

***Drosophila* stocks and genetics:** All crosses and stocks were maintained in standard conditions unless otherwise noted. The w^{1118}, *Df(21)crb-F89-4*, *Df(21)Exel6199*, *Df(21)Exel8178, P{EPgy2} EY08422* and *P{EP}CG5720[EP3716]* stocks were obtained from Bloomington Drosophila Stock Center (BDSC) at Indiana University. Excision alleles were generated by mobilization of *P{EPgy2}EY08422* using standard techniques. Mosaic wing discs were generated by crossing *UbxFLP;;ubi>GFP,FRT82B* and $dNab2^{ex3}$*,FRT82B/TM6B* animals. Genotypes used for eye modifier screen were: $Pabp2^{55}$ (22), $Pabp2^{EP2264}$ (22), abs^{00620}, abs^{f01698}, $bru\text{-}2^{EY18918}$, $bru\text{-}2^{MB00431}$, hrg^{10}, $Hrb87F^{KG02089}$, $Hrb98DE^{ZCL0558}$, $Nxt1^{f04855}$, $Nxt1^{DG05102}$, pum^{3}, pum^{bem}, sbr^{1}, sbr^{mgln}. *ram* alleles (23) were gift of Kevin Moses. sbr^{ts} allele (24) was a gift of Ilan Davis. *Pabp2* alleles were gift of Martine Simonelig. The RNAi line for *dNab2* was obtained from Vienna Drosophila Research Center.

***Drosophila* behavioral assays:** The negative-geotaxis assay was performed as previously described (25) with the following modifications: Briefly, newly eclosed flies (day 0) were collected daily, divided into groups of ten, and kept in separate vials for aging. Cohorts of age-matched flies were then transferred to a 25-mL graduated cylinder, gently tapped to the bottom, and analyzed for a climbing response. The number of flies reaching the top after 5, 10, 15, and 30 seconds post-tap was recorded. At least 10 groups were tested for each genotype and sex (where indicated). The flight assay was performed as previously described (26) with the following modifications: Briefly, newly eclosed flies (day 0) were collected daily, separated into groups of twenty, and aged as indicated. These groups

were funneled into a 500-mL graduated cylinder coated with paraffin oil and the number of flies landing within each of the ten 50-mL sections of the cylinder was recorded. At least 10 groups were tested for each genotype.

References

1. Ropers HH (2008) Genetics of intellectual disability. *Curr Opin Genet Dev* 18(3):241-250
2. Najmabadi H*, et al.* (2007) Homozygosity mapping in consanguineous families reveals extreme heterogeneity of non-syndromic autosomal recessive mental retardation and identifies 8 novel gene loci. *Hum Genet* 121(1):43-48
3. Kuss AW*, et al.* (2011) Autosomal recessive mental retardation: homozygosity mapping identifies 27 single linkage intervals, at least 14 novel loci and several mutation hotspots. *Hum Genet* 129(2):141-148
4. Kelly SM*, et al.* (2007) Recognition of polyadenosine RNA by zinc finger proteins. *Proc Natl Acad Sci U S A* 104(30):12306-12311
5. Leung SW*, et al.* (2009) Splice variants of the human ZC3H14 gene generate multiple isoforms of a zinc finger polyadenosine RNA binding protein. *Gene* 439(1-2):71-78
6. Hector RE*, et al.* (2002) Dual requirement for yeast hnRNP Nab2p in mRNA poly(A) tail length control and nuclear export. *Embo J* 21(7):1800-1810.
7. Anderson JT, Wilson SM, Datar KV, & Swanson MS (1993) NAB2: a yeast nuclear polyadenylated RNA-binding protein essential for cell viability. *Mol. Cell Biol.* 13:2730-2741
8. Suthers GK & Wilson SR (1990) Genetic counseling in rare syndromes: a resampling method for determining an approximate confidence interval for gene location with linkage data from a single pedigree. *Am J Hum Genet* 47(1):53-61

9. Durbin RM, *et al.* (2010) A map of human genome variation from population-scale sequencing. *Nature* 467(7319):1061-1073

10. Kelly SM, *et al.* (2010) Recognition of polyadenosine RNA by the zinc finger domain of nuclear poly(A) RNA-binding protein 2 (Nab2) is required for correct mRNA 3'-end formation. *J Biol Chem* 285(34):26022-26032

11. McNeil GP, Schroeder AJ, Roberts MA, & Jackson FR (2001) Genetic analysis of functional domains within the Drosophila LARK RNA-binding protein. *Genetics* 159(1):229-240

12. Zhang YQ, *et al.* (2001) Drosophila fragile X-related gene regulates the MAP1B homolog Futsch to control synaptic structure and function. *Cell* 107(5):591-603

13. Juge F, Zaessinger S, Temme C, Wahle E, & Simonelig M (2002) Control of poly(A) polymerase level is essential to cytoplasmic polyadenylation and early development in Drosophila. *EMBO J* 21(23):6603-6613

14. Kerwitz Y, *et al.* (2003) Stimulation of poly(A) polymerase through a direct interaction with the nuclear poly(A) binding protein allosterically regulated by RNA. *Embo J* 22(14):3705-3714

15. Murata T, *et al.* (2001) The hiiragi gene encodes a poly(A) polymerase, which controls the formation of the wing margin in Drosophila melanogaster. *Dev Biol* 233(1):137-147

16. Benoit B, *et al.* (2005) An essential cytoplasmic function for the nuclear poly(A) binding protein, PABP2, in poly(A) tail length control and early development in Drosophila. *Dev Cell* 9(4):511-522

17. Wilkie GS*, et al.* (2001) Small bristles, the Drosophila ortholog of NXF-1, is essential for mRNA export throughout development. *Rna* 7(12):1781-1792

18. van den Bogaart G, Meinema AC, Krasnikov V, Veenhoff LM, & Poolman B (2009) Nuclear transport factor directs localization of protein synthesis during mitosis. *Nat Cell Biol* 11(3):350-356

19. Lukong KE, Chang KW, Khandjian EW, & Richard S (2008) RNA-binding proteins in human genetic disease. *Trends Genet* 24(8):416-425

20. Apponi LH*, et al.* (2010) Loss of nuclear poly(A)-binding protein 1 causes defects in myogenesis and mRNA biogenesis. *Hum Mol Genet* 19(6):1058-1065

21. Mallon BS, Shick HE, Kidd GJ, & Macklin WB (2002) Proteolipid promoter activity distinguishes two populations of NG2-positive cells throughout neonatal cortical development. *J Neurosci* 22(3):876-885

Supplemental References (For experimental procedures)

1. Najmabadi H, *et al*. (2007) Homozygosity mapping in consanguineous families reveals extreme heterogeneity of non-syndromic autosomal recessive mental retardation and identifies 8 novel gene loci. *Hum Genet* 121(1):43-48
2. Abecasis GR, Cherny SS, Cookson WO, & Cardon LR (2001) GRR: graphical representation of relationship errors. *Bioinformatics* 17(8):742-743
3. O'Connell JR & Weeks DE (1998) PedCheck: a program for identification of genotype incompatibilities in linkage analysis. *Am J Hum Genet* 63(1):259-266
4. Abecasis GR, Cherny SS, Cookson WO, & Cardon LR (2002) Merlin--rapid analysis of dense genetic maps using sparse gene flow trees. *Nat Genet* 30(1):97-101
5. Ruschendorf F & Nurnberg P (2005) ALOHOMORA: a tool for linkage analysis using 10K SNP array data. *Bioinformatics* 21(9):2123-2125
6. Kruglyak L, Daly MJ, Reeve-Daly MP, & Lander ES (1996) Parametric and nonparametric linkage analysis: a unified multipoint approach. *Am J Hum Genet* 58(6):1347-1363
7. Gudbjartsson DF, Jonasson K, Frigge ML, & Kong A (2000) Allegro, a new computer program for multipoint linkage analysis. *Nat Genet* 25(1):12-13
8. Thiele H & Nurnberg P (2005) HaploPainter: a tool for drawing pedigrees with complex haplotypes. *Bioinformatics* 21(8):1730-1732
9. Brais B, *et al*. (1998) Short GCG expansions in the PABP2 gene cause oculopharyngeal muscular dystrophy. *Nat Genet* 18(2):164-167

10. Kelly SM, *et al*. (2007) Recognition of polyadenosine RNA by zinc finger proteins. *Proc Natl Acad Sci U S A* 104(30):12306-12311

11. Muddashetty RS, Kelic S, Gross C, Xu M, & Bassell GJ (2007) Dysregulated metabotropic glutamate receptor-dependent translation of AMPA receptor and postsynaptic density-95 mRNAs at synapses in a mouse model of fragile X syndrome. *J Neurosci* 27(20):5338-5348

12. Antar LN, Afroz R, Dictenberg JB, Carroll RC, & Bassell GJ (2004) Metabotropic glutamate receptor activation regulates fragile x mental retardation protein and FMR1 mRNA localization differentially in dendrites and at synapses. *J Neurosci* 24(11):2648-2655

13. Leung SW, *et al*. (2009) Splice variants of the human ZC3H14 gene generate multiple isoforms of a zinc finger polyadenosine RNA binding protein. *Gene* 439(1-2):71-78

14. Apponi LH, *et al*. (2010) Loss of nuclear poly(A)-binding protein 1 causes defects in myogenesis and mRNA biogenesis. *Hum Mol Genet* 19(6):1058-1065

15. Christie SB, Akins MR, Schwob JE, & Fallon JR (2009) The FXG: a presynaptic fragile X granule expressed in a subset of developing brain circuits. *J Neurosci* 29(5):1514-1524

16. Moberg KH, Mukherjee A, Veraksa A, Artavanis-Tsakonas S, & Hariharan IK (2004) The Drosophila F box protein archipelago regulates dMyc protein levels in vivo. *Curr Biol* 14(11):965-974

17. Zhang YQ, *et al*. (2001) Drosophila fragile X-related gene regulates the MAP1B homolog Futsch to control synaptic structure and function. *Cell* 107(5):591-603

18. Merabet S, Ebner A, & Affolter M (2005) The Drosophila Extradenticle and Homothorax selector proteins control branchless/FGF expression in mesodermal bridge-cells. *EMBO Rep* 6(8):762-768

19. Bassell GJ, Singer RH, & Kosik KS (1994) Association of poly(A) mRNA with microtubules in cultured neurons. *Neuron* 12(3):571-582

20. Chekanova JA, Shaw RJ, & Belostotsky DA (2001) Analysis of an essential requirement for the poly(A) binding protein function using cross-species complementation. *Curr Biol* 11(15):1207-1214

21. Longtine MS, *et al*. (1998) Additional modules for versatile and economical PCR-based gene deletion and modification in Saccharomyces cerevisiae. *Yeast* 14(10):953-961.

22. Benoit B, *et al*. (2005) An essential cytoplasmic function for the nuclear poly(A) binding protein, PABP2, in poly(A) tail length control and early development in Drosophila. *Dev Cell* 9(4):511-522

23. D'Costa A, Reifegerste R, Sierra S, & Moses K (2006) The Drosophila ramshackle gene encodes a chromatin-associated protein required for cell morphology in the developing eye. *Mech Dev* 123(8):591-604

24. Wilkie GS, *et al*. (2001) Small bristles, the Drosophila ortholog of NXF-1, is essential for mRNA export throughout development. *Rna* 7(12):1781-1792

25. Ganetzky B & Flanagan JR (1978) On the relationship between senescence and age-related changes in two wild-type strains of Drosophila melanogaster. *Exp Gerontol* 13(3-4):189-196

26. Benzer S (1973) Genetic dissection of behavior. *Sci Am* 229(6):24-37

CHAPTER 3

Defining the role of *dNab2* in the *Drosophila* mushroom bodies

ChangHui Pak, John J. Noto, Daniel Marenda, Anita H. Corbett, Kenneth H. Moberg

CP performed all fly experiments, immunostaining and qRT-PCR. Fly locomotor behavior tests were performed in collaboration with John J. Noto in the Department of Biochemistry at Emory University (Atlanta, GA). Courtship suppression assay was done in the laboratory of Dr. Daniel Marenda, Department of Neurobiology, Drexel University (Philadelphia, PA).

Introduction

Understanding normal brain function is critical for identifying the molecular etiology behind neurological disorders and therapeutic treatments. One way to understand brain function is to identify and characterize genes that when mutated impair normal human intellectual development. Intellectual disability (ID), previously referred to as mental retardation, is characterized by limited intellectual capacities reflected by an IQ below 70 and major constraints in adaptive behavior (1). Therapeutic options for the treatment of ID are extremely limited and its comparatively high prevalence of about 2% renders this disorder a major socioeconomic burden (1).

The fruit fly *Drosophila melanogaster* has served as an excellent system in which to model and dissect conserved aspects of human disease biology. In particular, fly models of ID have shed much light on our current understanding of the genetic factors involved in learning and cognition. For example, fly models of Fragile X syndrome (FXS), the most commonly inherited X-linked ID affecting about 1 in 4000 males and 1 in 6000 females, have been extensively studied and provided clues as to how the *Fmr1* disease gene controls neuronal cell biology and brain function (2).

Recently, we identified a locus for unspecific or non-syndromic autosomal recessive ID (NS-ARID) on chromosome 14q31.3 corresponding to mutation of the *ZC3H14* gene in two independent Iranian families (Chapter 2). *ZC3H14* encodes an evolutionarily conserved Cys_3His tandem zinc finger polyadenosine RNA-binding protein (3, 4). The founding member of this protein family, *S. cerevisiae* Nab2, is essential for viability and required for proper 3'end formation and poly(A) RNA export from the nucleus (5-7). Although multiple tissue-specific splice variants of human

ZC3H14 have been described (4), their function in multicellular organisms has not been examined.

We therefore took advantage of *Drosophila* as a model system to examine how mutations in the *ZC3H14* gene might alter brain development or function, and to identify molecular roles for the *Drosophila* ZC3H14 orthologue in neurons. We have previously described a *Drosophila* orthologue of ZC3H14 termed dNab2 that is required specifically in neurons for normal development and behavior (Chapter 2). Pan-neuronal knockdown of *dNab2* results in locomotor and flight defects similar to genomic null allele $dNab2^{ex3}$ and reciprocally, pan-neuronal expression of wild type *dNab2* in animals otherwise null for *dNab2* completely restores locomotor defects and rescues pupal lethality.

Based on these observations, we hypothesized that the role of *dNab2* in controlling normal behavior could be further refined to specific regions of the fly brain, which would then allow us to identify a more manageable region to study the molecular mechanisms of *dNab2* in neurons. In support of this hypothesis, we have accumulated a significant body of experimental evidence showing that *dNab2* loss results in defects in the structure of the mushroom bodies (MBs), (S.M. Kelly, unpublished data), which are highly specialized axonal structures in the CNS that are required for learning and memory (8). Given that the invertebrate MBs and the vertebrate hippocampus are both implicated in cognition, and thus show some degree of functional similarity, we chose the MBs as a system to examine how *dNab2* controls normal behavior.

Here we show that dNab2 protein is expressed in the cell bodies of MB neurons and is required specifically in the MBs to promote normal behavior including locomotion and memory formation. Flies depleted of dNab2 in the MBs show poor locomotor

activity that progressively worsens with age. Moreover, these flies also exhibit defective short-term memory in a courtship suppression assay (9, 10). Finally, we identify *mushroom-body expressed* (*mub*), which encodes a KH-type poly(C) RNA-binding protein, as a potential target of dNab2 whereby *mub* mRNA transcript levels are altered in *dNab2* mutant heads. Taken together, these findings reveal an MB-specific requirement of dNab2 and for the first time, provide potential insight into cognitive deficits that are present in patients with ZC3H14-associated NS-ARID.

Results

dNab2 is expressed in the MB neurons

The mushroom body (MB) consists of ~2500 Kenyon cells (the intrinsic neurons of the MB) whose axons project anteriorly from cell bodies located in the posterior dorsal region of the fly brain (8, 11). Near the site of synapse, axons extending from properly developed Kenyon cells bifurcate into vertical (α/α') and horizontal (β, β', and γ) lobes (8, 11). Interestingly, the MB receives inputs from a diverse set of interneurons, including projection neurons from the olfactory lobes, and has been extensively implicated in both short- and long-term memory (8, 11). Loss of *dNab2* perturbs normal architecture of the mushroom body (MB) neurons, including β-lobe fusions and missing α-lobes (S.M. Kelly, unpublished data). In light of these defects in the CNS, we sought to investigate the role of *dNab2* in the MB neurons.

To test whether dNab2 protein is actually expressed in Kenyon cells of the MBs, we immunostained for dNab2 protein in adult brains expressing an MB-specific *201Y-Gal4*-driven nuclear LacZ, which can be detected by co-staining with anti-β-galactosidase antibody (Figure 3.1). This analysis shows that dNab2 protein is widely expressed within the β-gal positive cell bodies of Kenyon neurons in the adult brain.

dNab2 acts in the MB neurons to promote normal locomotor activity

Flies with pan-neuronal (*Elav-Gal4*) depletion of dNab2 show defects in locomotor activity that are very similar to locomotor defects documented in *dNab2* genomic null animals with the negative geotaxis assay (Chapter 2). Moreover, when we tested flies of each of these genotypes at 2, 5 or 10 days post-eclosion, we observed an

age-dependent decline in locomotor activity in *dNab2* genomic mutants as well as in pan-neuronal *dNab2* knockdown flies, suggesting that reduced *dNab2* expression in the nervous system leads to a physiologically significant age-related decline in neuronal function (Chapter 2). We also tested for locomotor defects in flies that were knocked down for *dNab2* specifically in the MBs by using four independent MB *Gal4* lines: *201Y-Gal4*, *1471-Gal4*, *7B-Gal4* and *OK107-Gal4* (12, 13). The partially overlapping expression patterns of these drivers permit us to deplete dNab2 from some or all of MB neurons: *201Y-Gal4* is expressed in γ and α/β neurons (12). *1471-Gal4* is preferentially expressed in the γ lobes (12). *7B-Gal4* is expressed strongly in α/β lobes (13). *OK107-Gal4* is expressed in all the lobes of the MB, including α/β, α'/β' and γ lobes (12). In four independent *dNab2*-knockdown experiments using these MB *Gal4* drivers, we observed moderate decreases in adult locomotor activity compared to control flies (Figure 3.2), suggesting that *dNab2* is required in the MB cells to promote normal behavior. Moreover, this MB-specific locomotor defect worsened with age (e.g. at Day 10 post-eclosion) in a manner similar to $dNab2^{ex3}$ null flies or those with pan-neuronal depletion of dNab2 (data not shown).

Although the MB *Gal4* lines used were specific to the MB cells, they do exhibit a low level of expression of these *Gal4* lines elsewhere outside the MBs (12). To control for the possibility that the locomotor defects seen in the 'MB' *dNab2* knockdown flies might be due to non-specific effects of dNab2 depletion in cells other than Kenyon cells, we utilized a transgene that expresses the *Gal4* inhibitor *Gal80* specifically in the MB αβ, α'β', and γ neurons [*MBGal80;*(14);(15)]. By combining the *MBGal80* transgene with *Elav-Gal4*-driven *dNab2* RNAi (pan-neuronal knockdown), we can selectively

block RNAi of *dNab2* only in the MB cells, thereby allowing *dNab2* knockdown in all neurons except the MB neurons. As shown in Figure 3.3, flies with dNab2 depletion in all neurons (*Elav-Gal4/+;dNab2-IR/+*) show a severe decrease in locomotor activity compared to control (*Elav-Gal4*). However, when *dNab2* was knocked down in all neurons except the MBs (*Elav-Gal4/+;dNab2-IR/MBGal80*), adult flies show improved climbing performance relative to pan-neuronal *dNab2* knockdown flies (*Elav-Gal4/+;dNab2-IR/+*), indicating that *dNab2* is required in the MB neurons for optimal locomotor behavior (Figure 3.3).

dNab2 is necessary but not sufficient to promote locomotor activity in the MBs

While MB-specific knockdown of *dNab2* expression caused a significant decrease in locomotor activity, we hypothesized that re-introducing *dNab2* in only the MBs might rescue the locomotor and eclosion defects seen *dNab2* null flies. In order to test this hypothesis, we generated flies harboring a *UAS-dNab2-Flag* expression construct and tested for rescue of the eclosion and locomotor defects observed in *dNab2* null flies (Chapter 2). To ensure that expression of the *dNab2* transgene did not adversely affect wild type flies, we first examined whether overexpressing wild type *dNab2* (using the P-insertion *EP3716*) in the MBs resulted in any observable phenotype. As shown in Figure 3.4, wild type flies overexpressing *dNab2* in the MBs do not display any gross abnormalities or locomotor defects as assayed by the negative geotaxis assay. By contrast, *201Y-Gal4*-driven re-expression of dNab2 in the MB cells of *dNab2* null flies increased survival to adulthood such that 66% of flies eclosed compared to rescue control (39%), which expresses a leaky wild type *dNab2* transgene in the absence of a *Gal4*

driver (Table 3.1). Notably, expression of *dNab2* from another MB driver *1471-Gal4* did not significantly alter the survival to adulthood of *dNab2* null flies compared to rescue control (Table 3.1). This difference in phenotypic rescue between the *201Y-Gal4* and *1471-Gal4* drivers could be explained by differential expression patterns in the MB neurons as *1471-Gal4* is primarily expressed in the γ lobes whereas *201Y-Gal4* expression extends to γ and α/β lobes (12).

Given that expression of dNab2 in the MBs significantly rescued the eclosion defect seen in homozyogous dNab2 null flies, we determined whether re-expressing dNab2 only in the MBs is sufficient to rescue the locomotor defect of *dNab2* mutants. To test this idea, we used negative geotaxis assay to measure climbing ability in flies that express wild type *dNab2* transgene under the control of the following MB drivers: *1471-Gal4*, *201Y-Gal4* and *OK107-Gal4*. At both Day 2 and Day 5, the *1471-Gal4* driver was unable to rescue the locomotor defect of *dNab2* null flies, and flies expressing *UAS-dNab2-Flag* from the *201Y-Gal4* and *OK107-Gal4* drivers actually performed worse than *UAS-dNab2-Flag* control flies (Figure 3.5). Therefore, although dNab2 is required in the MBs for normal locomotor behavior, adding it back only in the MB neurons of otherwise *dNab2* mutant flies is not sufficient to restore normal behavior. This finding contrasts with the locomotor rescue produced by pan-neuronal re-expression of *dNab2* (Chapter 2), suggesting that dNab2 acts at multiple points within multicellular circuits that control behavior and that adding it back to a single cell type is not sufficient to restore overall function of the circuit.

Depletion of dNab2 in the MBs alters memory formation

Based on the previous finding that mutations in *ZC3H14* lead to NS-ARID in humans (Chapter 2), we hypothesized that dNab2 might function in the basis of learning and memory. Since MBs are the center for higher order brain function involved in memory consolidation and memory recall (8, 11), we subjected flies that are knocked down for *dNab2* in the CNS to a 'courtship suppression' learning paradigm. Courtship behaviors in *Drosophila* are innate and involve a complex set of behaviors that can be modified through training (16). Conditioned courtship suppression is one method to modify these behaviors and has been used as an associative learning paradigm (9, 10, 16). Briefly, in the initial training phase, a male is placed into a small mating chamber with a female that is already mated and thus unreceptive to courtship advances. While he initially courts the female vigorously, he learns that she is unreceptive and his courtship activity declines over time (the 'learning' component). In the next phase (the 'memory' component), this 'trained' male is placed into a mating chamber with a virgin female that is receptive to courtship advances. Despite her receptiveness, a pre-trained wild type male will show a suppressed courtship index toward this virgin female because of his 'learned' experience with the previous unreceptive female. This type of behavioral learning (termed 'courtship suppression') usually lasts 2-3 hrs and thus provides a useful test of short-term memory (9, 10).

Initially, we observed that males with pan-neuronal *dNab2* knockdown (*Elav-Gal4*) did not even court virgin females (personal communication, D. Marenda), suggesting perhaps a defect in recognition of female olfactory and gustatory cues. Consequently, we knocked down *dNab2* specifically in the MBs using *OK107-Gal4* and tested the ability of these males to court. In the initial training phase, both control and

dNab2 MB knockdown flies behaved similarly, suggesting that the learning component in these flies is intact (Figure 3.6). However, when tested for their short-term memory, *dNab2* MB knockdown males showed a profound defect in remembering the suppressed courtship behavior they had just learned (Figure 3.6). Thus, *dNab2* is required in the MB neurons to promote memory formation and for the first time model cognitive deficits present in patients with *ZC3H14*-mediated NS-ARID.

mub is a potential target of dNab2

In order to determine the molecular mechanism by which *dNab2* contributes to normal MB structure and promotes normal behavior and memory recall, we searched for possible *dNab2* target(s). We had previously found that a loss-of-function allele of the gene *mushroom-body expressed* (*mub*) acts as a dominant suppressor of a rough-eye phenotype produced by overexpression of *dNab2* in the developing eye (Chapter 2). *mub* encodes a KH-type poly(C) RNA-binding protein (Mub) that is implicated in regulation of alternative nuclear splicing of nuclear mRNAs (17) and *mub* mutant flies display similar phenotypes to *dNab2* mutants, including defective eclosion, flightless behavior and defective locomotor activity (18). Based on these similar phenotypes and the strong genetic interaction between *mub* and *dNab2* in the adult eye, we hypothesized that dNab2 might control *mub* expression. To test this, we determined whether the levels of *mub* mRNA transcripts are altered in *dNab2* mutant heads relative to control heads. By qRT-PCR analysis, we observed an approximately 50% reduction in *mub* mRNA transcripts in *dNab2* mutant flies (Figure 3.7), suggesting that *dNab2* might be required for efficient expression, processing or stability of *mub* mRNA transcripts. This finding along with

genetic data suggests that *mub* might be a downstream target of dNab2 to mediate memory formation in the MB neurons.

Discussion

Our recent discovery that mutations in *ZC3H14/dNab2* lead to non-syndromic autosomal recessive intellectual disability (NS-ARID) in humans (Chapter 2) prompted us to further refine the neuronal requirement of dNab2 in the *Drosophila* CNS. We find that dNab2 is required in the mushroom body (MB), a highly specialized structure involved in learning and memory (8, 11), for efficient short-term memory, and that complete loss of dNab2 leads to defects in the structure of the MBs. We show that dNab2 protein is expressed in the cell bodies of MB neurons and that flies depleted of dNab2 in the MBs show poor locomotor activity that progressively worsens with age. Moreover, these flies also exhibit defects in short-term memory without any learning deficit tested by conditioned courtship suppression assay. Finally, we identify *mub* as a potential target of dNab2 whereby *mub* mRNA transcript levels are altered in *dNab2* mutant heads.

In contrast to the requirement of dNab2 in controlling locomotion in the MBs, we unexpectedly found that re-expression of dNab2 in animals null for *dNab2* is not sufficient to rescue locomotor defects. This finding suggests that dNab2 might act at multiple points within multicellular circuits that control behavior and that adding it back to a single cell type is not sufficient to restore overall function of the circuit. Although complete ablation of the MB cells leads to enhanced walking activity (19, 20), the precise cellular circuits involved in MB-regulated locomotion are unclear. Future studies examining the neuronal circuitry of the MBs will allow us to address how dNab2 acts within this specialized circuit to promote normal locomotion.

Using the courtship suppression learning paradigm, we observed a short-term memory deficit in flies that are knocked down for dNab2 specifically in the MBs. This

finding is significant as it appears to parallel cognitive deficits that are present in patients with ZC3H14-mediated NS-ARID. Having established this link between dNab2 and memory formation, it would be interesting to examine whether the memory defect seen in *dNab2* knockdown flies is due to defective MB development, or to defects in synaptic signaling within MB neurons, or both. Since our studies utilized knockdown of *dNab2* via *OK107-Gal4*, which expresses from early development and throughout adult (12), temporal control of RNAi via temperature-sensitive *Gal80* repressor (21, 22) will allow one to dissect the critical time points dNab2 is needed for either development of MB neuronal differentiation or maintenance of MB neuronal function, e.g. learning and memory, or both.

The identification of *mub* as a potential mRNA target for dNab2 may provide a molecular link between dNab2 and MB function. A *mub* allele scored as a strong genetic modifier of dNab2-mediated overexpression phenotype in the adult eye (Chapter 2) and shows decrease in mRNA transcript level in *dNab2* mutant heads. Mub is a KH-type poly(C) RNA-binding protein that is strongly expressed in the MBs and has been implicated in regulation of alternative splicing of nuclear mRNAs (17, 23). Mub has also been identified as a suppressor of neurodegeneration induced by ataxin-1 (24). Since many genes that control learning and memory show localized expression in the MBs (25, 26), Mub could potentially be involved in mediating learning and memory by acting downstream of dNab2. One remaining question then is how dNab2 regulates *mub* expression post-transcriptionally. Future studies should include examining transcript-specific poly(A) tail length (LM-PAT) (27) and testing whether dNab2 physically interacts with *mub* transcripts by RNA-immunoprecipitation (28). In addition, whether

the misregulation of *mub* levels is the underlying cause for memory defects seen in *dNab2* MB-knockdown flies should be addressed.

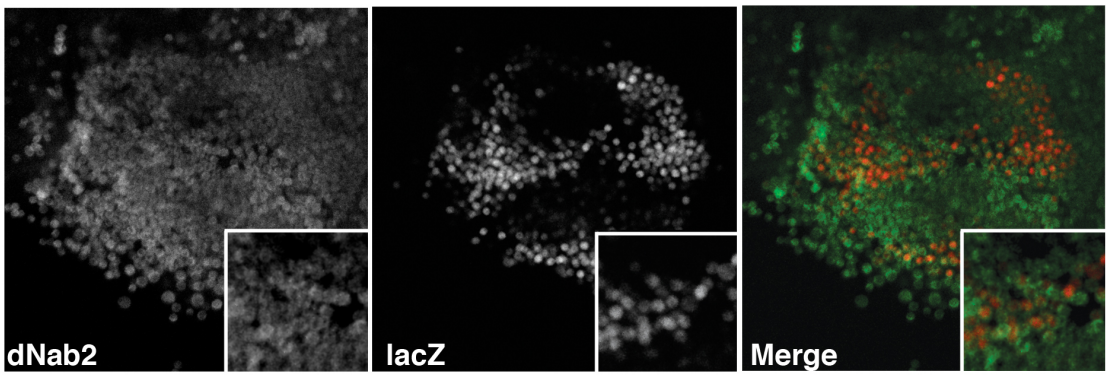


Figure 3.1. Localization of dNab2 protein in MB neurons of the adult fly brain. *201Y-Gal4*-driven expression of MB-specific nuclear lacZ is shown in red. Anti-dNab2 staining is shown in green.

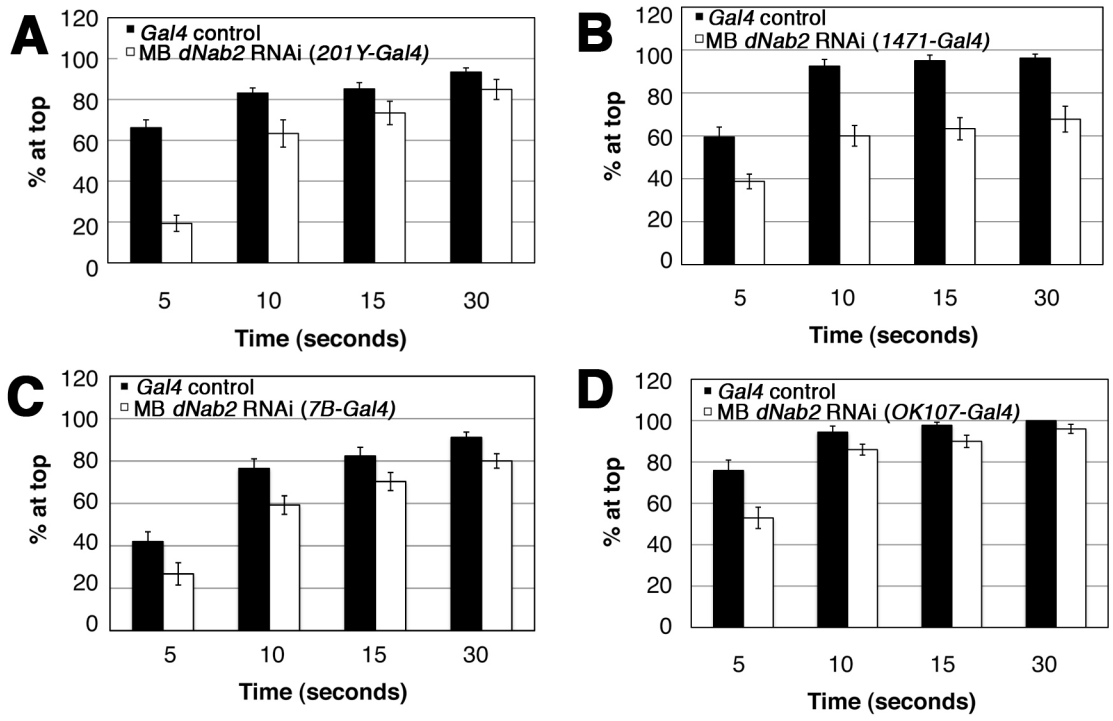


Figure 3.2. *dNab2* is required in the MB neurons for normal locomotor activity. Results of the negative geotaxis assay for four different MB *dNab2* RNAi flies (*201Y-Gal4>IR*, *1471-Gal4>IR*, *7B-Gal4>IR*, *OK107-Gal4>IR*) with respect to their *Gal4* controls (*201Y-Gal4>+*, *1471-Gal4>+*, *7B-Gal4>+*, *OK107-Gal4>+*) at Day 10 are shown. Data are presented as the average percentage of flies reaching the top of a 25-mL graduated cylinder after 5, 10, 15 and 30 seconds across all trials. Groups of ten flies were tested in at least ten independent trials for each genotype. $p<0.05$ (two-tailed *t*-test) when comparing *Gal4* control with MB *dNab2* RNAi at 5 seconds across all genotypes. Error bars=s.e.m.

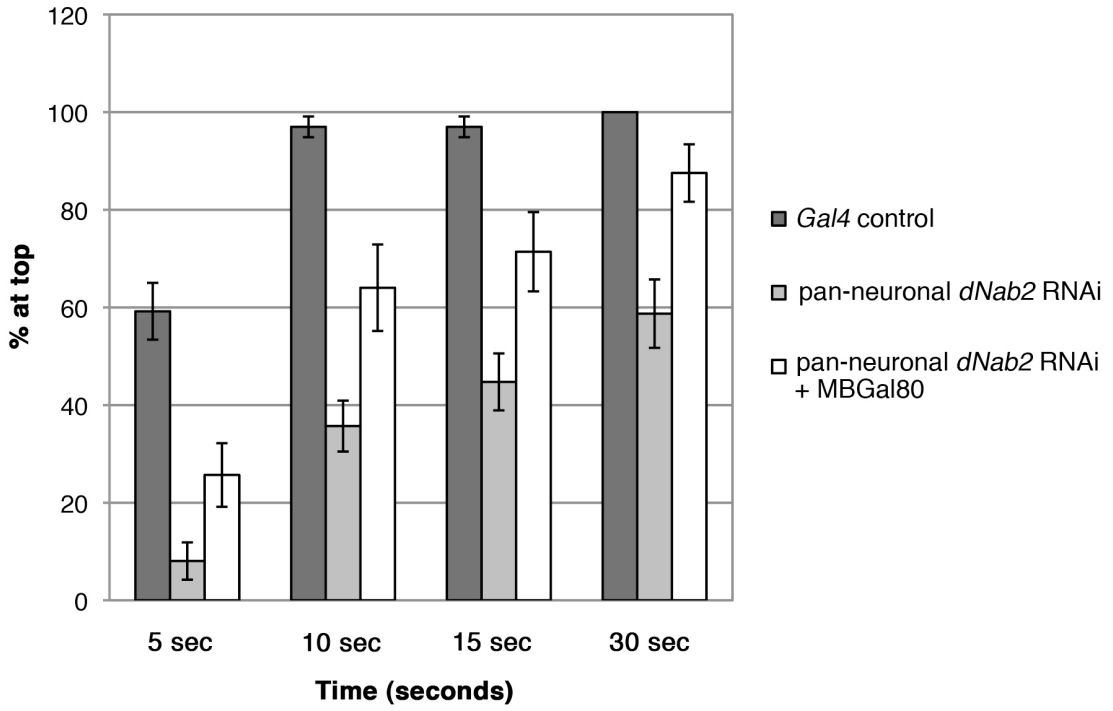


Figure 3.3. Specificity of *dNab2*-mediated behavior in the MBs. Results of the negative geotaxis assay for *Gal4* control (*Elav-Gal4*), pan-neuronal *dNab2* RNAi (*Elav-Gal/+;dNab2-IR>+*) and pan-neuronal *dNab2* RNAi with mushroom body-specific *Gal80* (*Elav-Gal4/+;dNab2-IR/MBGal80*) at mixed ages are shown. Data are presented as the average percentage of flies reaching the top of a 25-mL graduated cylinder after 5, 10, 15 and 30 seconds across all trials. Groups of ten flies were tested in at least ten independent trials for each genotype. $p<0.05$ (two-tailed *t*-test) when comparing *Gal4* control to pan-neuronal *dNab2* RNAi and pan-neuronal *dNab2* RNAi to pan-neuronal *dNab2* RNAi + MBGal80. Error bars=s.e.m.

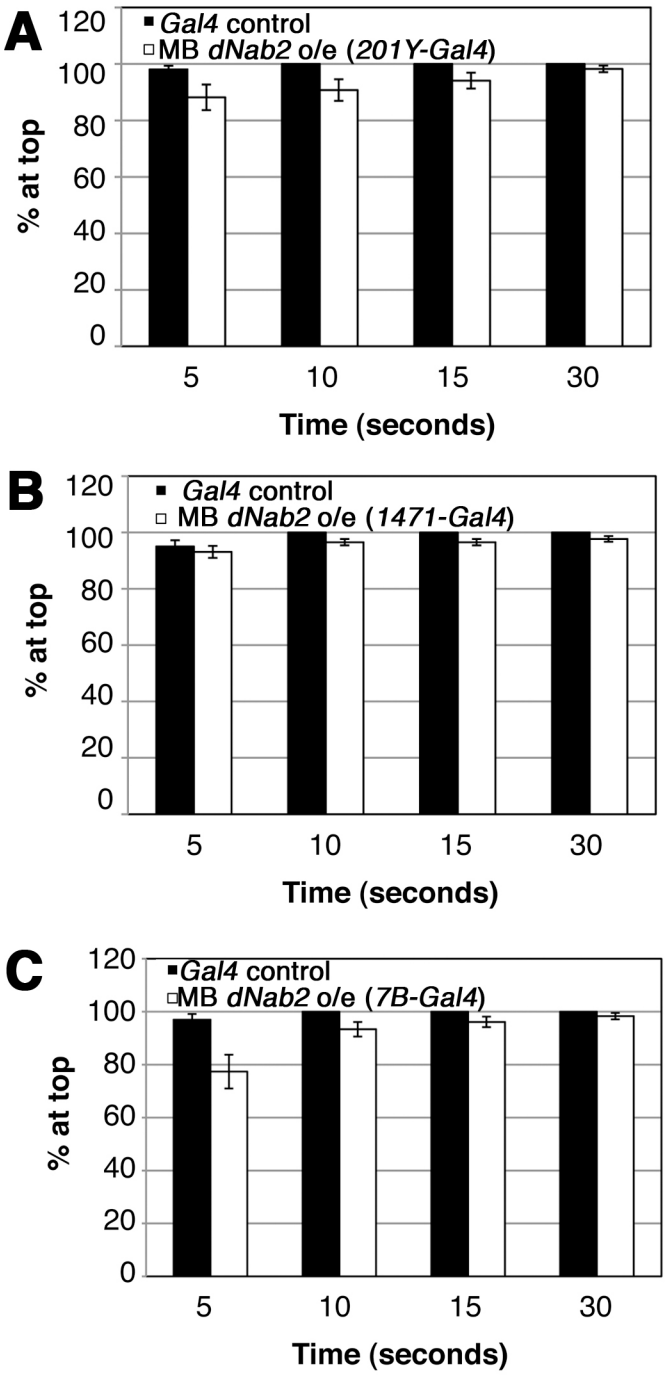


Figure 3.4. MB-specific overexpression of *dNab2* in wild type flies does not alter locomotor/climbing behavior. A negative geotaxis assay measuring climbing ability was performed on control (*Gal4* driver alone in wild type background) and mushroom body-specific *dNab2* overexpression flies (*Gal4* driving $dNab2^{EP3716}$) at Day 2 using *201Y-Gal4*

(**A**), *1471-Gal4* (**B**) and *7B-Gal4* (**C**). Data are presented as the average percentage of flies reaching the top of a 25-mL graduated cylinder after 5, 10, 15 and 30 seconds across all trials. Groups of ten flies were tested in at least ten independent trials for indicated each genotype. Error bars=s.e.m.

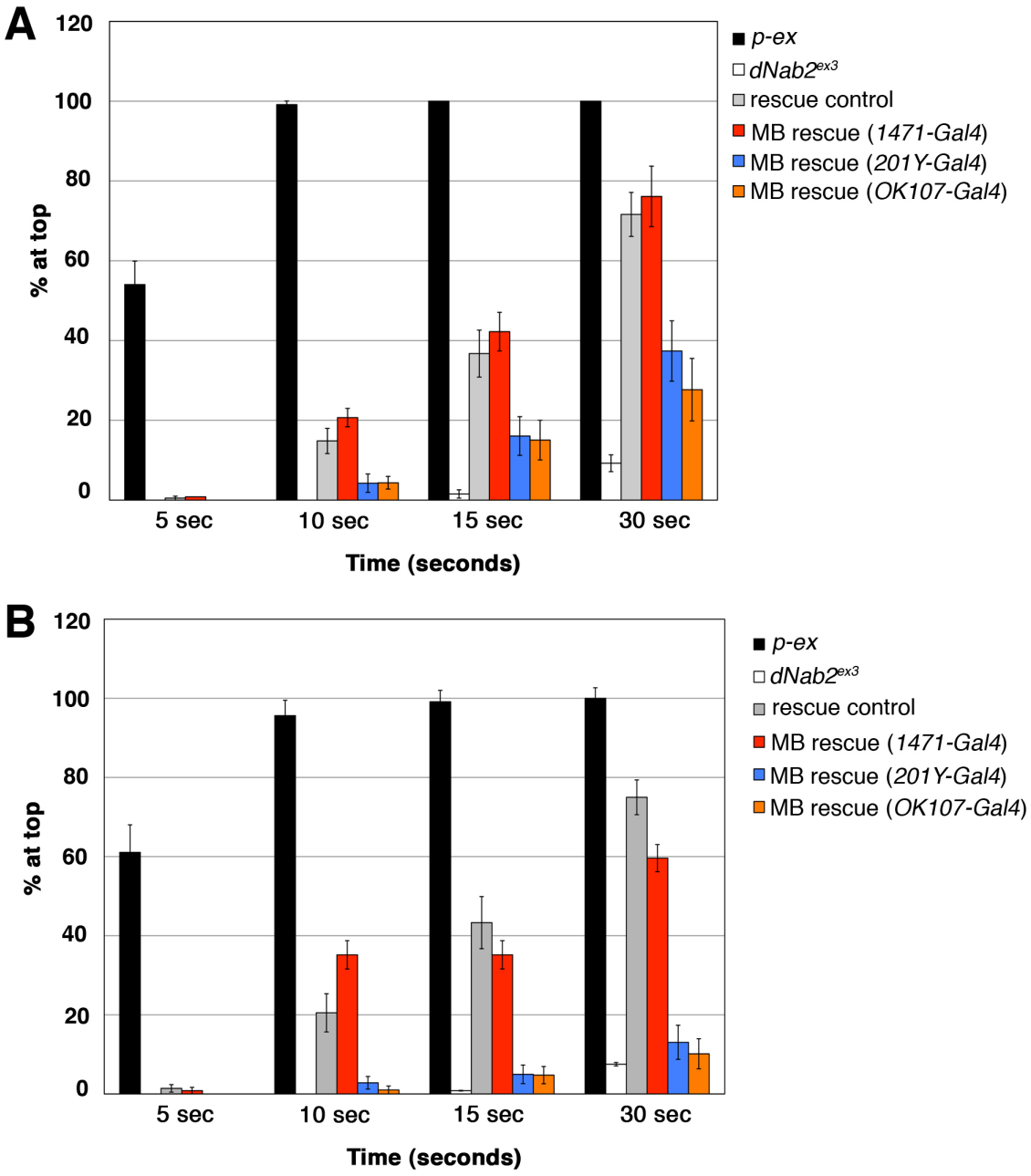


Figure 3.5. Re-expression of wild type dNab2 only in the MB neurons is not sufficient to rescue locomotor defect observed in *dNab2*^{ex3} zygotic mutants. Results of the negative geotaxis assay performed on the indicated genotypes highlighted by colored bars at Day 2 (**A**) and Day 5 (**B**). Data are presented as the average percentage of flies reaching the top of a 25-mL graduated cylinder after 5, 10, 15 and 30 seconds across all trials. Groups of

ten flies were tested in at least ten independent trials for indicated each genotype. Error bars=s.e.m.

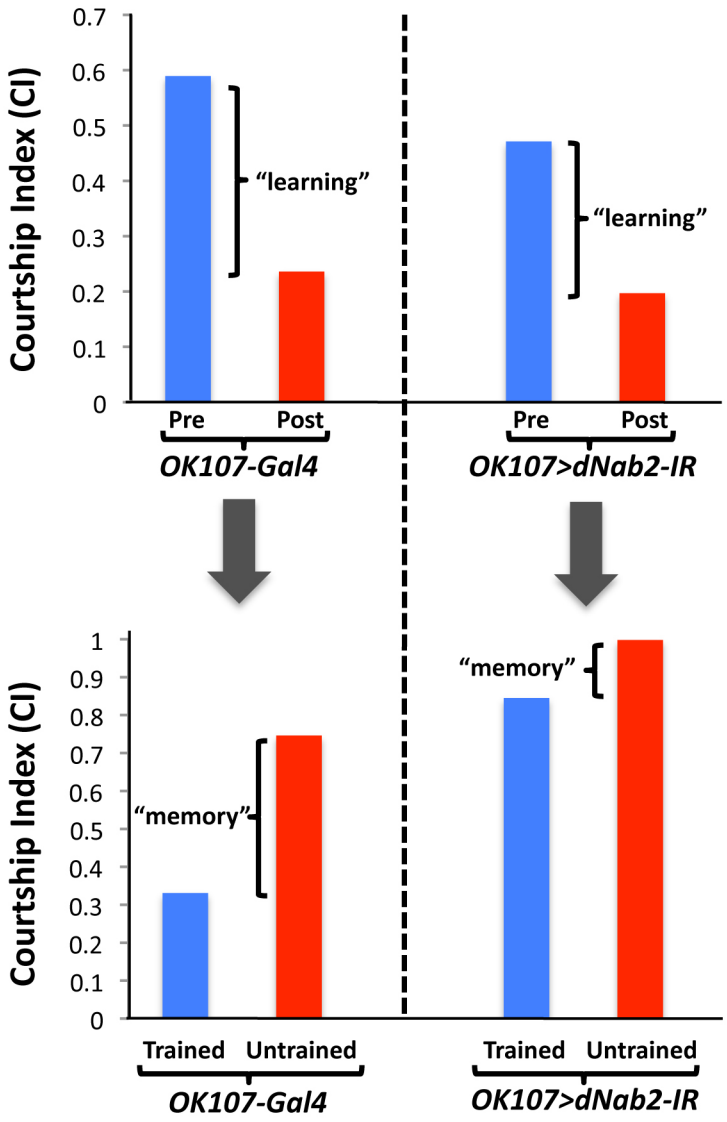


Figure 3.6. *dNab2* is required in the MB neurons for memory recall. Courtship conditioning assay of learning and memory in age-matched control (*OK107-Gal4*) and *dNab2* MB-depleted (*OK107>dNab2-IR*) male flies. N=3 Courtship index is determined by the fraction of time a male spends in courtship activity during the observation period (10). In the learning phase, 'pre' refers to mating behavior in the initial observation period and 'post' refers to mating behavior observed in later time point.

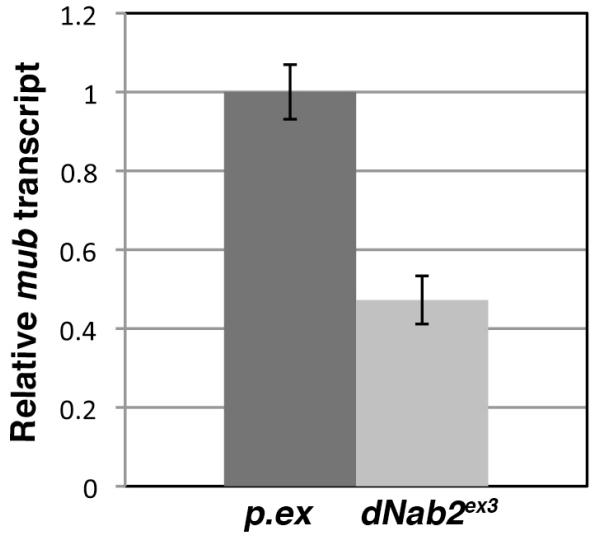


Figure 3.7. *dNab2* regulates *mub* transcript levels. Quantitative real-time RT-PCR analysis of *mub* transcript levels in adult heads of Day 5 old flies. All genotypes are normalized to *dNab2* transcript levels in *p.ex* control animals (set to 1.0). *β-tubulin* is an internal control. p=0.0004 (two-tailed *t*-test). Error bars=s.d. Representative of a single experiment.

Genotype	% eclosed (of expected)
dNab2^{ex3}/dNab2^{ex3} (zygotic mutant)	~3%
UAS-dNab2-Flag;; dNab2^{ex3}/dNab2^{ex3} (rescue control)	39%
201Y>UAS-dNab2-Flag;; dNab2^{ex3}/dNab2^{ex3} (MB rescue)	66%
1471>UAS-dNab2-Flag;; dNab2^{ex3}/dNab2^{ex3} (MB rescue)	36%

Table 3.1. Summary of the percentage of flies eclosed (of expected) for indicated genotypes.

Experimental procedures

***Drosophila* stocks and genetics:** All crosses and stocks were maintained in standard conditions unless otherwise noted. Precise and imprecise excision alleles (*p.ex* and *dNab2*^{ex3}) and transgenic rescue flies (*UAS-dNab2-Flag*) were obtained from previous studies (Chapter 2). *P{EP}CG5720[EP3716]* and stocks for genetic screen were obtained from Bloomington Drosophila Stock Center (BDSC) at Indiana University. *Pabp2* alleles were gift of Martine Simonelig. *MBGal80* was a gift of Scott Waddell. The RNAi line for *dNab2* was obtained from Vienna Drosophila Research Center.

Immunostaining: Immunostaining was performed as described previously for *Drosophila* tissues (29). Antibodies used were rabbit anti-dNab2 (1:10 for brain) (Chapter 2) and mouse anti-beta-galactosidase (1:1000).

Microscopy: Fluorescent images were acquired with a Zeiss LSM510 confocal laser-scanning microscope and Olympus IX-51 fluorescent microscope. Postacquisition image processing was done using Adobe Photoshop.

Quantitative real-time RT-PCR (qRT-PCR): Total RNA from *Drosophila* samples was extracted using TRIzol reagent (Invitrogen). cDNAs were generated by using the SuperScript III Reverse Transcriptase (Invitrogen). Quantitative Real-time PCR was performed according to standard protocols with LightCycler 480 (30). Primer sequences for *mub*: gccctggcccaatacttaat (left), ctgggcgcaatgtagtgata (right).

***Drosophila* behavioral assays:** The negative-geotaxis assay was performed as previously described (31) with the following modifications: Briefly, newly eclosed flies (day 0) were collected daily, divided into groups of ten, and kept in separate vials for aging. Cohorts of age-matched flies were then transferred to a 25-mL graduated cylinder, gently tapped to

the bottom, and analyzed for a climbing response. The number of flies reaching the top after 5, 10, 15, and 30 seconds post-tap was recorded. At least 10 groups were tested for each genotype and sex (where indicated). Courtship suppression assay was done as previously described (9, 10).

Reference

1. Ropers HH (2010) Genetics of early onset cognitive impairment. *Annu Rev Genomics Hum Genet* 11:161-187.
2. Bhogal B & Jongens TA (2010) Fragile X syndrome and model organisms: identifying potential routes of therapeutic intervention. *Dis Model Mech* 3(11-12):693-700.
3. Kelly SM, *et al*. (2007) Recognition of polyadenosine RNA by zinc finger proteins. *Proc Natl Acad Sci U S A* 104(30):12306-12311.
4. Leung SW, *et al*. (2009) Splice variants of the human ZC3H14 gene generate multiple isoforms of a zinc finger polyadenosine RNA binding protein. *Gene* 439(1-2):71-78.
5. Anderson JT, Wilson SM, Datar KV, & Swanson MS (1993) NAB2: a yeast nuclear polyadenylated RNA-binding protein essential for cell viability. *Mol Cell Biol* 13(5):2730-2741.
6. Green DM, *et al*. (2002) Nab2p is required for poly(A) RNA export in Saccharomyces cerevisiae and is regulated by arginine methylation via Hmt1p. *J Biol Chem* 277(10):7752-7760.
7. Hector RE, *et al*. (2002) Dual requirement for yeast hnRNP Nab2p in mRNA poly(A) tail length control and nuclear export. *EMBO J* 21(7):1800-1810.
8. Heisenberg M (2003) Mushroom body memoir: from maps to models. *Nat Rev Neurosci* 4(4):266-275.
9. Ejima A & Griffith LC (2011) Assay for courtship suppression in Drosophila. *Cold Spring Harb Protoc* 2011:pdb prot5575.

10. Ejima A & Griffith LC (2007) Measurement of Courtship Behavior in Drosophila melanogaster. *CSH Protoc* 2007:pdb prot4847.

11. Zars T (2000) Behavioral functions of the insect mushroom bodies. *Curr Opin Neurobiol* 10(6):790-795.

12. Aso Y, *et al*. (2009) The mushroom body of adult Drosophila characterized by GAL4 drivers. *J Neurogenet* 23(1-2):156-172.

13. Nicolai M, Lasbleiz C, & Dura JM (2003) Gain-of-function screen identifies a role of the Src64 oncogene in Drosophila mushroom body development. *J Neurobiol* 57(3):291-302.

14. Lee T & Luo L (1999) Mosaic analysis with a repressible cell marker for studies of gene function in neuronal morphogenesis. *Neuron* 22(3):451-461.

15. Krashes MJ, Keene AC, Leung B, Armstrong JD, & Waddell S (2007) Sequential use of mushroom body neuron subsets during drosophila odor memory processing. *Neuron* 53(1):103-115.

16. Siegel RW & Hall JC (1979) Conditioned responses in courtship behavior of normal and mutant Drosophila. *Proc Natl Acad Sci U S A* 76(7):3430-3434.

17. Park JW, Parisky K, Celotto AM, Reenan RA, & Graveley BR (2004) Identification of alternative splicing regulators by RNA interference in Drosophila. *Proc Natl Acad Sci U S A* 101(45):15974-15979.

18. Geng C & Macdonald PM (2007) Identification of genes that influence gurken expression. *Fly (Austin)* 1(5):259-267.

19. Martin JR, Ernst R, & Heisenberg M (1998) Mushroom bodies suppress locomotor activity in Drosophila melanogaster. *Learn Mem* 5(1-2):179-191.

20. Serway CN, Kaufman RR, Strauss R, & de Belle JS (2009) Mushroom bodies enhance initial motor activity in Drosophila. *J Neurogenet* 23(1-2):173-184.

21. McGuire SE, Le PT, Osborn AJ, Matsumoto K, & Davis RL (2003) Spatiotemporal rescue of memory dysfunction in Drosophila. *Science* 302(5651):1765-1768.

22. McGuire SE, Mao Z, & Davis RL (2004) Spatiotemporal gene expression targeting with the TARGET and gene-switch systems in Drosophila. *Sci STKE* 2004(220):pl6.

23. Grams R & Korge G (1998) The mub gene encodes a protein containing three KH domains and is expressed in the mushroom bodies of Drosophila melanogaster. *Gene* 215(1):191-201.

24. Fernandez-Funez P, *et al*. (2000) Identification of genes that modify ataxin-1-induced neurodegeneration. *Nature* 408(6808):101-106.

25. Skoulakis EM, Kalderon D, & Davis RL (1993) Preferential expression in mushroom bodies of the catalytic subunit of protein kinase A and its role in learning and memory. *Neuron* 11(2):197-208.

26. Han PL, Levin LR, Reed RR, & Davis RL (1992) Preferential expression of the Drosophila rutabaga gene in mushroom bodies, neural centers for learning in insects. *Neuron* 9(4):619-627.

27. Murray EL & Schoenberg DR (2008) Assays for determining poly(A) tail length and the polarity of mRNA decay in mammalian cells. *Methods Enzymol* 448:483-504.

28. Ule J, Jensen K, Mele A, & Darnell RB (2005) CLIP: a method for identifying protein-RNA interaction sites in living cells. *Methods* 37(4):376-386.

29. Zhang YQ, *et al*. (2001) Drosophila fragile X-related gene regulates the MAP1B homolog Futsch to control synaptic structure and function. *Cell* 107(5):591-603.

30. Durbin RM, *et al*. (2010) A map of human genome variation from population-scale sequencing. *Nature* 467(7319):1061-1073.

31. Ganetzky B & Flanagan JR (1978) On the relationship between senescence and age-related changes in two wild-type strains of Drosophila melanogaster. *Exp Gerontol* 13(3-4):189-196.

CHAPTER 4

A genetic-modifier screen identifies factors that functionally interact with the *Drosophila dNab2* gene

ChangHui Pak, John J. Noto, Seth M. Kelly, Anita H. Corbett, Kenneth H. Moberg

CP devised the screen and performed the candidate-based EP screen, all fly experiments, qRT-PCR and immunoblotting. A discovery-based EP screen and fly locomotor behavior tests were performed in collaboration with John J. Noto in the Department of Biochemistry at Emory University (Atlanta, GA). Fly line bearing stable neuronal dNab2 RNAi (*Elav-Gal4;UAS-dNab2-IR)* was built by Seth M. Kelly in the Department of Cell Biology at Emory University (Atlanta, GA).

Introduction

Mechanisms regulating cognitive function are of considerable biological interest. Growing number of developmental brain disorders affecting cognition has allowed for identification and characterization of key regulatory genes important for normal human intellectual development. Intellectual disability (ID), previously referred to as mental retardation, is characterized by limited intellectual capacities reflected by an IQ below 70 and major constraints in adaptive behavior (1). Therapeutic options for the treatment of ID are extremely limited and its comparatively high prevalence of about 2% renders this disorder a major socioeconomic burden (1).

Recently, we identified a locus for unspecific or non-syndromic autosomal recessive ID (NS-ARID) on chromosome 14q31.3 corresponding to mutation of the *ZC3H14* gene in two independent Iranian families (Chapter 2). *ZC3H14* encodes an evolutionarily conserved Cys_3His tandem zinc finger polyadenosine RNA-binding protein (2, 3). The *Drosophila* orthologue of ZC3H14, dNab2, is required specifically in neurons for normal development and behavior (Chapter 2). Pan-neuronal knockdown of *dNab2* results in locomotor and flight defects similar to genomic null allele $dNab2^{ex3}$ and reciprocally, pan-neuronal expression of wild type *dNab2* in animals otherwise null for *dNab2* completely restores locomotor defects and rescues pupal lethality (Chapter 2). Biochemical and genetic evidence show that dNab2 preferentially binds to polyadenosine RNAs and acts to restrict bulk poly(A) tail length *in vivo* (Chapter 2), consistent with the previous findings that the budding yeast orthologue Nab2 is required for proper 3'end formation and poly(A) RNA export from the nucleus (4-6). These findings highlight the

significance of post-transcriptional gene regulatory mechanisms in controlling neural function.

How *dNab2/ZC3H14* controls nervous system function and whether this is mediated solely through the control of poly(A) tail length remains unclear and will certainly be a focus of future studies. At present, we hypothesize that *dNab2* controls neuronal function through specific target mRNAs and interactors as yet unidentified and that tissue-specific variation in the identity of these factors may underlie pleiotropic phenotypes observed in *dNab2* genomic mutants. In an initial attempt to identify these factors, we designed and implemented for a genetic screen intended to identify factors that either act in a complex with dNab2 or that are mRNA targets of dNab2. By utilizing a rough-eye phenotype caused by overexpression of *dNab2* in the developing *Drosophila* eye, we screened a total of 662 loss-of-function and gain-of-function alleles for their ability to modify the dNab2 induced rough-eye phenotype. Here we describe the results of this genetic screen and follow up experiments characterizing the interactions between *dNab2* and *dFmr1* (*Drosophila* fragile X mental retardation gene).

Results

Overexpression of dNab2 reveals tissue-specific phenotypes

To identify tissues that were sensitive to the dosage of dNab2 protein, we overexpressed dNab2 in a tissue-specific manner via the EP-element insertion *EP3716*, located upstream of the *dNab2* locus, in combination with multiple *GAL4* drivers (Table 4.1). Different temperatures were used to achieve varying strengths of expression of the *GAL4* transgenes and flies were then scored for gross organismal phenotypes. As shown in Table 4.1, dNab2 overexpression caused lethality when overexpressed in all tissues (*Act-Gal4*). Similarly, pan-muscle overexpression of *dNab2* (*Mef2-Gal4*) caused lethality at the pupal stage. By contrast, *dNab2* overexpression using various neuronal drivers (*Elav-Gal4, OK6-Gal4, CHA-Gal4*) did not cause any noticeable phenotypes. These results suggest that excess dNab2 protein level is more toxic to muscle cells compared to neuronal tissues. A partially penetrant (~60%) phenotype resulting in scutellar bristle defects was observed when *dNab2* was overexpressed with *Patched-Gal4*, which expresses in the epithelial anterior-posterior axis (Table 4.1). These flies displayed defects in the arrangement of scutellar bristles, including missing, excess bristles, or asymmetrically distributed bristles. Importantly, eye-specific *dNab2* overexpression in developing photoreceptor neurons via the *GMR-Gal4* driver resulted in a rough-eye phenotype that was easily visualized and fully penetrant (Table 4.1). Based on these observations, we chose the *GMR*-driven *dNab2* rough-eye phenotype as the basis of a screen to identify genes that interact with *dNab2* in neuronal cells.

Modification of dNab2 mediated rough-eye phenotype as the basis of a screen

In order to identify factors that genetically interact with *dNab2*, we performed both candidate-based and discovery-based screens for alleles that could modify the roughness and reduced size of *Drosophila* eyes expressing *dNab2* from the *GMR-Gal4* driver. A stable fly stock carrying the *GMR-Gal4* driver and $dNab2^{EP3716}$ insertion (Figure 4.1A) exhibits a 100% penetrant small rough-eye phenotype relative to *GMR-Gal4* control (Figure 4.2A). Candidate modifier alleles were introduced in the $GMR{>}dNab2^{EP3716}$ background by crossing $GMR{>}dNab2^{EP3716}$ flies with either known alleles of genes selected as likely *dNab2* interactors (candidate-based screen) or alleles in the Enhancer-Promoter (EP) collection(7) (discovery-based screen) (Figure 4.1B).

As a proof-of-principle, we also tested whether the $GMR{>}dNab2^{EP3716}$ rough-eye phenotype could be dominantly modified by modulating expression of genes known to act with *dNab2* in controlling poly(A) tail length. Indeed, we found that *dNab2* shows strong genetic interactions with both the poly(A) polymerase *hiragi* and the poly(A) RNA-binding protein *Pabp2* (Figure 4.2B). A single loss-of-function allele of *hiragi*, hrg^{10}, dominantly enhanced the $GMR{>}dNab2^{EP3716}$ rough-eye, resulting in even smaller, rougher, and necrotic eye (Figure 4.2B). In a similar manner, loss-of-function allele of *Pabp2*, $Pabp2^{55}$, dominantly enhanced the $GMR{>}dNab2^{EP3716}$ rough-eye phenotype (Figure 4.2B), whereas co-overexpression of *Pabp2* via $Pabp2^{EP2264}$ allele strongly suppressed the $GMR{>}dNab2^{EP3716}$ small, rough-eye phenotype (Figure 4.2B). These genetic relationships entirely match our previous finding that *dNab2* antagonizes poly(A) tail length (Chapter 2) while both *hrg* and *Pabp2* promote it (8, 9).

Candidate-based screen for dNab2-dominant interactors

We carried out a candidate-based screen of genes encoding RNA-binding proteins, RNA processing factors, translational regulatory proteins and proteins known to regulate neural function (Table 4.2). Loss-of-function and gain-of-function alleles were tested for their ability to modify the *GMR>dNab2*^{EP3716} rough-eye phenotype. Subsequent enhancers/suppressors of the rough-eye phenotype were then crossed to *GMR-Gal4* alone to eliminate alleles, which produced phenotypes independent of *dNab2*.

As shown in Table 4.2, we identified 24 alleles (out of 112 tested) that were able to dominantly modify the rough-eye phenotype produced by *GMR>dNab2*^{EP3716}. Three of these alleles produced eye phenotypes in combination with *GMR-Gal4* and were therefore eliminated from further analysis. The remaining 21 alleles correspond to 18 genes encoding factors involved in mRNA processing and translation (i.e. *mushroom-body expressed, Rm62, smaug, mitochondrial ribosomal protein S29* (*mRpS25*) and *Eukaryotic initiation factor 4E*) as well as factors involved in RNA localization, long-term memory and synaptic transmission (i.e. *couch potato, dFmr1* and *staufen*) (Table 4.2).

Discovery-based screen for dNab2-dominant interactors

In parallel to the candidate-based screen above, we also carried out a discovery-based screen for *dNab2* interacting genes. For this, we chose to screen against a small subset of (EP) collection (7) consisting of 550 lines randomly distributed throughout the *Drosophila* genome. Depending upon the orientation of the EP insertion in the genome, the affected gene can either be overexpressed or disrupted. As noted in Table 4.3, "sense

orientation" designates an EP element that is inserted in the same orientation of the gene, which is then able to overexpress the gene in the presence of *Gal4*. The "antisense orientation" designates an EP element that is inserted in the opposite orientation of the gene, which could potentially disrupt gene function. Of 550 lines tested, 85 EP lines dominantly modified the *GMR>dNab2*^{EP3716} rough-eye phenotype (25 suppressors, 60 enhancers) but had no eye phenotypes on their own when crossed to *GMR-Gal4*. It is notable that an EP element in the *mRpS25* gene scored as a strong modifier of the *GMR>dNab2*^{EP3716} rough-eye phenotype and that an allele of *mRpS25* also scored in our parallel candidate screen (Table 4.2). This result validates each screening approach and suggests that the results from one may be used to corroborate the other.

Interestingly, multiple genes involved in neuronal functions such as learning and memory were strong modifiers of *dNab2*-mediated rough eye phenotype. These genes include *leak* (*robo2), rhomboid (rho), silver (svr)* and *rutabaga* (*rut)*. These findings suggest that *dNab2* might act in pathways that play critical roles learning and memory formation either together or upstream of these learning and memory genes. Consistent with the neuronal function of *dNab2*, *Rab-protein 6 (Rab6)* and *egalitarian (egl)*, which are important for RNA localization and RNA transport, were also strong modifiers of *GMR>dNab2*^{EP3716} rough-eye phenotype, suggesting that *dNab2* might function to affect localization of mRNAs.

dFmr1 dominantly modifies dNab2 mediated locomotor defect

In the course of the candidate-based screen, we found that introducing one null allele of *dFmr1* (Δ50M) was sufficient to suppress the rough-eye phenotype caused by

GMR>dNab2^{EP3716} (see Table 4.2; Figure 4.3A), suggesting that *dFmr1* might act in the same or in parallel genetic pathways as *dNab2*. Due to this observation, we decided to further characterize the genetic relationships between *dNab2* and *dFmr1*. *dFmr1* encodes the *Drosophila* orthologue of Fragile X mental retardation protein (FMRP) (10), in which mutations affecting FMRP levels lead to the most commonly inherited X-linked human intellectual disability (11). FMRP is an mRNA-binding protein that binds to higher-order structure RNAs (12-14) and negatively regulate translation (15, 16). Interestingly, *dFmr1* mutant flies exhibit similar developmental phenotypes that resemble those of *dNab2* zygotic mutants, including defects in wing position, locomotor activity, flight, courtship-mediated short-term memory and MB β-lobe fusion (17-20), further strengthening the argument that both genes might act in similar neurodevelopmental pathways.

To pursue this hypothesis further, we tested whether heterozygosity for a loss-of-function allele of *dFmr1* could dominantly modify a *dNab2* loss-of-function locomotor phenotype. Both complete loss-of-function and pan-neuronal knockdown of *dNab2* results in poor locomotor activity (Chapter 2). Using a fly line that stably expresses *dNab2* RNAi transgene in pan-neuronal manner (*Elav-Gal4;UAS-dNab2-IR)*, we tested dominant modification by *dFmr1* measured by negative geotaxis assay. Significantly, introducing a single loss-of-function allele of *dFmr1* (Δ113M) enhanced the locomotor defect of flies lacking neuronal *dNab2* but had no effect on *Elav-Gal4* control flies (Figure 4.4), confirming that both *dNab2* and *dFmr1* act in similar genetic pathways or alternatively, *dFmr1* acts downstream of *dNab2* to modulate normal locomotor behavior in neurons.

Loss of dNab2 affects dFmr1 expression

As dNab2 is a polyadenosine RNA-binding protein and regulates tail length of poly(A) RNAs (Chapter 2), we tested whether dNab2 regulates *dFmr1* mRNA abundance. By qRT-PCR, we observed a modest downregulation of *dFmr1* mRNA levels (~20%) in *dNab2* null heads compared to control (Figure 4.5A). This downregulation of *dFmr1* mRNA levels was more prominent in older animals (Day 11-13) showing further decrease in *Fmr1* mRNA level (~70%) (Figure 4.5A'). Parallel age-dependent decline in locomotor activity was observed in both *dNab2* genomic null mutants and pan-neuronal *dNab2* knock down flies (Chapter 2), which suggests that the loss of *dNab2* may lead to a progressive neurodegeneration phenotype facilitated by progressive loss or decrease of *dFmr1* transcripts.

In parallel to the analysis of *dFmr1* mRNA, we also looked at protein levels in *dNab2* null heads by immunoblotting with an anti-dFMRP antibody (Developmental studies hybridoma bank). This analysis produced a very interesting result: when RIPA buffer (see Experimental procedures) was used to solubilize protein, we detected less dFMRP in *dNab2* mutant heads relative to control heads (Figure 4.5B), but when a more stringent SDS-sample buffer was used to extract proteins, dFMRP levels were unchanged (Figure 4.5B'). Loading was controlled by immunoblotting with a β-tubulin antibody. This link between the *dNab2* genotype and dFMRP extractability was also observed in human immortalized cell lines comparing human FMRP in *ZC3H14* null and control patient lymphoblasts (R154X) (S.W. Leung, unpublished data). Moreover, preliminary evidence from *Fmr1* knockout mice suggests that the lack of FMRP shifts ZC3H14 protein into a more soluble fraction (C. Gross, unpublished data). While these findings

are very preliminary, they could be indicative of reciprocal requirements for dFMRP/FMRP and dNab2/ZC3H14 in regulating the subcellular distribution of the other.

Discussion

In order to gain insight into the molecular mechanisms underlying ZC3H14-associated human intellectual disability, we sought out to identify the molecular partners or downstream targets of dNab2, the *Drosophila* orthologue of ZC3H14. To this end, we devised and implemented a screen for genetic modifiers a rough-eye phenotype caused by overexpression of *dNab2* in the developing *Drosophila* eye. A total of 662 loss-of-function and gain-of-function alleles were tested and of those, 106 alleles were found to be suppressors or enhancers of the *dNab2* mediated rough-eye phenotype. Among these modifiers are factors involved different aspects of post-transcriptional regulation of gene expression (e.g. polyadenylation, nuclear alternative splicing, translation, and RNA localization) as well as factors important for learning and memory. These results are consistent with the previous findings that place dNab2 in post-transcriptional regulatory steps to modulate gene expression (Chapter 2; Appendix). Moreover, our identification of factors that regulate learning and memory as modifiers of *dNab2* overexpression provides additional support for a role for dNab2 in short-term memory formation (Chapter 3).

To extend from the results of the screen, we chose to focus on the genetic interaction between *dNab2* and *dFmr1*, the *Drosophila* fragile X mental retardation gene (10), which encodes an mRNA-binding protein, dFMRP, that acts as a negative regulator of translation (15, 16). Loss of function allele of *dFmr1* dominantly suppresses the *GMR>dNab2*^{EP3716} gain-of-function adult eye phenotype and dominantly enhances the dNab2 loss-of-function locomotor phenotype (Chapter 2). This pattern of *dFmr1* modification of the *dNab2* gain-of-function and loss-of-function phenotypes indicates that dFmr1 and dNab2 act synergistically *in vivo*. How they do so is not yet clear.

However, in addition to our genetic data, we also have molecular evidence showing that dNab2 regulates *dFmr1* mRNA abundance as well as dFMRP extractability and/or stability. Yet, the precise molecular mechanism by which dNab2 regulates *dFmr1* expression is still unclear. Mechanistically, *dFmr1* and *dNab2* could act together, in parallel, or alternatively *dFmr1* could be a downstream target of dNab2 function. Thus, one focus of future work will be to determine whether this interaction is via the effects on poly(A) tail length or via a direct protein-protein interaction. In addition, finding out the subcellular localizations of dNab2 and dFMRP in cultured neurons would allow better understanding of spatial regulation of these two proteins in such highly polarized cells. Preliminary data suggest that ZC3H14 localizes to both the nucleus and axonal cytoplasm in primary rat hippocampal cultured neurons (C. Gross, unpublished data). Perhaps we can speculate a model where ZC3H14/dNab2 have dual functions in both the nucleus and cytoplasm where they can contact messenger ribonucleoprotein (mRNP) complexes associated with FMRP/dFMRP and affect translation and/or mRNA transport including *FMR1/dFmr1* transcripts. At a genetic level, it will be interesting to test whether loss of either factor affects the localization of the other in cultured *Drosophila* brain neurons, and whether overexpression of dFMRP is sufficient to rescue behavioral phenotypes associated with neuronal loss of dNab2 (Chapter 2, 3).

In summary, we have designed and implemented a genetic screen that has identified numerous candidate *dNab2*-interacting genes that could potentially encode factors that complex with dNab2 protein or that represent mRNAs that are by dNab2 in neurons. This work provides a strong foundation for future studies of molecular

mechanisms underlying the neuronal role of *Drosophila* dNab2 and the etiology of ZC3H14-associated human intellectual disability.

Table 4.1. Tissue-specific misexpression of *dNab2*.

GAL4 driver	**Expression pattern**	**Phenotype**	**Temperature (°C)**
Actin	All cell types	Pupal lethal	25
β-tubulin	All cell types	Larval lethal	25
		Larval lethal	21
		Lethal	18
MEF2	Pan-muscle	Pupal lethal	25
		Pupal lethal	21
		Pupal lethal	18
24B	Pan-muscle	Reduced viability; thinner body	25
		Larval lethal	29
5058	Larval muscle #12	-	25
Mhc	Muscle	-	25
Elav (C155)	Pan-neuronal	-	25
		-	29
OK6	Motor neurons	-	25
		-	29
CHA	Cholinergic neurons	-	25
		-	29
Patched	Epithelial anterior-posterior axis	Defects in scutellar bristles	25
Engrailed	Epithelial posterior compartment	-	25
Wingless	Epithelial dorsal-ventral axis	-	25
C96	Wing margin precursors	-	25
GMR	Eye	Rough-eye	25
		Rough-eye	29
Eyeless	Eye	-	25
		-	29

(-) denotes no phenotypes observed.

Table 4.2. Candidate genes tested for genetic interaction with *dNab2* in the eye

Gene	Allele§	E or S*	Function and/or associated biological process
abstrakt	abs^{00620}	-	ATP-dependent helicase; nervous system development, mRNA splicing
	abs^{f01698}	-	
	$abs^{EY04479}$	-	
Adar	$Adar^{BG02235}$	-	double-stranded RNA adenosine deaminase; adult behavior
Argonaute 2	$AGO2^{EY04479}$	-	RNA interference
Ataxin-2	$Atx2^{06490}$	-	regulation of actin filament polymerization
	$Atx2^{DG08112}$	S	
Aubergine	aub^{QC42}	-	piRNA binding
	aub^{HN}	-	
	$aub^{KG05389}$	-	
belle	bel^{6}	-	ATP-dependent RNA helicase
	bel^{NEO30}	S	
	bel^{CAP1}	-	
Bicaudal D	$BicD^{r5}$	-	mRNA transport
bancal	$bl^{KG02524}$	-	alternative nuclear mRNA splicing
	$bl^{EY09813}$	-	
bruno-2	$bru\text{-}2^{f00171}$	-	negative regulation of translation
	$bru\text{-}2^{EY18918}$	-	
	$bru\text{-}2^{MB00431}$	-	
	$bru\text{-}2^{G5819}$	-	
Clipper	Clp^{G2556}	-	mRNA cleavage; mRNA polyadenylation;Dm CPSF
couch potato	cpo^{1432}	S	mRNA binding; synaptic transmission
CCR4/twin	$CCR4/twin^{8115}$	-	nuclear-transcribed mRNA poly(A) tail shortening
discs overgrown	dco^{J3B9}	-	kinase activity
Dodeca-satellite-binding protein 1	$Dp1^{BG0145B}$	-	mRNA 3'UTR-binding
	$Dp1^{BG02288}$	-	
Elongation factor 1α48D	$Ef1\alpha48D^{1275}$	-	translation elongation factor activity
Eukaryotic initiation factor 4a	$eIF\text{-}4a^{K01501}$	-	translation initiation factor activity; RNA helicase activity
Eukaryotic initiation factor 4E	$eIF\text{-}4E^{7238}$	S	translation initiation factor activity; RNA cap binding

Fmr1	UAS-Fmr1.Z§	lethal	synaptic transmission; locomotory behavior; learning or memory
	$Fmr1^{\Delta 113M}$	S	
	$Fmr1^{\Delta 50M}$	S	
found in neurons	UAS-fne.C 4-10B§	lethal	poly(U) RNA binding
Glutamate receptor IIA	$GluRIIA^{AD9}$	-	glutamate receptor activity; synaptic transmission
	$GluRIIA^{Sp16}$	-	
Helicase at 25E	$Hel25E^{e02545}$	S	ATP-dependent RNA helicase activity
hoi-polloi	$hoip^{K07104}$	S	mRNA binding; nervous system development
held out wings	how^{J5B5}	-	mRNA 3'-UTR binding
Heterogeneous nuclear ribonucleoprotein at 27C	$Hrb27C^{2647}$	-	regulation of nuclear mRNA splicing; axon guidance
Heterogeneous nuclear ribonucleoprotein at 87F	$Hrb87F^{KG02089}$	-	regulation of alternative nuclear mRNA splicing
Heterogeneous nuclear ribonucleoprotein at 98DE	$Hrb98DE^{ZCL0558}$	-	negative regulation of RNA splicing
hiragi	hrg^{10}	E+++	mRNA polyadenylation
	hrg^{p1}	-	
	hrg^{1}	-	
IGF-II mRNA-binding protein	Imp^{EP760}	-	mRNA binding; nervous system development
lark	UAS-lark-3HA 23A§	E	mRNA binding
	$lark^{EY00297}$	-	
	$lark^{DG23107}$	-	
	$lark^{EY23084}$	-	
maelstrom	$mael^{R20}$	-	intracellular mRNA localization
	$mael^{KG03309}$	-	
	$mael^{EY08554}$	-	
mitochondrial ribosomal protein S29	$mRpS29^{KG07362}$	S	structural constituent of ribosome; translation
musashi	msi^{1}	-	mRNA binding; negative regulation of translation

mushroom-body expressed	mub^{4093}	S	poly(C) RNA binding; regulation of alternative nuclear mRNA splicing
muscleblind	mbl^{E27}	-	Zinc finger, CCCH-type
no on or off transient A	$nonA^{4B18}$	-	poly-pyrimidine tract binding; mRNA binding
nanos	nos^{L7}	-	dendrite morphogenesis
NTF2-related export protein 1	$Nxt1^{f04855}$	-	mRNA export
	$Nxt1^{DG05102}$	-	
oo18 RNA-binding protein	orb^{dec}	-	mRNA localization; mRNA polyadenylation
	$orb^{EY08547}$	-	
orb2	$orb2^{BG02373}$	-	long-term memory; male courtship behavior
polyA-binding protein	$pAbp^{K10109}$	-	positive regulation of translation
	$pAbp^{k10109}$	-	
	$pAbp^{EP310}$	-	
	$pAbp^{EY11561}$	-	
Pabp2	$Pabp2^{EP2264}$	S+++	mRNA polyadenylation
	$Pabp2^{55}$	E+++	
	$Pabp2^{6}$	E	
Parkin	$park^{25}$	-	protein autoubiquitination
pasilla	ps^{10615}	-	mRNA binding; nuclear mRNA splicing
	$ps^{MB04043}$	-	
poly U binding factor 68kD	$pUf68^{EY07952}$	-	nuclear mRNA splicing
pumilio	pum^{13}	-	mRNA 3'-UTR binding; long term memory
	pum^{3}	-	
	pum^{BEM}	-	
Purine-rich binding protein-α	$Pur\text{-}\alpha^{KG05743}$	S	transcription activator activity
quaking related 58E-2	$qkr58E\text{-}2^{KG07766}$	-	mRNA binding
	$qkr58E\text{-}2^{EP2103}$	E+++	
quaking related 58E-3	$qkr58E\text{-}3^{EY02038}$	-	regulation of alternative nuclear mRNA splicing

RNA-binding protein 9	Rbp9^{P2690}	-	mRNA binding
Resistant to dieldrin	Rd^{1}	-	GABA-A receptor activity
Rm62	Rm62^{1086}	-	RNA interference; regulation of alternative nuclear mRNA splicing
	Rm62^{EY06795}	S	
	Rm62^{DG12402}	-	
	Rm62^{EY10915}	-	
	Rm62 (excision L3)	-	
SC35	SC35^{KG02986}	-	regulation of alternative nuclear mRNA splicing
Stem-loop binding protein	Slbp^{EP1045}	-	histone mRNA 3'-end processing
smaug	smg^{1}	E++	nuclear-transcribed mRNA poly(A) tail shortening
small ribonucleoprotein particle U1 subunit 70K	snRNP-U1-70K^{2107}	-	negative regulation of nuclear mRNA splicing
split ends	spen^{3}	-	transcription regulator activity
	spen^{3350}	S	
squid	sqd^{J6E3}	-	mRNA binding
staufen	stau^{RY9}	-	RNA localization; long term memory
	stau^{1}	S++	
TBPH	TBPH^{KG08578}	-	mRNA binding; adult locomotor behavior
Tis11 homolog	Tis11^{BG00309}	-	RNA interference; Zinc finger, CCCH-type
	Tis11^{EY09433}	-	
	Tis11^{EY09107}	-	
transformer 2	tra2^{KG08361}	-	mRNA processing; regulation of nuclear mRNA splicing
Trf4-2	Trf4-2^{EY05585}	S	polyadenylation of snRNAs
tsunagi	tsu^{EO567}	-	mRNA binding
	tsu^{KG04415}	-	
tudor	tud^{KG10175}	-	P granule organization
vasa intronic gene	vig^{EY07816}	-	heterochromatin organization; RNA interference
Zn72D	Zn72D^{BG02677}	-	mRNA binding

small bristles	sbr^{1}	-	mRNA export
	$sbr^{Magellan}$	-	
wispy	$wisp^{12-3147}$	-	polynucleotide adenylyltransferase activity
	$wisp^{KG05287}$	-	

(-) denotes no interaction.
*E=enhancement, S=suppression
+ indicates modification strength (+ = mild, ++ = moderate, +++ = strong).
§ denotes modifier alleles that produced phenotype with *GMR-Gal4*.

Table 4.3. Modifiers of *dNab2*-mediated rough-eye phenotype in a discovery-based screen

Gene	**EP tested (BL stock#)§**	**s or as**	**E or S***	**Function and/or associated biological process**
Rab-protein 6 (Rab6)	17051	as	E++	mRNA localization during oogenesis
leak (robo2)	17071	as	S++	axon guidance receptor activity
XNP/dATRX	17188	s	E+++	ATPase activity; chromatin remodeling
egalitarian (egl)	17205	as	E	mRNA transport
rhomboid (rho)	17276	as	E++	serine-type peptidase activity; learning and memory
TBP-associated factor 1 (Taf-1)	17281	as	S+++	positive regulation of transcription
mitochondrial ribosomal protein S25 (mRpS25)	11418	as	S+	structural constituent of ribosome
silver (svr)	10042	s	S++	carboxypeptidase activity; long term memory
longitudinals lacking (lola)	17254	as	E++	neuromuscular process
Protein tyrosine phosphatase 4E (Ptp4E)	10088	s	S+	motor axon guidance; central nervous system development
rutabaga (rut)	10126	s	S+	calcium- and calmodulin-responsive adenylate cyclase activity; learning and memory

(-) denotes no interaction.
BL = Bloomington
s=sense orientation of EP insertion with respect to the affected gene
as=antisense orientation of EP insertion with respect to the affected gene
*E=enhancement, S=suppression
+ indicates modification strength (+ = mild, ++ = moderate, +++ = strong).
§ denotes alleles that produced phenotype with *GMR-Gal4*.

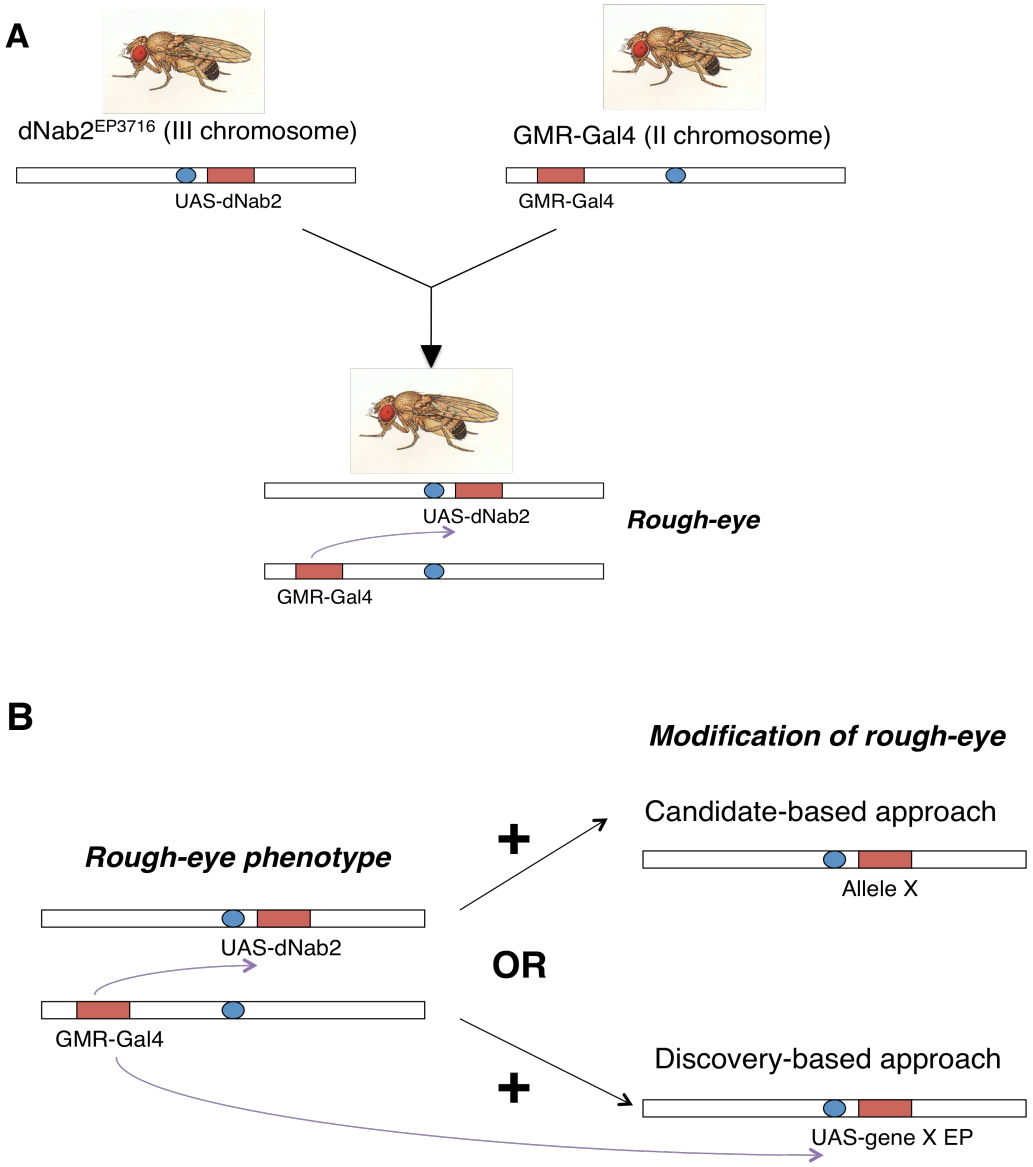


Figure 4.1. A schematic of the genetic screen. (**A**) A chromosome bearing an EP-element (*dNab2*^{EP3716}) and a chromosome bearing a *GMR-Gal4* were combined to produce a fly line that stably expresses *GMR>dNab2*^{EP3716} with a rough-eye phenotype. (**B**) dNab2 overexpressing flies (*GMR>dNab2*^{EP3716}) were crossed with either collection of a known allele (Allele X) or a subset of the EP collection (UAS-gene X EP) in a candidate-based screen or a discovery-based screen, respectively.

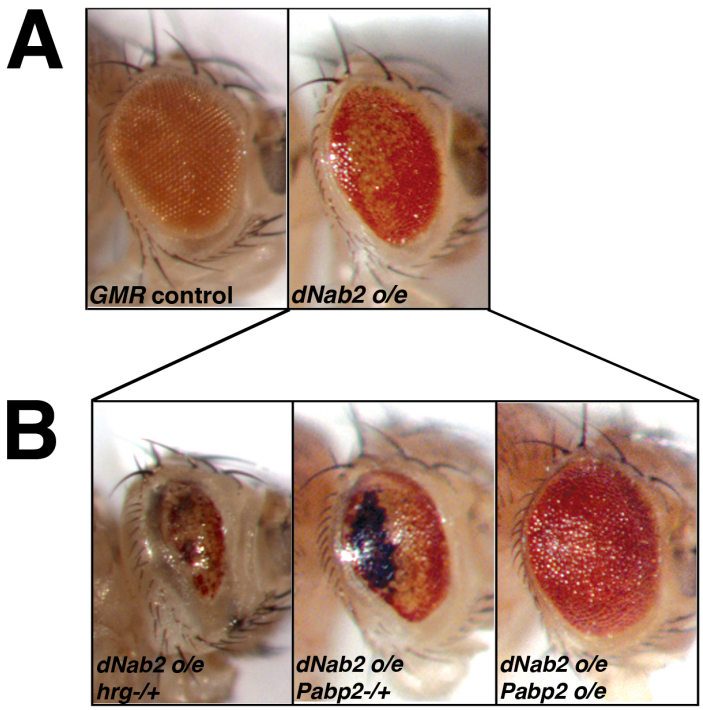


Figure 4.2. Modification of *GMR>dNab2*^{EP3716} rough-eye phenotype. Light microscopic images of adult fly eye for the following genotypes: (**A**) *GMR-Gal4*/+ (*GMR* control) and *GMR-Gal4*/+;*dNab2*^{EP3716}/+ (*dNab2 o/e*), (**B**) *GMR-Gal4*/*hrg*^{10};*dNab2*^{EP3716}/+ (*dNab2 o/e hrg-*/+), *GMR-Gal4*/*Pabp2*^{55};*dNab2*^{EP3716}/+ (*dNab2 o/e Pabp2-*/+), *GMR-Gal4*/*Pabp2*^{EP2264};*dNab2*^{EP3716}/+ (*dNab2 o/e Pabp2 o/e*). o/e=overexpression.

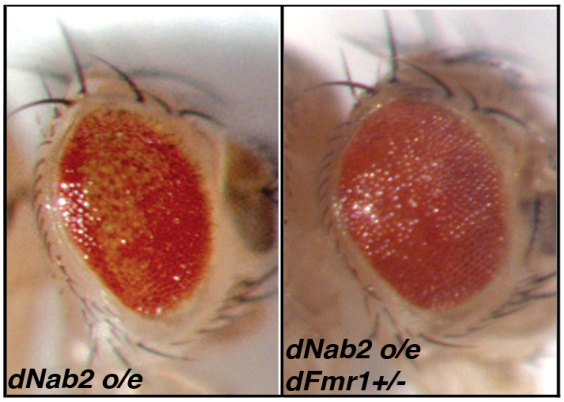


Figure 4.3. Loss-of-function allele of *dFmr1* dominantly suppresses *dNab2*-mediated rough-eye phenotype. Light microscopic images of adult fly eye for the following genotypes: *GMR-Gal4/+;dNab2*^{EP3716}*/+* (*dNab2 o/e*) and *GMR-Gal4/+;dNab2*^{EP3716}*/dFmr1*^{delta50M} (*dNab2 o/e dFmr1 +/-*).

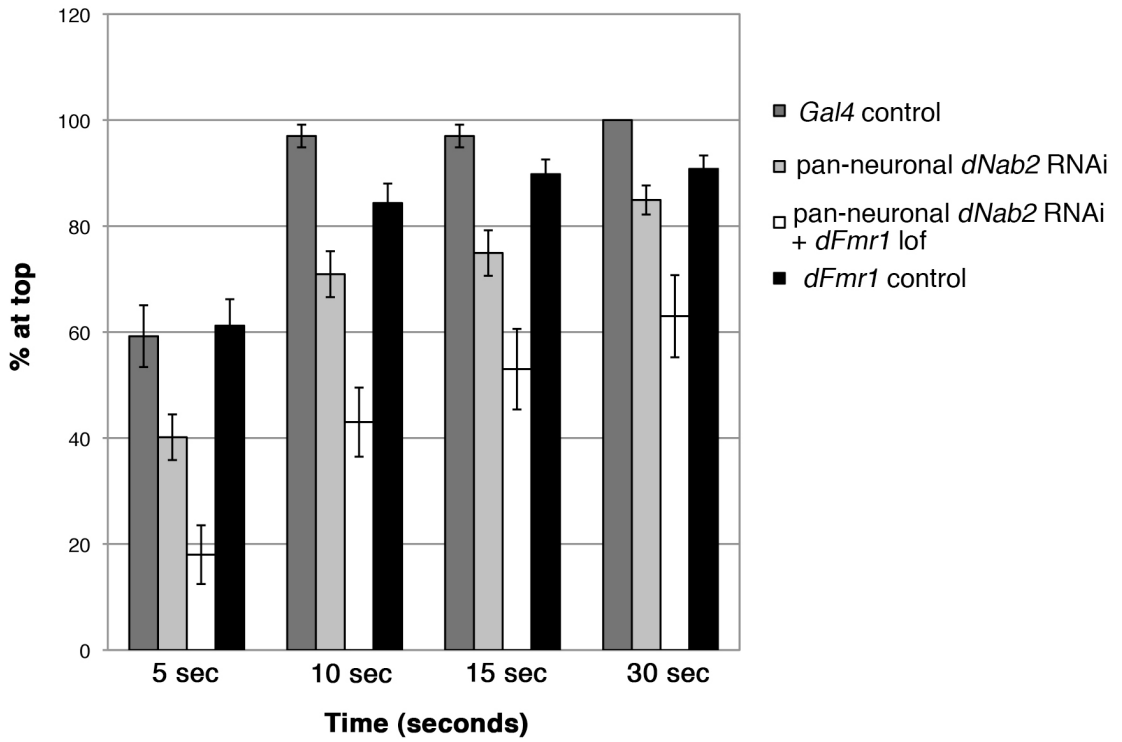


Figure 4.4. Loss-of-function allele of *dFmr1* dominantly suppresses the locomotor defect in flies knocked down for *dNab2* in neurons. Results from negative geotaxis assay for *Gal4* control (*Elav-Gal4*), pan-neuronal *dNab2* RNAi (*Elav-Gal4/+;dNab2-IR/+*), pan-neuronal *dNab2* RNAi + *dFmr1* lof (*Elav-Gal4/+;dNab2-IR/+;dFmr1*^{delta113M}*/+*), and *dFmr1* control (*Elav-Gal4/+; dFmr1*^{delta113M}*/+*). Data are presented as the average percentage of flies reaching the top of a 25-mL graduated cylinder after 5, 10, 15 and 30 seconds across all trials. Groups of ten flies with mixed ages were tested in at least ten independent trials for each genotype. $p<0.05$ (two-tailed *t*-test) when comparing *Gal4* control to pan-neuronal *dNab2* RNAi and pan-neuronal *dNab2* RNAi to pan-neuronal *dNab2* RNAi + *dFmr1* lof at 5 seconds. Error bars=s.e.m.

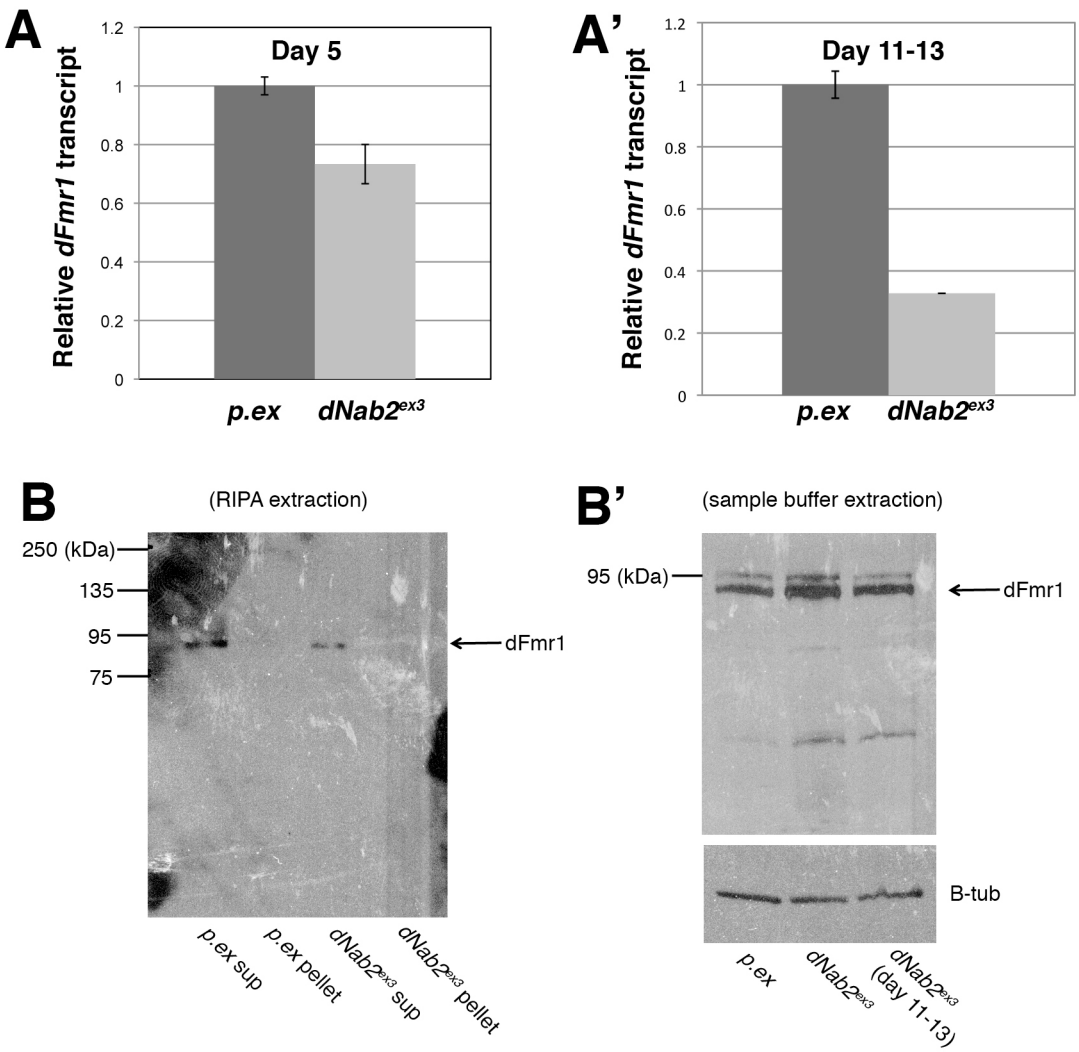


Figure 4.5. Expression levels of *dFmr1* mRNA and dFMRP are altered in *dNab2* zygotic mutant heads. (**A-A'**) Quantitative RT-PCR analysis of *dFmr1* transcript levels in adults fly heads. All genotypes were normalized to *dFmr1* transcript levels in *p. ex* control animals (set to 1.0). *β-tub* is an internal control. Error bars=s.d. Representative of a single experiment. (**B-B'**) Anti-dFMRP immunoblot of *p.ex* control and *dNab2*^{ex3} mutant adult heads. Proteins were extracted with RIPA buffer (**B**) or SDS-sample buffer (**B'**). Anti-β-tubulin is shown as a loading control.

Experimental procedures

***Drosophila* stocks and genetics:** All crosses and stocks were maintained in standard conditions unless otherwise noted. Precise and imprecise excision alleles (*p.ex* and *dNab2*^{ex3}) were obtained from previous studies (Chapter 2). *P{EP}CG5720[EP3716]* and stocks for genetic screen were obtained from Bloomington Drosophila Stock Center (BDSC) at Indiana University. *Pabp2* alleles were gift of Martine Simonelig. The RNAi line for *dNab2* was obtained from Vienna Drosophila Research Center.

***Drosophila* behavioral assays:** The negative-geotaxis assay was performed as previously described (21) with the following modifications: Briefly, newly eclosed flies (day 0) were collected daily, divided into groups of ten, and kept in separate vials for aging. Cohorts of age-matched flies were then transferred to a 25-mL graduated cylinder, gently tapped to the bottom, and analyzed for a climbing response. The number of flies reaching the top after 5, 10, 15, and 30 seconds post-tap was recorded. At least 10 groups were tested for each genotype and sex (where indicated).

Immunoblotting: Immunoblotting was performed as described previously (3). Mouse anti-dFMRP (Developmental Studies Hybridoma Bank, 1:1000) and goat anti-β-tubulin (Santa Cruz, 1:10,000) were used. RIPA buffer (150 mM NaCl, 1% NP40, 0.5% deoxycholate, 0.1% SDS and 50 mM Tris 8.0).

Microscopy: *Drosophila* eye images were photographed with a Leica DFC500 charge-coupled device digital camera. Postacquisition image processing was done using Adobe Photoshop.

Quantitative real-time RT-PCR (qRT-PCR): Total RNA from *Drosophila* samples was extracted using TRIzol reagent (Invitrogen). cDNAs were generated by using the

SuperScript III Reverse Transcriptase (Invitrogen). Quantitative Real-time PCR was performed according to standard protocols with LightCycler 480 (22). Primer sequences for *dFmr1*: caaacgaagtgcaaaagtgc (left), tgtactcaaatcagtcttttgcttg (right).

Reference

1. Ropers HH (2010) Genetics of early onset cognitive impairment. *Annu Rev Genomics Hum Genet* 11:161-187.
2. Kelly SM, *et al*. (2007) Recognition of polyadenosine RNA by zinc finger proteins. *Proc Natl Acad Sci U S A* 104(30):12306-12311.
3. Leung SW, *et al*. (2009) Splice variants of the human ZC3H14 gene generate multiple isoforms of a zinc finger polyadenosine RNA binding protein. *Gene* 439(1-2):71-78.
4. Anderson JT, Wilson SM, Datar KV, & Swanson MS (1993) NAB2: a yeast nuclear polyadenylated RNA-binding protein essential for cell viability. *Mol Cell Biol* 13(5):2730-2741.
5. Green DM, *et al*. (2002) Nab2p is required for poly(A) RNA export in Saccharomyces cerevisiae and is regulated by arginine methylation via Hmt1p. *J Biol Chem* 277(10):7752-7760.
6. Hector RE, *et al*. (2002) Dual requirement for yeast hnRNP Nab2p in mRNA poly(A) tail length control and nuclear export. *EMBO J* 21(7):1800-1810.
7. Rorth P (1996) A modular misexpression screen in Drosophila detecting tissue-specific phenotypes. *Proc Natl Acad Sci U S A* 93(22):12418-12422.
8. Murata T, *et al*. (2001) The hiiragi gene encodes a poly(A) polymerase, which controls the formation of the wing margin in Drosophila melanogaster. *Dev Biol* 233(1):137-147.

9. Benoit B, *et al*. (2005) An essential cytoplasmic function for the nuclear poly(A) binding protein, PABP2, in poly(A) tail length control and early development in Drosophila. *Dev Cell* 9(4):511-522.

10. Wan L, Dockendorff TC, Jongens TA, & Dreyfuss G (2000) Characterization of dFMR1, a Drosophila melanogaster homolog of the fragile X mental retardation protein. *Mol Cell Biol* 20(22):8536-8547.

11. Hagerman PJ (2008) The fragile X prevalence paradox. *J Med Genet* 45(8):498-499.

12. Darnell JC, *et al*. (2001) Fragile X mental retardation protein targets G quartet mRNAs important for neuronal function. *Cell* 107(4):489-499.

13. Schaeffer C, *et al*. (2001) The fragile X mental retardation protein binds specifically to its mRNA via a purine quartet motif. *EMBO J* 20(17):4803-4813.

14. Siomi H, Siomi MC, Nussbaum RL, & Dreyfuss G (1993) The protein product of the fragile X gene, FMR1, has characteristics of an RNA-binding protein. *Cell* 74(2):291-298.

15. Laggerbauer B, Ostareck D, Keidel EM, Ostareck-Lederer A, & Fischer U (2001) Evidence that fragile X mental retardation protein is a negative regulator of translation. *Hum Mol Genet* 10(4):329-338.

16. Li Z, *et al*. (2001) The fragile X mental retardation protein inhibits translation via interacting with mRNA. *Nucleic Acids Res* 29(11):2276-2283.

17. Dockendorff TC, *et al*. (2002) Drosophila lacking dfmr1 activity show defects in circadian output and fail to maintain courtship interest. *Neuron* 34(6):973-984.

18. Michel CI, Kraft R, & Restifo LL (2004) Defective neuronal development in the mushroom bodies of Drosophila fragile X mental retardation 1 mutants. *J Neurosci* 24(25):5798-5809.

19. Morales J, *et al*. (2002) Drosophila fragile X protein, DFXR, regulates neuronal morphology and function in the brain. *Neuron* 34(6):961-972.

20. Zhang YQ, *et al*. (2001) Drosophila fragile X-related gene regulates the MAP1B homolog Futsch to control synaptic structure and function. *Cell* 107(5):591-603.

21. Ganetzky B & Flanagan JR (1978) On the relationship between senescence and age-related changes in two wild-type strains of Drosophila melanogaster. *Exp Gerontol* 13(3-4):189-196.

22. Durbin RM, *et al*. (2010) A map of human genome variation from population-scale sequencing. *Nature* 467(7319):1061-1073.

CHAPTER 5

Conclusion and Discussion

Conclusion

Here we describe the first identification and characterization of dNab2, a *Drosophila* orthologue of ZC3H14/Nab2 class of Cys_3His (CCCH) tandem zinc finger (ZnF) polyadenosine RNA-binding protein (1, 2). We show for the first time that mutations in the human *ZC3H14* lead to non-syndromic autosomal recessive intellectual disability (NS-ARID) and use *Drosophila melanogaster* to model key aspects of the disease. Our results reveal that dNab2 is essential for development and required in neurons for normal locomotion and flight (Chapter 2). Furthermore, biochemical and genetic data indicate that dNab2 restricts bulk poly(A) tail length *in vivo*, suggesting that this function may underlie its role in development and disease (Chapter 2). We go on to refine the role of dNab2 in controlling locomotor activity and memory formation in the *Drosophila* mushroom bodies (MBs), a highly specialized structure involved in higher cognitive functions and locomotion (3) (Chapter 3). Finally, using a genetic modifier screen, we identify putative dNab2 targets and/or interacting proteins that modulate dNab2-mediated neuronal function (Chapter 4). These studies reveal a conserved requirement for *ZC3H14/dNab2* in the metazoan nervous system and allow for future studies on the molecular mechanisms underlying ZC3H14-associated human intellectual disability.

A model

Our findings that *ZC3H14* is mutated in patients with NS-ARID and that *dNab2* is both necessary and sufficient to promote normal behavior in all neurons, including the MB neurons, illustrate the overall significance of *dNab2/ZC3H14* in neuronal function. We have evidence supporting the role of *dNab2* in controlling poly(A) tail length and

suggest that misregulation of poly(A) tails of mRNAs could potentially underlie the developmental defects observed by loss of *dNab2*. However, whether misregulated poly(A) tails are direct cause of the neuronal defects seen in *dNab2* mutants or patients are yet to be proven.

The precise molecular function of *dNab2/ZC3H14* in neurons is still unclear; however, we propose a model based upon our genetic, biochemical and cell biological studies whereby *dNab2/ZC3H14* plays an important role in regulating translation of specific neuronal mRNAs and proper transport of these mRNAs along the axonal cytoplasm (Figure 5.1). Our model is based on the following observations: dNab2 acts to restrict poly(A) tail length *in vivo* (Chapter 2); *dNab2* acts in similar genetic pathways as *dFmr1*, a known negative regulator of translation important for neuronal function (4, 5) (Chapter 4); *dNab2* genetically interacts with known components of RNA transport and translation as well as specific genes involved in learning and memory (Chapter 4); ZC3H14 (isoforms 1-3) localizes to both poly(A) RNA containing nuclear speckles and the axonal cytoplasm including the growth cones (C. Gross, unpublished data); and ZC3H14 (isoforms 1-3) associates with 80S ribosomes (C. Gross, unpublished data). We suggest that dNab2/ZC3H14 binds to poly(A) tails of mRNAs in the nucleus and interacts with other poly(A)-binding proteins (Pabs), such as Pabp2/PABPN1. Since dNab2/ZC3H14 function is antagonistic to Pabp2/PABPN1 and limiting the length of poly(A) tails, its activity is probably inhibited by the polyadenylation machinery including Pabp2/PABPN1, which acts positively to lengthen poly(A) tails in the nucleus (6, 7). Once properly polyadenylated transcripts are exported out into the cytoplasm, dNab2/ZC3H14 associates with translation factors and other translational regulators, such

as dFMRP/FMRP, to repress translation. It is tempting to spectulate that dNab2/ZC3H14 acts in concert with factors, such as CPEB and Maskin, to regulate translation through controlling polyadenylation. Moreover, possibly with dFMRP/FMRP and RNA transport factors such as Staufen (Chapter 4), dNab2/ZC3H14 might act to maintain translationally repressed state of mRNAs while they are transported along the axons.

The fact that we find nuclear isoforms of ZC3H14 (isoforms 1-3) in both the nucleus and cytoplasm is not unprecedented. Studies of the budding yeast orthologue Nab2 show that Nab2 shuttles between nucleus and cytoplasm in a transcription-dependent manner (8). In addition, Nab2 bound mRNAs locate to the bud tip and are displaced by Kap104-dependent mechanism (9). Although dNab2 locates to the nucleus at steady-state, subcellular distribution in neurons might mimic that of ZC3H14. In fact, we have preliminary evidence suggesting that dNab2 localizes to axonal granules in cultured *Drosophila* neurons (S.M. Kelly, unpublished data).

A growing literature now provides evidence for unexpected roles of nuclear Pabs in the cytoplasm and reciprocally, cytoplasmic Pabs in the nucleus (10). One specific example is the *Drosophila* nuclear Pab, Pabp2. In addition to its canonical nuclear role in poly(A) tail lengthening in nuclei of somatic tissues, Pabp2 also regulates poly(A) tails of specific cytoplasmic transcripts important in ooctyes and early embryos, as mutations in *Pabp2* cause elongated poly(A) tails of key maternal mRNAs correlated with early developmental arrest (11). Our data suggest that dNab2 might also play a dynamic role in controlling poly(A) tail length throughout development, as *dNab2* mutant embryos and larvae show decreased poly(A) tail length compared to *p.ex* controls (Figure 5.2). In contrast to a role in limiting bulk poly(A) tail length in adult tissues (Chapter 2), this

finding suggests that dNab2 has an opposing role in stimulating polyadenylation during early development. How the function of dNab2 switches between a positive regulator of polyadenylation in early embryonic and larval stages of development to a negative regulator of polyadenylation in the adult remains to be defined.

Remaining questions

From my dissertation work, several questions arise. The first is, are hyperadenylated poly(A) RNAs biologically relevant? In other words, can these hyperadenylated transcripts explain the defects seen in the development of *dNab2* mutants and/or the neurological defects seen in NS-ARID patients? In our *dNab2* mutant flies, we have clear evidence of hyperadenylation of poly(A) RNAs. However, we do not formerly know whether patient cells exhibit increased poly(A) tail length. Preliminary data revealed that in Family-1 lymphoblast cell line, which is null for isoforms 1-3 but has isoform 4 intact, there is no change in bulk poly(A) tail length compared to control (S.W. Leung, unpublished data). Whether this is due to the presence of isoform 4 remains to be tested. Transcripts with shorter poly(A) tails are targeted for rapid degradation by nuclear quality control pathways such as the exosome. However, what happens to transcripts with poly(A) tails that extend beyond the canonical length, termed hyperadenylated, is unclear. One hypothesis would be that hyperadenylated transcripts could be more stable and thus, be translated to a greater extent. Alternatively, hyperadenylated transcripts could be subject to degradation through as yet unidentified degradation machinery. Yet, the answer to this puzzling question is not so straightforward.

Mutations in several genes, in addition to Nab2/dNab2, lead to an increase in bulk poly(A) tail length. mRNA export mutants in budding yeast show hyperadenylation of transcripts (12, 13). In *Drosophila*, mutants in the deadenylation machinery lead to hyperadenylated transcripts (14). In addition, overexpression of poly(A) polymerase results in elongated poly(A) tails (15). Depending upon the developmental context, these hyperadenylated transcripts were shown to either induce or not induce translation (15). Despite these numerous reports, the functional consequences of hyperadenylated transcripts have not been studied. In explaining the neurological phenotypes seen in patients, we can speculate that likely consequences of hyperadenylated mRNAs could include altered transcript stability, titration of critical poly(A) RNA-binding proteins and/or bypass of cytoplasmic polyadenylation necessary for activity-dependent translation of neuronal mRNAs. Individually or in combination, these defects could disrupt spatiotemporal control of gene expression needed for development of the nervous system and higher order brain function.

The second question is, if both Nab2 and dNab2 are essential, why do patients display only neurological defect? Deletion of Nab2 in budding yeast is lethal (16) and germline and zygotic loss of dNab2 results in embryonic lethality (Chapter 2). Yet, mutations in *ZC3H14* lead to non-syndromic neurological deficit. One possibility is that human mutations in *ZC3H14* are hypomorphic rather than null mutations. Family-1 mutation affects expression of isoforms 1-3 but presumably isoform 4 is present in these patients (Chapter 2). Moreover, although Family-2 mutation affects all *ZC3H14* splice variants and patients display a more severe intellectual disability than Family-1 (Chapter 2), the exact molecular defect is not yet known as cells from these patients in rural Iran

are not yet available for analysis. Alternatively, we cannot rule out the possibility that a protein that is functionally redundant with ZC3H14 in higher eukaryotes exists; however, the human genome does not encode any apparent sequence orthologues of ZC3H14. Future studies exploiting mammalian model systems will be required to address the functional requirements for ZC3H14 as they relate to human intellectual disability.

Future directions

The molecular mechanism of dNab2/ZC3H14

In order to understand the molecular mechanism of dNab2/ZC3H14, we should test whether dNab2/ZC3H14-mediated phenotypes are due to functional consequences of hyperadenylated transcripts or by misregulation of specific target mRNAs that control neuronal function. First, we can test whether hyperadenylated mRNAs contribute to neuronal phenotypes by creating a transgenic fly expressing a UAS-reporter construct (e.g. GFP) with a stretch of abnormally lengthened adenosines at the 3'-end. This transgene could be driven by tissue-specific *Gal4* lines to express hyperadenylated GFP transgene in the tissue of choice. If hyperadenylated transcripts are causative, flies should have a phenotype similar to *dNab2* mutants. As mentioned above, one likely consequence of hyperadenylated mRNAs include titration of critical poly(A)-binding proteins (Pabs). To test this hypothesis, we can examine the overall levels of Pabs (e.g. PABPN1) by western blot analysis. If these Pabs are bound to these abnormally hyperadenylated reporter construct and sequestered away from their normal functions, the overall levels of Pabs should be decreased. In addition, the amount of Pabs bound to hyperadenylated GFP mRNAs could be indirectly assayed by RNA-immunoprecipitation. If Pabs bind to the

long stretches of adenosines on the reporter, the amount of RNA that is pulled down should exceed that of a transgenic control reporter with normal length of adenosines.

Another possibility is that upon loss of dNab2/ZC3H14, target mRNAs could be misregulated. To test this hypothesis, a direct approach such as RNA-immunoprecipitation should be used. Either endogenous dNab2/ZC3H14 or overexpression constructs of dNab2/ZC3H14 can be used to pull down the protein in fly heads or primary rat hippocampal cultures, and RNAs that are bound to the protein should be analyzed by either microarray or pyrosequencing. This approach will allow us to find direct mRNA targets that are bound to dNab2/ZC3H14. Upon validation of certain target mRNAs, these transcripts could be tested for their effects on neuronal function regulated by dNab2/ZC3H14 in flies or in mammalian models.

Placing ZC3H14 in the spectrum of RNA-binding proteins causing intellectual disability

Many RNA-binding proteins are associated with intellectual disability (ID) (17). Among those proteins, FMRP (Fragile X mental retardation protein) is best studied (18). In the course of this dissertation, we uncovered an interesting link between *dNab2/ZC3H14* and *dFmr1/Fmr1*. We show genetic evidence that *dNab2* and *dFmr1* act in the similar genetic pathways to control locomotor activity in neurons (Chapter 4). In addition, in *dNab2* null heads, we find that *dFmr1* transcripts are downregulated and show differential extractability of dFMRP in RIPA buffer while the total protein level is unchanged (Chapter 4). Based on these data, we speculate on different models that could act independently or in combination: 1) dNab2 and dFMRP work cooperatively, together in a complex; 2) dNab2 and dFMRP work in parallel pathways to control neuronal

function; or 3) *dFmr1* is downstream target of dNab2. In order to test these models in the future, co-immunoprecipitation experiments should be performed to determine the direct interaction between dNab2 and dFMRP. In addition, RNA-immunoprecipitation and transcript-specific poly(A) tail length assay (LM-PAT) should be performed to address whether dNab2 regulates *dFmr1* mRNAs directly.

Since there is less RIPA-soluble dFMRP in *dNab2* mutants and dNab2/ZC3H14 is found in the axonal cytoplasm (Chapter 4), it is also tempting to speculate that dNab2 might be required for targeting dFMRP-containing RNA granules to the dendrites. In the absence of dNab2, dFMRP might not be properly targeted and confer differential solubility. Moreover, there could exist a feedback mechanism whereby dFMRP might autoregulate its own mRNAs; therefore, in the absence of dNab2, mistargeting of dFMRP might explain the downregulation of dFmr1 transcript levels with no change in total protein. In the future, we should test whether this model is correct by determining the requirement of dNab2/ZC3H14 in proper transport/targeting of dFMRP/FMRP-containing RNA granules in cultured neurons. Cell biological experiments in combination with live cell imaging will allow for better understanding of the role dNab2/ZC3H14 plays in RNA transport.

Due to the interesting biological link between *FMRP* and *ZC3H14* and the fact that mutations in both of these genes lead to intellectual disability, both *FMRP* and *ZC3H14* could possibly be involved in common signaling pathways that are important for normal brain function. However, even if this is the case, mutations in other RNA-binding proteins suggest that there are other signaling pathways that are affected and give rise to intellectual disability. Future studies examining the functions of these RNA-binding

proteins and different pathways that affect normal cognitive function will together allow for better understanding of the complex networks underlying brain function.

Drug discovery to ameliorate intellectual disability

Our findings indicating that dNab2 is required for neuronal function suggest that potential signaling pathways could be altered in *dNab2* mutants and may provide insight into parallel mechanisms that might contribute to neurological dysfunction in ZC3H14-mediated NS-ARID. Future studies following up on genetic interactors of *dNab2* (Chapter 4) and dNab2-regulated mRNA transcripts (Appendix) would allow for discovery of signaling pathways and networks that are important for neuronal function. Once dNab2 is placed within such a signaling pathway, a small molecule screen identifying potential inhibitor or agonist that would affect dNab2-mediated function could be useful. Since depletion of dNab2 in the *Drosophila* MB neurons leads to defects in short-term memory (Chapter 3), the courtship suppression assay could be used to search for small molecules that could rescue the courtship-mediated memory. Such studies performed on fly models of Fragile X syndrome confirm the possibility of such an approach (19). In addition, other learning paradigms such as olfactory discrimination assay (20, 21) could be used for dissecting both short-term and long-term memory deficits in flies with altered dNab2 function and for identification of small molecules that modify the learning and memory phenotypes. Furthermore, with our current efforts in creating a mouse model for ZC3H14, characterization of both cellular and behavioral phenotypes will allow for better understanding of the molecular pathologies present in human patients, as alternative splice variants of *ZC3H14* are present only in mammals.

In conclusion, both fly and future mouse models will serve as excellent tools in modeling NS-ARID and provide insight into understanding the complex networks underlying normal brain function.

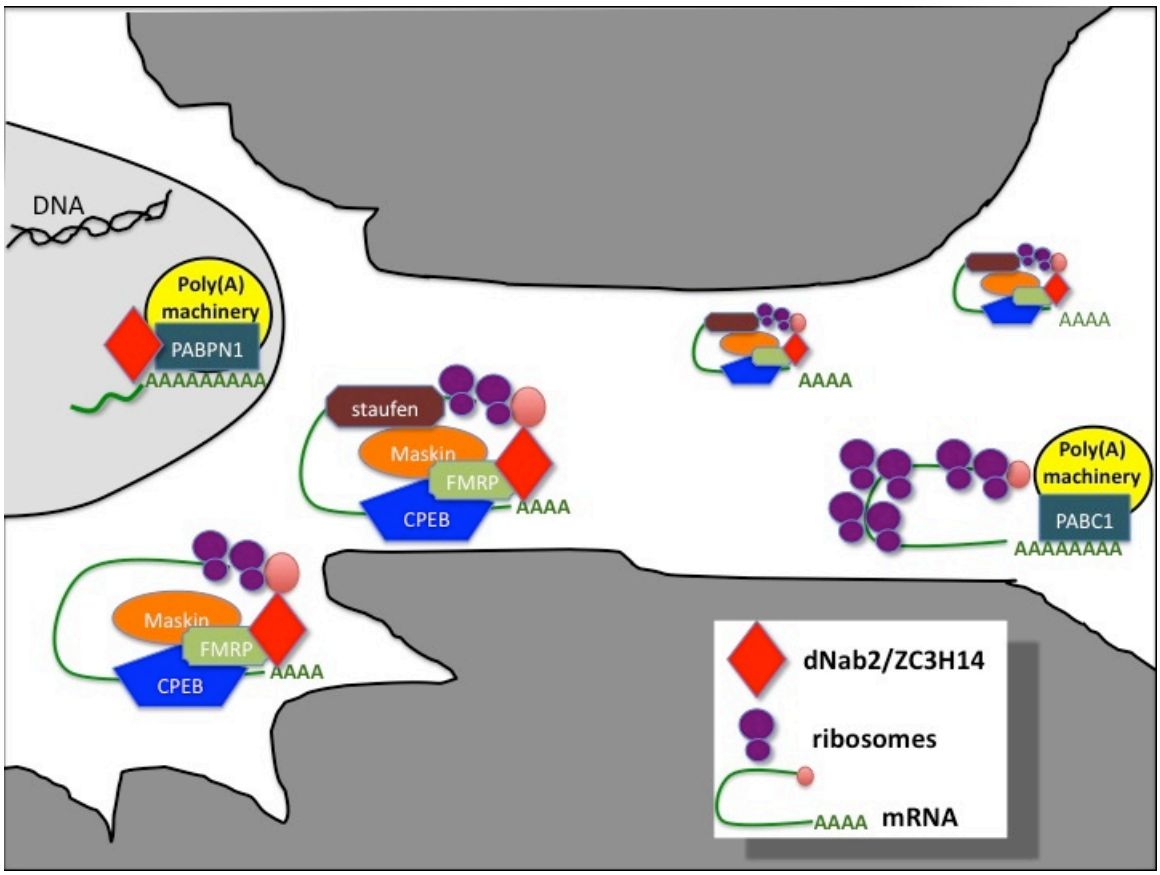


Figure 5.1. A model of dNab2/ZC3H14 function in neurons. In the nucleus, dNab2/ZC3H14 binds to poly(A) tails of mRNAs, functioning to antagonize PABPN1 in promoting polyadenylation. Upon nuclear export of mRNAs into the cytoplasm, PABPN1 is possibly displaced from the poly(A) tails but dNab2/ZC3H14 is still bound to its target transcripts. dNab2/ZC3H14 associates with translation factors (e.g. ribosomes and cap-binding eIF4E) and other translational regulators (e.g. dFMRP/FMRP). This model speculates that dNab2/ZC3H14 might also interact with CPEB and Maskin in order to maintain the translationally dormant state of mRNAs. Since dNab2/ZC3H14 acts to limit poly(A) tails (Chapter 2) and the footprint of yeast Nab2 has been shown to bind to ~10 adenosines (unpublished data), it is likely that dNab2/ZC3H14 binds to the short

poly(A) tails that are maintained in the CPEB-Maskin/PARN complex (22). Moreover, possibly with RNA transport factors, such as dFMRP/FMRP and Staufen, dNab2/ZC3H14 might act to properly transport RNA granules along the axon and finally target these complexes to the dendrites for translation. Upon a signaling event, translation occurs and dNab2/ZC3H14 could be replaced by a cytoplasmic Pab, such as PABC1 which then binds to elongated poly(A) tails and promote translation.

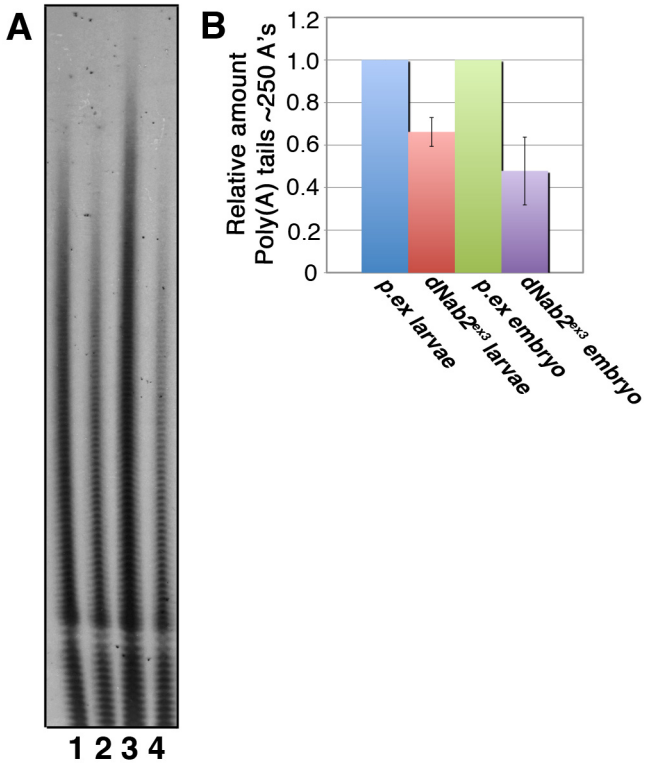


Figure 5.2. dNab2 acts to lengthen bulk poly(A) tail length during early development. (**A**) Bulk poly(A) tail length measurements in whole animals from 1) *p.ex* control larvae 2) *dNab2*^{ex3} larvae 3) *p.ex* control embryo 4) *dNab2*^{ex3} maternal and zygotic null embryo. (**B**) Densitometric quantification of poly(A) tracts (Image J) of ~250 nucleotides (nt) normalized to poly(A) tracts of ~100 nt for each genotype. Each *p.ex* controls were set to 1.0. $p<0.05$ (two-tailed *t*-test) when each *p.ex* controls were compared to *dNab2* mutants. n=4. Error bars=s.d.

References

1. Leung SW, *et al*. (2009) Splice variants of the human ZC3H14 gene generate multiple isoforms of a zinc finger polyadenosine RNA binding protein. *Gene* 439(1-2):71-78.
2. Kelly SM, *et al*. (2007) Recognition of polyadenosine RNA by zinc finger proteins. *Proc Natl Acad Sci U S A* 104(30):12306-12311.
3. Heisenberg M (2003) Mushroom body memoir: from maps to models. *Nat Rev Neurosci* 4(4):266-275.
4. Laggerbauer B, Ostareck D, Keidel EM, Ostareck-Lederer A, & Fischer U (2001) Evidence that fragile X mental retardation protein is a negative regulator of translation. *Hum Mol Genet* 10(4):329-338.
5. Li Z, *et al*. (2001) The fragile X mental retardation protein inhibits translation via interacting with mRNA. *Nucleic Acids Res* 29(11):2276-2283.
6. Wahle E (1991) A novel poly(A)-binding protein acts as a specificity factor in the second phase of messenger RNA polyadenylation. *Cell* 66(4):759-768.
7. Bienroth S, Keller W, & Wahle E (1993) Assembly of a processive messenger RNA polyadenylation complex. *EMBO J* 12(2):585-594.
8. Green DM, *et al*. (2002) Nab2p is required for poly(A) RNA export in Saccharomyces cerevisiae and is regulated by arginine methylation via Hmt1p. *J Biol Chem* 277(10):7752-7760.
9. van den Bogaart G, Meinema AC, Krasnikov V, Veenhoff LM, & Poolman B (2009) Nuclear transport factor directs localization of protein synthesis during mitosis. *Nat Cell Biol* 11(3):350-356.

10. Lemay JF, Lemieux C, St-Andre O, & Bachand F (2010) Crossing the borders: poly(A)-binding proteins working on both sides of the fence. *RNA Biol* 7(3):291-295.

11. Benoit B, *et al*. (2005) An essential cytoplasmic function for the nuclear poly(A) binding protein, PABP2, in poly(A) tail length control and early development in Drosophila. *Dev Cell* 9(4):511-522.

12. Hilleren P & Parker R (2001) Defects in the mRNA export factors Rat7p, Gle1p, Mex67p, and Rat8p cause hyperadenylation during 3'-end formation of nascent transcripts. *RNA* 7(5):753-764.

13. Jensen TH, Patricio K, McCarthy T, & Rosbash M (2001) A block to mRNA nuclear export in S. cerevisiae leads to hyperadenylation of transcripts that accumulate at the site of transcription. *Mol Cell* 7(4):887-898.

14. Temme C, Zaessinger S, Meyer S, Simonelig M, & Wahle E (2004) A complex containing the CCR4 and CAF1 proteins is involved in mRNA deadenylation in Drosophila. *EMBO J* 23(14):2862-2871.

15. Juge F, Zaessinger S, Temme C, Wahle E, & Simonelig M (2002) Control of poly(A) polymerase level is essential to cytoplasmic polyadenylation and early development in Drosophila. *EMBO J* 21(23):6603-6613.

16. Anderson JT, Wilson SM, Datar KV, & Swanson MS (1993) NAB2: a yeast nuclear polyadenylated RNA-binding protein essential for cell viability. *Mol Cell Biol* 13(5):2730-2741.

17. Lukong KE, Chang KW, Khandjian EW, & Richard S (2008) RNA-binding proteins in human genetic disease. *Trends Genet* 24(8):416-425.

18. Bassell GJ & Warren ST (2008) Fragile X syndrome: loss of local mRNA regulation alters synaptic development and function. *Neuron* 60(2):201-214.

19. McBride SM, *et al*. (2005) Pharmacological rescue of synaptic plasticity, courtship behavior, and mushroom body defects in a Drosophila model of fragile X syndrome. *Neuron* 45(5):753-764.

20. Bitterman ME, Menzel R, Fietz A, & Schafer S (1983) Classical conditioning of proboscis extension in honeybees (Apis mellifera). *J Comp Psychol* 97(2):107-119.

21. Tully T & Quinn WG (1985) Classical conditioning and retention in normal and mutant Drosophila melanogaster. *J Comp Physiol A* 157(2):263-277.

22. Richter JD (2007) CPEB: a life in translation. *Trends Biochem Sci* 32(6):279-285.

APPENDIX Part 1

Knockdown of *dNab2* affects transcript abundance of specific mRNAs in S2 cells.

Todd Bradley, ChangHui Pak, Marco Blanchette, Kenneth H. Moberg, Anita H. Corbett

S2 cell knockdown, high-throughput sequencing and initial data analysis were performed by Todd Bradley and Marco Blanchette at Stower's Institute for Medical Research (Kansas City, MO). CP performed analysis of knockdown by qRT-PCR.

Brief summary

In order to identify dNab2-specific target transcripts, we utilized knockdown of *dNab2* in the haemocyte-driven *Drosophila* cell line, S2 cells (1), and subsequently analyzed relative abundance in transcript levels on a genome-wide scale by high-throughput sequencing. Upon treatment of S2 cells with double-stranded RNAs (dsRNAs) targeted to *dNab2* transcript, we observed greater than 50% knockdown in both protein and mRNA levels compared non-specific control (Figure A.1). After achieving *dNab2*-specific knockdown, we isolated total RNA and analyzed relative transcript abundance by high-throughput sequencing.

By this method, we have identified ~90 mRNAs that decline in abundance in S2 cells depleted of *dNab2* by RNAi (Table A.1). Interestingly, only 2 mRNAs were upregulated upon *dNab2* knockdown (Table A.2). Although dNab2 binds polyadenosine RNA tracts *in vitro* (Chapter 2), our data indicate that dNab2 appears to have transcript-specific regulatory roles *in vivo*. This result is consistent with our current findings that knockdown of the human *ZC3H14* in MCF7 cell line results in changes in transcript levels of ~200 mRNAs (C.P. Wigington, unpublished data).

A number of the dNab2-regulated mRNAs encode proteins with predicted neuronal function. *Gef64C* encodes a Rho GTP exchange factor, which has previously been reported to be involved promoting in axon guidance in *Drosophila* embryos (2). In addition, *Neuroglian*, which encodes a cell-adhesion molecule (CAM), functions in synapse formation and axon pathfinding (3, 4). Interestingly, mutations in the human gene, *L1-CAM*, result in a broad spectrum of neurological phenotypes (the CRASH syndrome), including intellectual disability (5, 6). Downregulation of these transcripts

could potentially explain the neurological phenotypes that we observe in both *dNab2* mutants and human patients (Chapter 2). Further validation of the data in the fly central nervous system would allow us to better understand whether these mRNAs are putative dNab2 targets and how they might contribute to dNab2-mediated phenotypes.

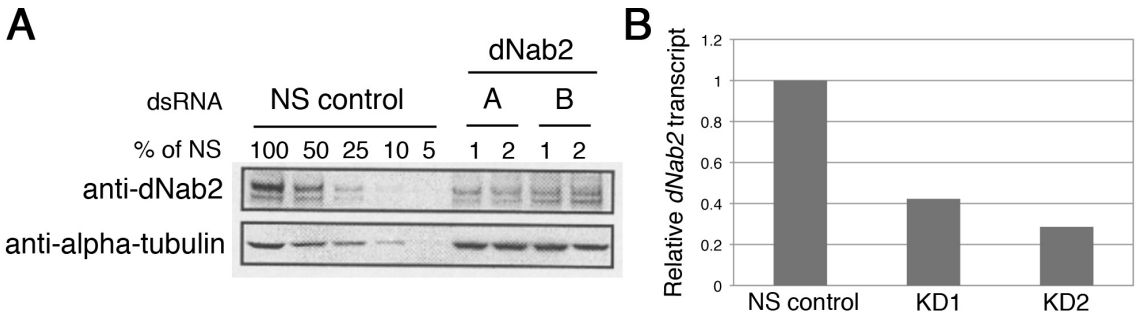


Figure A.1. Knockdown of *dNab2* expression. (**A**) Immunoblot detecting both dNab2 and α-tubulin in S2 cells that were treated with double-stranded RNAs (dsRNAs) targeted to either a non-specific control (NS control) or *dNab2* transcript. (**B**) Quantitative real-time RT-PCR analysis detecting *dNab2* transcripts for NS control (set as 1.0) compared to two independent *dNab2* knockdown experiments (KD1 and KD2). *β-tubulin* is an internal control. Representative of a single experiment.

Table A.1. Genes downregulated upon *dNab2* knockdown in S2 cells.

Rank	Symbol	Full name	Log2 fold change	q-value (BH adujsted)	Tags in experiment	Tags in control
1	CG33465		-5.16	7.47e-04	0	9
2	CAH2	Carbonic anhydrase 2	-3.44	8.16e-04	2	11
3	CG31028		-3.21	2.11e-04	3	14
4	CG6834		-3.10	4.74e-04	3	13
5	CG10814		-3.02	1.03e-12	12	49
6	tor	torso	-2.91	2.98e-05	5	19
7	Rya-r44F	Ryanodine receptor 44F	-2.75	1.45e-04	5	17
8	CG17650		-2.40	2.32e-05	9	24
9	Cyp4ad1	Cyp4ad1	-2.16	8.08e-04	8	18
10	CG30377		-2.14	4.10e-04	9	20
11	CG8420		-2.06	2.09e-07	20	42
12	CG7447		-1.93	1.24e-08	28	54
13	CG15412		-1.93	1.02e-04	14	27
14	CG2680		-1.93	2.03e-04	13	25
15	CG10352		-1.88	3.04e-06	21	39
16	CG4259		-1.74	3.72e-18	81	137
17	CG8369		-1.69	9.98e-07	30	49
18	lectin-33A	lectin-33A	-1.63	6.88e-14	73	114
19	GstD2	Glutathione S transferase D2	-1.62	3.93e-18	98	152
20	CG6499		-1.61	2.10e-10	55	85
21	CG14302		-1.61	6.35e-27	149	230
22	Tsf1	Transferrin 1	-1.58	1.66e-13	77	116
23	modSP	modular serine protease	-1.57	3.34e-14	82	123
24	CG10433		-1.57	4.55e-04	20	30
25	Hsp67Bb	Heat shock gene 67Bb	-1.55	4.52e-05	27	40
26	CG16959		-1.55	4.21e-04	21	31
27	CG12766		-1.51	1.67e-05	32	46
28	CG12520		-1.49	1.55e-25	173	245
29	GstD10	Glutathione S transferase D10	-1.47	2.56e-06	40	56
30	Kaz1-ORFB	Kaz1-ORFB	-1.46	2.65e-05	33	46
31	CG7724		-1.44	1.91e-06	43	59
32	CG13565		-1.42	1.42e-06	46	62
33	Vdup1	Vitamin D[[3]] up-regulated protein 1	-1.39	1.66e-20	163	216
34	CG13116		-1.37	2.99e-07	56	73
35	Prx2540-2	Peroxiredoxin 2540	-1.36	2.27e-05	40	52
36	Msr-110	Msr-110	-1.36	1.45e-16	138	179
37	Unc-89	Unc-89	-1.35	1.84e-05	42	54
38	Mlc2	Myosin light chain 2	-1.35	1.96e-07	60	77
39	Gef64C	Guanine nucleotide exchange factor GEF64C	-1.34	2.50e-04	32	41
40	CG5142		-1.33	9.55e-07	55	70
41	Snmp	Sensory neuron membrane	-1.30	3.35e-06	53	66

		protein				
42	Cad96Cb	Cad96Cb	-1.29	8.06e-09	80	99
43	alphaPS4	alphaPS4	-1.29	2.70e-24	234	289
44	l(2)03659	lethal (2) 03659	-1.29	1.86e-08	77	95
45	CG10924		-1.28	1.18e-04	39	48
46	GstE2	Glutathione S transferase E2	-1.28	3.78e-14	135	166
47	GstD5	Glutathione S transferase D5	-1.27	5.30e-49	494	603
48	Hsp67Ba	Heat shock gene 67Ba	-1.27	3.31e-05	46	56
49	Ance	Angiotensin converting enzyme	-1.27	3.98e-06	56	68
50	CG14567		-1.26	2.17e-04	38	46
51	CG6830		-1.25	5.15e-10	100	120
52	CG8398		-1.24	5.44e-05	46	55
53	CG34231		-1.24	4.42e-04	36	43
54	CG32625		-1.24	1.70e-05	52	62
55	Cyp4p1	Cytochrome P450-4p1	-1.22	1.38e-09	100	118
56	CG4301		-1.21	2.63e-06	64	75
57	Reg-2	Rhythmically expressed gene 2	-1.21	7.66e-06	59	69
58	Sp212	Serine-peptidase 212	-1.21	2.25e-04	42	49
59	betaInt-nu	beta[nu] integrin	-1.20	8.85e-27	306	356
60	CG16700		-1.20	2.02e-04	43	50
61	CG1090		-1.20	6.35e-04	37	43
62	CG6040		-1.19	4.76e-11	125	144
63	CG18622		-1.18	6.32e-08	88	101
64	CG17270		-1.16	9.14e-04	38	43
65	CG3940		-1.15	2.42e-18	231	259
66	CG9813		-1.15	1.18e-11	145	162
67	CG13102		-1.13	1.40e-05	66	73
68	CG13924		-1.12	9.09e-04	42	46
69	CG9573		-1.11	2.21e-05	66	72
70	Cht2	Chitinase 2	-1.10	7.48e-14	191	207
71	TfIIA-S-2	TfIIA-S-2	-1.10	4.48e-04	48	52
72	CG9270		-1.10	6.89e-15	207	224
73	mthl14	methuselah-like 14	-1.09	4.76e-20	287	309
74	CG5381		-1.07	5.76e-16	235	250
75	fon	fondue	-1.07	4.86e-04	51	54
76	CG11378		-1.07	3.74e-05	69	73
77	CG17124		-1.06	2.49e-14	215	227
78	CG30502		-1.06	3.04e-04	55	58
79	Cyp6a8	Cytochrome P450-6a8	-1.06	1.35e-26	415	436
80	Npc2b	Niemann-Pick type C-2b	-1.05	1.74e-12	191	200
81	CG10232		-1.05	7.45e-17	263	275
82	CG4804		-1.04	8.90e-14	219	227
83	CG5955		-1.04	1.89e-07	113	117
84	CG31116		-1.03	1.68e-16	267	276
85	CG12206		-1.03	1.14e-76	1300	1337
86	Nrg	Neuroglian	-1.02	6.11e-17	286	292

87	CG14545		-1.01	2.60e-17	293	299
88	stv	starvin	-1.01	8.04e-04	54	55
89	CG7702		-1.01	2.63e-10	171	174
90	dtr	defective transmitter release	-1.01	1.20e-04	70	71
91	ry	rosy	-1.00	4.24e-22	387	392
92	CG8136		-1.00	4.04e-11	190	192

Genes significantly affected with q-value (Benjamini-Hochberg adjusted) <= 0.001 that are affected at least by 2 fold are shown. 'CG' indicates uncharacterized gene.

Table A.2. Genes upregulated upon *dNab2* knockdown in S2 cells.

Rank	Symbol	Full name	Log2 fold change	q-value (BH adujsted)	Tags in experiment	Tags in control
1	Ser	Serrate	1.80	4.29e-20	289	42
2	Pxn	Peroxidasin	1.73	5.71e-183	2839	432

Genes significantly affected with q-value (Benjamini-Hochberg adjusted) <= 0.001 that are affected at least by 2 fold are shown. 'CG' indicates uncharacterized gene.

Experimental procedures

Standard procedures were used to achieve knockdown in S2 cells (7). Libraries for Illumina sequencing were created using standard protocols (Illumina mRNA seq kit). High-throughput sequencing was performed at the core facility at Stower's Institute for Medical Research (Kansas City, MO). For all other methods, refer to Experimental procedures in Chapter 2.

APPENDIX Part 2

Introduction to *Drosophila* genetic tools

In this section, the versatile gene-overexpression system, the GAL4-upstream activating sequence (UAS) system, and newly generated techniques and genetic screen methods based on GAL4-UAS system, and P-element mediated mutagenesis will be described.

The GAL4-UAS system

The most widely used technique in *Drosophila* to achieve spatially controlled gene expression is the GAL4-UAS system that originates from yeast (8). As shown in Figure A.2, a *P*-element carrying the GAL4 transcriptional activator is randomly distributed throughout the genome, allowing the expression of GAL4 under the control of endogenous promoters. In a different parental line, target transgene is cloned downstream of the UAS element. When GAL4 is introduced by crossing the two parental lines, target transgene is expressed in a tissue-specific pattern as the GAL4 activator.

Based on this over-expression system, variations of techniques have been manipulated to achieve both spatial and temporal expression of transgenes (9). Among them is TARGET (temporal and regional gene expression targeting) system, which makes use of a temperature-sensitive GAL80 (ts-GAL80) protein that acts as a repressor for GAL4 (10, 11). When introduced in flies, tsGAL80 represses GAL4 expression at 19°C and when flies are introduced to higher temperatures beginning at 30°C, it allows optimal de-repression of GAL4 expression (Figure A.3). Using this system, one can tightly control transgene expression for both time and space throughout development and beyond. Another great tool emerged after the discovery of RNA interference (RNAi) (12). In combination with GAL4-UAS system, double-stranded RNAs can be expressed in both spatial and temporal-controlled manner to knockdown a target gene (13). This method allows gene inactivation for both maternal and zygotic transcripts, thereby bypassing the need for germ-line clone analysis.

Enhancer-Promoter (EP) over-expression screens

One advantage of the *Drosophila* as a model system is the availability of established genetic screens. The Enhancer-Promoter (EP) over-expression system allows conditional expression of *Drosophila* genes based on their proximity to randomly inserted *P*-elements within the genome (14). These inserted *P*-elements each carry multiple copies of the UAS recognition site for the GAL4 transcription factor. When a library of such insertions scattered randomly throughout the genome is combined with a transgene that expresses GAL4 in a particular spatial or temporal pattern, genes near the insertion will be ectopically expressed in the same pattern. By using GAL4 'drivers' with known expression patterns, the EP system thus allows one to screen through an 'EP library' for genes that elicit a phenotype in the tissue of interest.

EP screens allow one to find genes and pathways that interact with the gene of interest. For example, an EP collection, covering 7000 *P*-elements, was used to screen modifiers of neurodegeneration caused by polyglutamine tracts in the fly eye to model Huntington's disease as well as other polygluatmine diseases (15). This study along with many other recent studies are identifying modifiers of polyglutamine toxicity using *Drosophila*, aiming to enhance our understanding of trinucleotide repeat instability in disease (16).

P-element excision mutagenesis

In a large-scale effort to achieve gene disruption, several systematic *P*-element gene disruption projects have generated thousands of stocks that harbor a single *P*-element construct inserted at a known location in the genome (17). Most often *P*-element

insertions cause hypomorphic mutations of genes nearby where they are inserted; however, it is not uncommon to find *P*-element insertions to fail to disrupt any gene. In addition, it is desirable to obtain complete loss-of-function and not hypomorphic alleles to study gene function. In order to achieve this, a mutagenesis strategy called 'imprecise excision' can be used, in which *P*-element insertions can be used to generate new mutations close to the original insertion (18).

The basis of this method is to generate many independent excisions of the *P*-element and to screen for those that remove some of the genomic sequence flanking the insertion site (Figure A.4). In an imprecise excision screen, flies that harbor both the *P*-element insertion and a chromosome that expresses high level of *P*-transposase are generated by an appropriate crossing scheme. Once the *P*-transposase catalyzes the excision reaction, the *P*-element is cut out of the donor site, leaving a double-stranded break. This site is normally repaired by DNA-repair mechanisms; however, the alternative outcome of the repair process is most useful, whereby enlarged deletions that extend in either direction are isolated. Because *P*-element insertions often are present near the promoter regions, one can obtain deletions that partially or completely disrupt gene function by removing coding sequences.

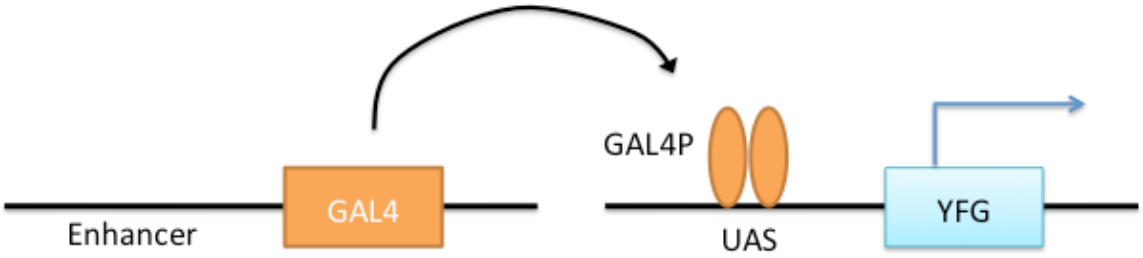


Figure A.2. The GAL4-UAS system. The yeast transcriptional activator GAL4 is expressed in a specific spatial pattern, driven by a defined promoter or an endogenous enhancer. The GAL4 protein (GAL4P) binds to the upstream activating sequence (UAS) and allows activation of transcription of your favorite gene (YFG) cloned downstream of the UAS. (Adapted from McGuire *et al*., 2004.)

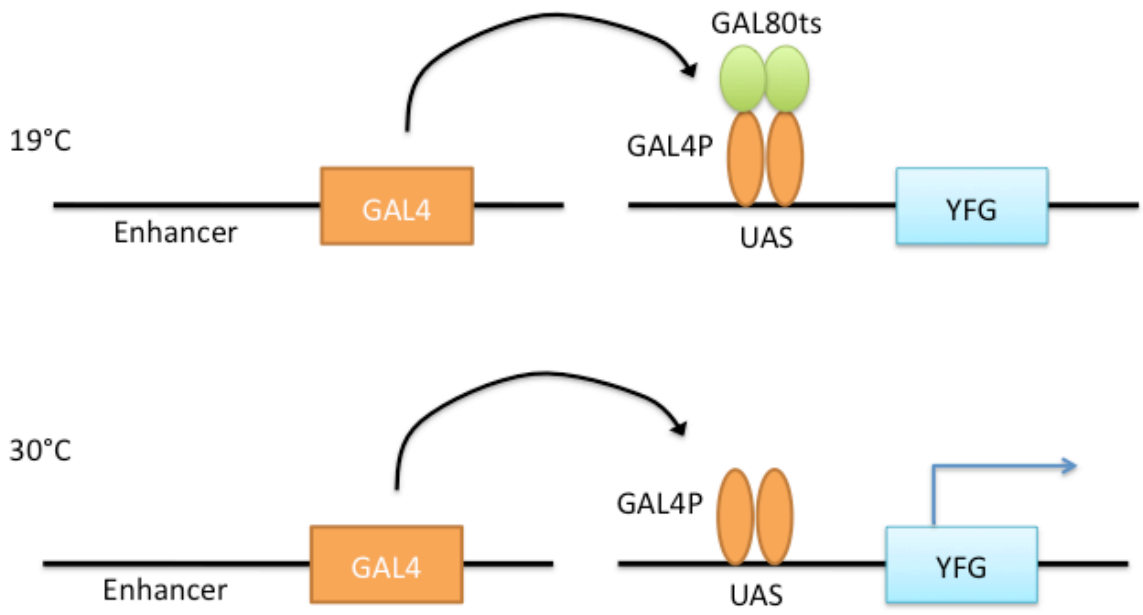


Figure A.3. TARGET (temporal and regional gene expression targeting) system. The conventional GAL4-UAS system is modified such that it is conditionally regulated by a temperature sensitive allele of GAL80 (GAL80ts) to provide spatial and temporal control of gene expression. At 19°C, YFG is repressed due to the binding of GAL80ts to GAL4P. With a temperature shift to 30°C, this repression is relieved, allowing for transcription and high level of expression of YFG in a specific tissue. (Adapted from McGuire *et al*., 2004.)

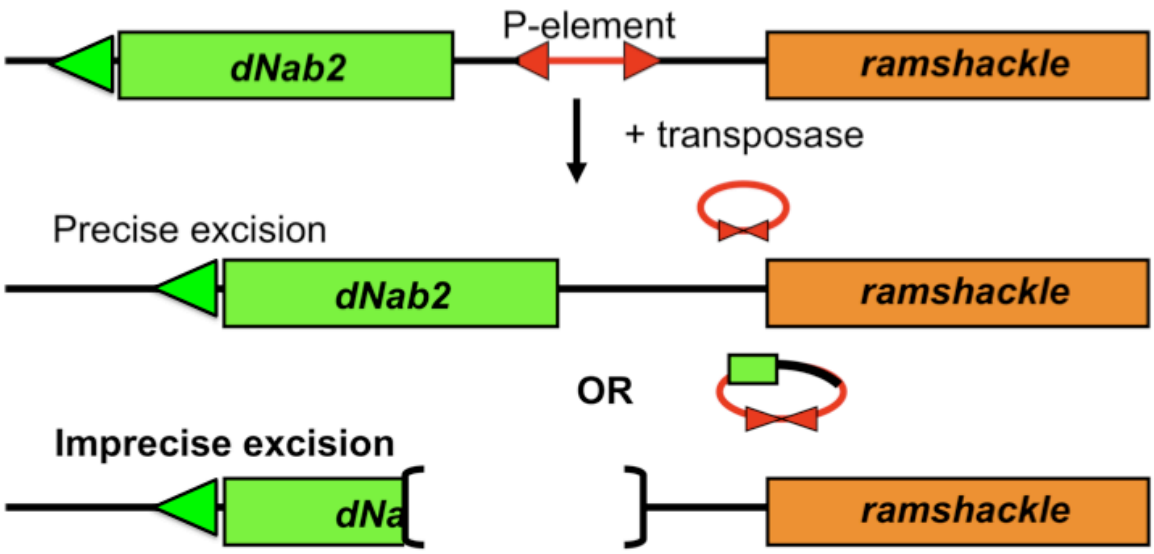


Figure A.4. A schematic of imprecise excision-based mutagenesis. A *P*-element insertion in the genome is mobilized by introducing a source of *P*-transposase, which catalyzes the excision of the element. Generally, precise excision occurs such that the *P*-element is precisely removed, restoring the normal genomic sequence. In rare instances, however, small deletions can be isolated whereby the *P*-element takes part of the flanking genomic sequences with it (imprecise excision). The genomic locus of *dNab2* is shown with the location of the neighboring gene, *ramschackle*.

References

1. Schneider I (1972) Cell lines derived from late embryonic stages of Drosophila melanogaster. *J Embryol Exp Morphol* 27(2):353-365.
2. Bashaw GJ, Hu H, Nobes CD, & Goodman CS (2001) A novel Dbl family RhoGEF promotes Rho-dependent axon attraction to the central nervous system midline in Drosophila and overcomes Robo repulsion. *J Cell Biol* 155(7):1117-1122.
3. Itoh K, *et al*. (2004) Brain development in mice lacking L1-L1 homophilic adhesion. *J Cell Biol* 165(1):145-154.
4. Hall SG & Bieber AJ (1997) Mutations in the Drosophila neuroglian cell adhesion molecule affect motor neuron pathfinding and peripheral nervous system patterning. *J Neurobiol* 32(3):325-340.
5. Bateman A, *et al*. (1996) Outline structure of the human L1 cell adhesion molecule and the sites where mutations cause neurological disorders. *EMBO J* 15(22):6050-6059.
6. Fransen E, Van Camp G, D'Hooge R, Vits L, & Willems PJ (1998) Genotype-phenotype correlation in L1 associated diseases. *J Med Genet* 35(5):399-404.
7. Blanchette M, Green RE, Brenner SE, & Rio DC (2005) Global analysis of positive and negative pre-mRNA splicing regulators in Drosophila. *Genes Dev* 19(11):1306-1314.
8. Brand AH & Perrimon N (1993) Targeted gene expression as a means of altering cell fates and generating dominant phenotypes. *Development* 118(2):401-415.

9. McGuire SE, Roman G, & Davis RL (2004) Gene expression systems in Drosophila: a synthesis of time and space. *Trends Genet* 20(8):384-391.

10. McGuire SE, Le PT, Osborn AJ, Matsumoto K, & Davis RL (2003) Spatiotemporal rescue of memory dysfunction in Drosophila. *Science* 302(5651):1765-1768.

11. McGuire SE, Mao Z, & Davis RL (2004) Spatiotemporal gene expression targeting with the TARGET and gene-switch systems in Drosophila. *Sci STKE* 2004(220):pl6.

12. Fire A, *et al*. (1998) Potent and specific genetic interference by double-stranded RNA in Caenorhabditis elegans. *Nature* 391(6669):806-811.

13. Lam G & Thummel CS (2000) Inducible expression of double-stranded RNA directs specific genetic interference in Drosophila. *Curr Biol* 10(16):957-963.

14. Rorth P (1996) A modular misexpression screen in Drosophila detecting tissue-specific phenotypes. *Proc Natl Acad Sci U S A* 93(22):12418-12422.

15. Kazemi-Esfarjani P & Benzer S (2000) Genetic suppression of polyglutamine toxicity in Drosophila. *Science* 287(5459):1837-1840.

16. Bilen J & Bonini NM (2005) Drosophila as a model for human neurodegenerative disease. *Annu Rev Genet* 39:153-171.

17. Spradling AC, *et al*. (1999) The Berkeley Drosophila Genome Project gene disruption project: Single P-element insertions mutating 25% of vital Drosophila genes. *Genetics* 153(1):135-177.

18. Voelker RA, *et al*. (1984) Frequent Imprecise Excision among Reversions of a P Element-Caused Lethal Mutation in Drosophila. *Genetics* 107(2):279-294.