Investigation of RNA-mediated pathogenic pathways in a *Drosophila* model of expanded repeat disease

A thesis submitted for the degree of Doctor of Philosophy, June 2010

Clare Louise van Eyk, B.Sc. (Hons.)

School of Molecular and Biomedical Science, Discipline of Genetics

The University of Adelaide

Table of Contents

Index of Figures and Tables

Chapter 1

Chapter 2

Chapter 3

Chapter 4

Chapter 5

Chapter 6

Chapter 7

Declaration

This work contains no material which has been accepted for the award of any other degree or diploma in any university or other tertiary institution and, to the best of my knowledge and belief, contains no material previously published or written by another person, except where due reference has been made in the text.

I give consent to this copy of my thesis, when deposited in the University Library, being available for loan and photocopying subject to the provisions of the Copyright Act 1968.

I also give permission for the digital version of my thesis to be made available on the web, via the University's digital research repository, the Library catalogue, the Australasian Digital Theses Program (ADTP) and also through web search engines, unless permission has been granted by the University to restrict access for a period of time.

Clare van Eyk

Acknowledgements

I would like to thank my supervisor, Rob Richards, for giving me support and advice as well as plenty of independence throughout this project. I am also very grateful to my co-supervisor, Louise O'Keefe, for the many helpful conversations in the fly room and for encouraging me to stay positive. Thanks also to the various people who have assisted in aspects of this project; particularly to Jo Milverton for performing microinjections and to Gareth Price for assistance with the microarray studies.

To all Richards lab and Genetics Discipline members, past and present, thanks for making the lab and the building such a fun and rewarding working environment. Special thanks to Saumya Samaraweera and Amanda Choo, for distracting me when I needed it most, and to Sonia Dayan for always taking the time to chat, even when you had a million things on the go.

I am also grateful to all of my family and friends who have given me love and support in too many ways to mention. In particular, I would like to thank my parents, Helen and Bernie van Eyk, who have always encouraged us to aim high in whatever we choose to do. Also to Simon Wells: for intellectual debate (and occasionally admitting that I'm right).

Abbreviations

℃: degrees Celsius

%: percentage

µA: microamps

µg: micrograms

µL: microlitre

A: Adenosine

ADAR: Adenosine deaminase acting on RNA

ADD1: Adducin 1

ALS: Amyotrophic lateral sclerosis

AMPA: α-amino-3-hydroxy-5-methyl-4-isoxazole propionateglutamate

AR: Androgen receptor

ATP: adenosine triphosphate

BDNF: Brain-derived neurotrophic factor

bp: base pairs

C: cytosine

cDNA: complementary DNA

CLCN-1: Chloride channel 1

CNS: central nervous system

CUG-BP: CUG binding protein

Cy: cyanine

CYFIP2: Cytoplasmic FMR1 interacting protein 2

da: daughterless

DAPI: 4'-6-Diamidino-2-phenylindole

DEPC: diethyl pyrocarbonate

DIGE: differential in-gel electrophoresis

DM: Myotonic dystrophy

DMF: dimethyl formamide

DMPK: Dystrophia myotonica protein kinase

DNA: deoxyribonucleic acid

DRPLA: Dentatorubral-pallidoluysian atrophy

dsRNA: double-stranded RNA

DTT: dithiothreitol

EDTA: ethylene diamine tetra-acetic acid

elav: embryonic lethal abnormal vision

emPAI: exponentially modified protein abundance index
ESI: electro-spray ionisation
FA: formic acid
FMR1: Fragile X mental retardation 1
FMRP: Fragile X mental retardation protein
FXTAS: Fragile X tremor-ataxia syndrome
G: guanosine
GABA: gamma-aminobutyric acid
GFP: green fluorescent protein
GluCl-α: Glutamate-gated chloride channel α
GluR-B: AMPA receptor subunit B
GMR: glass multimer reporter
GSK3: Glycogen synthase kinase 3
HD: Huntington's disease
HDL-2: Huntington's disease-like-2
hnRNP: Heterogenous ribonucleoprotein
Hr38: Hormone receptor-like in 38
Hts: Hu-li tai shao
HTT: Huntingtin
I: inosine
Insc: Inscuteable
IPTG: isopropyl β-D-1-thiogalactopyranoside
IR: Insulin receptor
JPH3: Junctophilin-3
kb: kilobase
kDa: kilodalton
KLHL1: Kelch-like 1
LB: Luria broth
M: Molar
Mbl: Muscleblind (*Drosophila)*
MBNL: Muscleblind-like
MEF: Myocyte enhancing factor
mg: milligrams
mGluRA: metabotropic glutamate receptor A
miRNA: microRNA

MJD: Machado Joseph disease

mL: millilitres

mM: millimolar

MQ: MilliQ™ purified water

mRNA: messenger RNA

MS: mass spectrometry

MS/MS: tandem mass spectrometry

MTMR1: Myotubularin-related protein 1

TOR: target of rapamycin

ng: nanograms

NGF: Nerve growth factor

NL IPG: non-linear immobilised pH gradient

NMDAR: N-methyl-D-aspartate receptor

NPC: Nuclear pore complex

dNTP: deoxyribonucleoside triphosphate

NUP: nucleoporin

NUR77: Nuclear receptor 77

OPMD: Oculopharyngeal muscular dystrophy

para: paralytic sodium channel

PBS: phosphate buffered saline

PBST: PBS + Tween

PKR: RNA regulated protein kinase

pmol: picomole

polyQ: polyglutamine

polyL: polyleucine

PP2A: Protein phosphatase 2A

PP2R2B: PP2A regulatory subunit 2B

PSF: Poly-pyrimidine-tract associated splicing factor

Q: glutamine

RISC: RNA-induced silencing complex

RNA: ribonucleic acid

RNAi: RNA interference

ROS: reactive oxygen species

Rp49: Ribosomal protein 49

rpm: revolutions per minute

RyR: Ryanodine receptor
SAP: Shrimp alkaline phosphatase
SBMA: Spinal bulbar muscular atrophy
SCA: Spinocerebellar ataxia
SDS: sodium dodecyl sulphate
SERCA: Sarcoplasmic/endoplasmic reticulum calcium ATPase
Sgg: Shaggy
siRNA: small interfering RNA
SOC: super-optimal broth with catabolite repression
SSC: saline sodium citrate
T: thymine
TAE: tris-acetate EDTA
TBE: tris-borate EDTA
TBP: TATA-box binding protein
TNNT: Troponin T
TudorSN: Tudor Staphylococcal nuclease
U: uracil
UAS: upstream activation sequence
UTR: untranslated region
UV: ultraviolet
V: Volts
VDRC: Vienna *Drosophila* RNAi Centre
X-gal: 5-bromo-4-chloro-3-indolyl-b-D-galactopyranoside
YAC: yeast artificial chromosome

Drosophila **nomenclature**

Throughout this thesis, *Drosophila* genes are represented by italicised lower-case text (for example "*htt*"), RNAs are represented by lower-case non-italicised text (for example "htt") and proteins are represented by non-italicised text with a capital first letter (for example "Htt").

Abstract

Expansion of a repeat sequence beyond a pathogenic range has been identified as the cause of a group of neurodegenerative diseases known as the expanded repeat diseases. Disease-associated repeat tracts have been found both within the coding region of genes, such as the CAG repeat coding for polyglutamine, or within non-coding regions. Despite the identification of the mutation involved in these diseases, the mechanism by which this type of mutation leads to cell death remains unclear. There is a substantial amount of evidence to suggest that RNA-mediated toxicity plays a role in pathogenesis of both the polyglutamine diseases and the untranslated dominant expanded repeat diseases. A common feature of the expanded repeats involved in each of these diseases is the ability of the repeat-containing RNA to form a hairpin secondary structure and therefore it has been predicted that similar mechanisms may be responsible for initiating cellular dysfunction and death in each case. This study uses a *Drosophila* model to investigate the intrinsic, RNA-mediated toxicity of three repeat sequences (CUG, CAG and AUUCU) associated with degeneration in human disease. Using a combination of hypothesis-driven and non-biased approaches, early changes elicited in response to neuronal expression of these expanded repeat tracts have been investigated. A hypothesis of a role for RNA editing in CAG repeat pathogenesis was explored using this *Drosophila* model. Microarray and proteomic approaches were also utilised to identify pathways which are perturbed by the expression of these repeat sequences. The results described in this thesis demonstrate a degree of sequence- and context-independent toxicity of expanded repeat RNA in this model, suggesting that this kind of effect may also be a component of pathogenesis in the disease situation. Pathways commonly perturbed in response to expression of these RNA species may represent particularly valuable therapeutic targets, since preventing this type of effect could provide positive outcomes in a number of diseases.

Chapter 1: Introduction

The identification of disease-causing mutations and investigation of their impact on cellular function are vital to the development of therapeutic strategies for human disease. Advances in the fields of molecular and cellular biology have greatly increased our understanding of the link between genotype and phenotype in disease progression however in a number of neurodegenerative diseases, including the expanded repeat diseases, the primary cellular dysfunction resulting from the mutation has remained elusive. This study uses a *Drosophila* model of expanded repeat disease to investigate early cellular changes which may contribute to pathogenesis.

1.0 Expanded repeat diseases

It has been nearly two decades since the first description of human disease caused by the expansion of trinucleotide repeats in the genome (1-2). There are now at least 17 diseases resulting in neurological or neuromuscular degeneration which can be attributed to the expansion of repeat sequences (3-6). The mechanism of expansion of these repeats has been suggested to involve slippage during replication of the repeat sequences (7). Typically the expanded repeat diseases manifest later in life, they show anticipation in families - that is increasing severity of pathology through the generations - and they demonstrate repeat number-dependent age of onset (8). Instability of the repeats and a tendency towards expansion of repeat number in the germ-line has been connected to the phenomenon of anticipation (7).

There is a large amount of overlap in the clinical features of the expanded repeat diseases, despite the presence of the repeats in unrelated genes, suggesting that the manifestation of expanded repeats in disease is not simply through altered expression of the repeat-containing gene. Disease-causing expanded repeats can be classified into two categories: those which fall in translated regions of a gene, generally consisting of CAG repeats encoding a polyglutamine tract, and those which fall in an untranslated region of the genome (7-8). Both translated and untranslated repeats generally show dominant inheritance, despite the fact that the untranslated repeats do not alter the protein itself.

1.1 Translated repeat diseases

1.1.1 Polyglutamine diseases

To date there are 9 diseases known to be caused by expansion of a CAG repeat encoding glutamine (Table 1.1). These diseases are Huntington's disease (HD), spinal bulbar muscular atrophy (SBMA), dentatorubral-pallidoluysian atrophy (DRPLA) and a number of spinocerebellar ataxias (SCA1,2,3,6,7&17). Despite their presence in entirely unrelated proteins, overlapping features of the polyglutamine diseases suggest that there are dominant toxic properties of expanded polyglutamine. These features include progressive neurological degeneration which is generally late-onset, repeat number-dependent age-of-onset and the formation of aggregates by the polyglutamine-containing proteins (9). Differences in the specificity of affected cells are thought to be a result of the gene in which the repeat falls, however it is unclear why mutant forms of ubiquitously expressed proteins should elicit effects on such a limited population of cells.

Disease	Polyglutamine-containing protein	Proposed Function	CAG Repeat Number	
			Disease	Normal
HD	Huntingtin	Organelle trafficking & axonal transport	36-121	6-34
SBMA	Androgen Receptor	Ligand-activated Transcription Factor	38-62	9-36
DRPLA	Atrophin-1	Nuclear receptor Co-repressor	49-88	6-35
SCA1	Ataxin-1	RNA processing	39-82	6-44
SCA2	Ataxin-2	RNA metabolism	36-63	15-31
SCA3/MJD	Ataxin-3	Protein turnover/ DNA Repair?	55-84	12-40
SCA6	CACNA1A	Voltage-gated calcium channel sub-unit	21-33	4-18
SCA7	Ataxin-7	Transcriptional regulation	37-306	4-35
SCA17	TATA box-binding protein	Transcriptional initiation	47-55	27-42

Table 1.1: Polyglutamine diseases. HD – Huntington's disease, SBMA –spinal bulbar muscular atrophy, DRPLA - dentatorubral-pallidoluysian atrophy, SCA- spinocerebellar ataxia, MJD- Machado Joseph Disease. Larger repeats within the normal range are normally interrupted by CAT for SCA1 and CAA for SCA2 while disease-causing alleles consist of pure CAG repeats.

Huntington's disease

Huntington's disease (HD) is the most common of the polyglutamine diseases, and is caused by expansion of a CAG repeat resulting in an expanded N-terminal polyglutamine tract in the Huntingtin (HTT) protein. The disease presents as progressive neurodegeneration resulting in loss of motor and cognitive function. This neurodegeneration preferentially affects the medium spiny gamma-aminobutyric acid (GABA) neurons in the striatum (9). The function of the 350kDa HTT protein is not known, although its association with microtubules and synaptic vesicles suggests a role in organelle trafficking and axonal transport (10). A direct association of mutant HTT with clathrin-coated membranes has also been demonstrated suggesting that perturbation of endocytic pathways may play a role in HD pathogenesis (11). Recent studies have demonstrated axonal transport defects in both mammalian neurons and a *Drosophila* model expressing the mutant HTT protein, as well as in HD brains, characterised by aggregation of vesicles, mis-localisation of mitochondria and apoptosis (12-14).

HTT protein shows ubiquitous expression in neurons and is also expressed at low levels throughout the body (15). A number of neuronal cell types which express HTT at all times survive in Huntington's patients – for example the striatal cholinergic interneurons – while the striatal spiny neurons most affected in Huntington's disease do not consistently express HTT (16) and it is therefore unclear how selective neurodegeneration is elicited. Furthermore, the mutant protein is expressed at similar or even slightly reduced levels in comparison to the normal protein in affected regions of the brain (9) and therefore there does not appear to be a simple relationship between expression of the mutant protein and vulnerability.

Spinal bulbar muscular atrophy (SBMA)

SBMA or Kennedy's disease is caused by a CAG repeat expansion in the androgen receptor (*AR*) gene on the X chromosome and is the only one of the polyglutamine diseases not autosomally dominantly inherited (17). The AR is a ligand activated transcription factor which is translocated to the nucleus in response to testosterone; a process which is perturbed by the expansion of the glutamine tract in the AR (18). It has been suggested that higher circulating androgen and testosterone

levels could explain why only males are affected by SBMA (19). Disease is not simply a result of loss-of-function of the *AR* gene - since mutations to *AR* do not result in SBMA and SBMA individuals show only limited androgen insensitivity - however recent reports suggest that loss-of-function may contribute to disease progression, since the AR is essential for neuronal health in a YAC mouse model (20). The disease predominantly affects spinal and bulbar motor neurons, resulting in muscular atrophy and weakness (19).

Dentatorubral-pallidoluysian atrophy (DRPLA)

DRPLA individuals commonly show seizures, involuntary muscle movement and chorea as well as dementia (21) as a result of generalised neurodegeneration in the cortex, globus pallidus, striatum and cerebellum (22). The polyglutamine tract in DRPLA is located in Atrophin-1 (23-24) which has been characterised as a nuclear receptor co-repressor and is predicted to interact with a number of transcription factors and members of the histone de-acetylase family (25). Expression of expanded polyglutamine-containing Atrophin-1 is sufficient to induce symptoms of DRPLA through a process which involves proteolytic processing and aggregation of an 120 kDa fragment containing the expanded polyglutamine tract (26). Pathogenesis does not appear to be mediated through a functional interaction of the mutant Atrophin-1 protein with the wild-type protein, since homozygous deletion of the C-terminal region of the wild-type protein does not induce neurodegeneration and also fails to modify polyglutamine-induced phenotypes (27).

The spinal cerebellar ataxias (SCAs)

The SCAs are a group of diseases characterised by progressive degeneration of the cerebellum resulting in late-onset ataxia and lack of coordination. 70% of SCA patients also show degeneration of the peripheral nervous system involving both axonal and primary neuropathy (28). There are currently 28 autosomal dominant SCAs recognised, 17 of which are caused by a known mutation and 6 of which are recognised as polyglutamine diseases (29) as listed in Table 1.1.

SCA1 is caused by a CAG expansion in the *Ataxin-1* gene which encodes a protein containing an RNA binding motif, with a proposed role in RNA metabolism (30). The RNA binding capacity of Ataxin-1 has been shown to be dependent on the length of the CAG repeat and therefore it has been suggested that loss of Ataxin-1 protein function, perhaps resulting in aberrant RNA metabolism, may play a role in disease progression (30). The expanded polyglutamine-containing protein is found in nuclear aggregates in disease models (31) however, while nuclear localisation of the polyglutamine-containing protein is required for disease progression, nuclear aggregation is not (31) and it is therefore unclear whether aggregates are a component of pathogenesis. While disease-associated *Ataxin-1* alleles consist of pure CAG repeats, normal alleles of *Ataxin-1* contain CAT interruptions in the CAG repeat which encode histidine and are thought to reduce aggregation of the protein and expansion of the repeat tract (32).

SCA2 is caused by a CAG expansion within the *Ataxin-2* gene. It is characterised by degeneration in the cerebellum and brainstem, although it is unclear how specificity is elicited since the Ataxin-2 protein is widely expressed in the brain (33). The function of the Ataxin-2 protein is unknown although it has been implicated in RNA metabolism – a role which is supported by its interaction with cytoplasmic poly-A-binding protein (34-35) and poly-ribosomes (34) – and a *Drosophila* orthologue has been characterised as a regulator of actin filament formation (36). Both mutant and wild-type Ataxin-2 are found exclusively in the cytoplasm and mutant Ataxin-2 does not form nuclear inclusions such as those thought to be pathogenic in a number of the other polyglutamine diseases (37). Normal alleles of *Ataxin-2* are frequently interrupted with CAA repeats which, while still encoding glutamine, are thought to alter the structure at the DNA and RNA level (33).

SCA3 or Machado Joseph Disease is the result of a CAG expansion in the gene encoding Ataxin-3, *MJD1*. The phenotype involves progressive ataxia as well as peripheral neuropathy which affects both motor and sensory neurons (38). The function of Ataxin-3 is unknown, however interactions with DNA repair proteins have been reported. HHR23 proteins, which are important in nucleotide excision repair, have been shown to be localised to nuclear inclusions in MJD individuals via their interaction with the mutant protein (39). Interactions with components of the proteasome and ubiquitin-binding factors, which are also present in nuclear inclusions

in MJD, suggest a role for Ataxin-3 in protein turnover (40). More recent studies also implicate Ataxin-3 in Ca^{2+} signalling via an interaction of the mutant protein with the type 1 inositol 1,4,5-trisphosphate receptor, an intracellular calcium channel (41). Interestingly, nuclear inclusions in MJD have also been demonstrated to contain wild-type Ataxin-3 protein which is recruited in a polyglutamine length-dependent manner, suggesting that the polyglutamine expansion may affect the normal function of the protein (42).

SCA6 is the result of a C-terminal polyglutamine expansion in the alpha (1A) subunit of the neuronal P/Q-type voltage-gated calcium channel encoded by the *CACNA1A* gene (43). Neurodegeneration in SCA6 is primarily localised to the Purkinje cells of the cerebellum, which coincides with high expression of the *CACNA1A* gene (44). There are two known disorders caused by missense mutations in the *CACNA1A* gene – episodic ataxia type 2 (EA2) and familial hemiplegic migraine – both of which give phenotypes similar to SCA6. It is therefore unclear whether there is toxic gain-of-function in SCA6, or whether alteration of the kinetic properties of the encoded channel is sufficient to explain the neurodegeneration observed (43). This evidence, along with the observation that the pathogenic threshold for CAG repeats is much lower in SCA6 than the other polyglutamine diseases, has led to questions about the classification of SCA6 as a polyglutamine disease (43). More recently, aggregation of the mutant CACNA1A calcium channel has been demonstrated in a knock-in mouse model of SCA6, coinciding with age-dependent neurodegeneration without a concurrent change in electrophysiological function of neurons (45). This evidence suggests that a simple change to calcium channel function is unlikely to explain SCA6 pathology.

The polyglutamine tract in **SCA7** is located within Ataxin-7. SCA7 individuals show neuronal loss in the cerebellum, brainstem and spinal cord, as well as retinal degeneration which typically leads to blindness (46). The regions where neurodegeneration is observed coincide with regions where Ataxin-7 is expressed. Ataxin-7 is thought to play a role in transcriptional regulation since it forms complexes with other transcriptional regulators including histone acetyl-transferases (47-48). Transcriptional dysregulation has been observed in the absence of ataxia in mice expressing expanded *Ataxin-7,* perhaps suggesting that this is an early component of pathology (49).

SCA17 is caused by an expanded CAG repeat in the gene encoding the TATA box-binding protein (TBP) which is a transcription initiation factor. TBP is ubiquitously expressed, yet pathogenesis in SCA17 is limited to the cerebellum and particularly affects the Purkinje cells (50). Expression of TBP containing an expanded polyglutamine tract in both cellular and mouse models has been demonstrated to result in decreased expression of the nerve growth factor (NGF) receptor, TrkA, which is required for Purkinje neuron survival (51). This relationship may go some way to explaining specificity of neurodegeneration in SCA17.

Despite the presence of the expanded polyglutamine tract in unrelated genes and the degeneration of a specific sub-set of neurons in each disease, there are a number of features of the polyglutamine diseases which suggest that common pathogenic mechanisms may be involved. Whilst a number of pathways have been demonstrated to be perturbed in models of polyglutamine pathogenesis – including axonal transport, Ca^{2+} homeostasis, transcription, protein turn-over, RNA metabolism and mitochondrial function –the primary cause of dysfunction remains unclear.

1.1.2 Pathogenesis and aggregate formation

Aggregates of polyglutamine-proteins are a feature of all of the polyglutamine diseases, however their role as a protective or pathogenic agent is contentious. They are present in neurons which undergo degeneration in each disease, and are also commonly found in neurons which do not undergo degeneration. Both cytoplasmic and intranuclear inclusions are seen in HD (52), while only intranuclear inclusions are seen in brains of SBMA, DRPLA and SCA1,3,6,7 and 17 patients. In SCA6 and some cases of SCA2 only cytoplasmic aggregates have been reported (53).

The role of the nuclear and cytoplasmic aggregates formed by mutant HTT in neurodegeneration has been much debated. It has been reported that the N-terminal fragment of the mutant HTT protein forms aggregates selectively in striatal neurons and is predictive of cell death (54). However, striatal cells transfected with mutant HTT in conditions where inclusions cannot form actually show an increase in associated cell death suggesting that formation of nuclear inclusions can actually play a protective role, perhaps by sequestering soluble forms of the mutant protein (55). There is some debate over the validity of cell culture models for neurodegeneration and it has been

suggested that animal models, which typically show increased incidence of cell death correlating with formation of aggregates, are more representative of the situation in HD individuals (56). Other expanded repeat diseases tell a different story: there is no formation of nuclear inclusions in some cases of SCA2 (37), yet many clinical manifestations of the disease overlap with the other SCAs where nuclear inclusions are observed and in the case of SCA1, are even necessary for pathogenesis (31). This observation further supports the idea that aggregates may not actually play a causative role in neurodegeneration, however the role of nuclear and cytoplasmic aggregates in pathogenesis remains unclear.

1.1.3 Polyalanine diseases

There is a class of diseases which are associated with congenital malformations which are known to be the result of the presence of polyalanine tracts in proteins. The most studied of these is oculopharyngeal muscular dystrophy (OPMD), however polyalanine tracts have been found in 494 human proteins and are highly polymorphic. They are thought to play a role in transcriptional regulation and are known to be regions of DNA binding in many proteins (57). While many polyalanine tracts are encoded by GCC or GCG codons, a –1 frameshift of a CAG or CUG repeat would also result in translation of alanine and therefore a role for polyalanine in pathogenesis of the polyglutamine diseases has been proposed.

It has been shown that the expanded CAG repeat tract in the *Ataxin-3* gene is prone to –1 frameshifts, resulting in hybrid proteins containing alanine and glutamine tracts (58). Frameshifts are thought to be elicited through the formation of secondary structures, particularly hairpins, in RNA containing long repeat tracts. In a cellular model of SCA3 treatment with anisomycin, a ribosome-interacting drug which reduces –1 frameshift, results in a reduction of the cellular toxicity of CAG repeats, suggesting a role for these polyalanine tracts in toxicity of polyglutamine diseases (58). Very low levels of +1 and +2 frameshifts, which would result in polyalanine or polyserine tracts, have been demonstrated in mutant HTT (59). Polyalanine and polyserine containing proteins are also among identified modifiers for mutant HTT, further supporting a role for frameshift in the pathogenic pathway of HD (60).

However, there is some evidence to suggest that polyalanine tracts are unlikely to play a major role in intrinsic toxicity of polyglutamine. Both CAG and CAA-encoded polyglutamine tracts have been demonstrated to show a similar range of pathogenic phenotypes in a *Drosophila* model (61), and therefore the ability of the repeat tract to undergo a frameshift to encode polyalanine does not appear to play a major role in toxicity in all cases. It is not clear whether this result indicates specific properties of polyglutamine tracts in *Drosophila* and therefore the possibility that frameshifts are an important component of toxicity in mammalian cells cannot be ruled out.

1.2 Untranslated expanded repeat diseases

There are nine diseases currently attributed to the presence of expanded untranslated repeats in the genome. Three of these – Fragile X, XE and Friedrich's ataxia – are known to be the result of loss-of-function of the gene in which the repeat falls. In the case of Fragile X and XE, methylation of the expanded repeat (CGG and GCC respectively) results in transcriptional silencing (62) while in the case of Friedrich's ataxia, a GAA expansion in an intron is thought to prevent transcriptional elongation by formation of a secondary structure in the DNA (63). These diseases are recessive and are therefore thought to have a different pathogenic mechanism to the dominant untranslated repeat diseases.

1.2.1 Dominant untranslated expanded repeat diseases

Dominant diseases caused by expanded untranslated repeats are thought to result from RNA gain-of-function. Unlike the translated-repeat diseases there is no toxic protein encoded by the repeat, yet these diseases show dominance and similar neurodegenerative phenotypes to the polyglutamine diseases. Table 1.2 shows a complete list of the dominant untranslated repeat diseases.

Disease	Repeat	Gene & region	Mechanism	Repeat Number	
				Disease	Normal
FXTAS	CGG	*FMR1,* 5'UTR	Unknown	60-200	6-53
DM1	CTG	*DMPK*, 3'UTR	RNA G-O-F	50-3000	5-37
DM2	CCTG	*ZNF9*, intron	RNA G-O-F	75-11000	10-26
SCA8	CTG/CAG	*Ataxin-8*OS, 3'UTR *Ataxin-8,* polyQ	Unknown	107-127	16-37
SCA10	ATTCT	*Ataxin-10*, intron	Unknown	800-4500	10-29
SCA12	CAG	*PPP2R2B*, 5'UTR	Unknown	55-78	9-28
HDL2	CTG	*Junctophilin-3*, depends on splicing	Unknown	51-57	14-19

Table 1.2: Dominant untranslated repeat diseases. FXTAS- Fragile X-associated tremor-ataxia syndrome, DM- Myotonic dystrophy, SCA- spinal cerebellar ataxia, HDL- Huntington's disease-like, G-O-F – gain-of-function.

Myotonic dystrophy type 1 and 2 (DM1 and 2) are the best characterised of the untranslated repeat diseases. DM1 was the earliest known example of a dominant disorder caused by repeat expansion in a non-coding region of a gene. It results from a CTG expansion above around 50 repeats in the 3'UTR of the *Dystrophia myotonica-protein kinase* (*DMPK)* gene. The disease manifests as myopathy and progressive cardiac defects, often coupled with the formation of characteristic posterior iridescent cataracts and insulin resistant diabetes (64). There is also evidence of some CNS pathology, including damage to cortical regions of the brain (65), resulting in variable and progressive cognitive impairment in individuals with DM1 (66). It is clear that DM1 is not simply the result of loss-of-function of *DMPK* since mouse *DMPK* knock-out models do not recapitulate all aspects of the disease and show much milder cardiac and muscular defects even when both alleles are knocked-out (67). A role for loss-of-function of the nearby *SIX5* gene has also been suggested however, while knock-out models of *SIX5* do have cataracts, they are not the posterior iridescent sort seen in DM1 (68-69) and therefore even a reduction in expression of both genes cannot explain all aspects of pathology.

Transgenic mice expressing 45 kb of the human *DM1* region containing at least 300 repeats also display symptoms of DM1, including CNS pathology (67), supporting a toxic gain-of function mechanism in the disease. The most significant evidence for an RNA gain-of-function mechanism in DM1 is that expression of a CUG repeat in a completely unrelated mRNA has also been shown to result in myopathy and myotonia

in a mouse transgenic model (70). It has been demonstrated by *in situ* analysis that the expanded repeat in DM1 is transcribed and mRNAs containing the repeat are spliced in the normal manner but retained in the nucleus in foci associated with nuclear components, including the splicing factor Muscleblind (MBNL) (3). A number of targets of MBNL are aberrantly spliced in DM1 individuals including cardiac troponin T (TNNT2) (71), insulin receptor (IR) (3, 71), chloride channel-1 (CLCN-1) (3), tau (72), myotubularin-related protein-1 (MTMR1), fast skeletal troponin T (TNNT3), N-methyl-D-aspartate receptor 1 (NMDAR1) (72), amyloid precursor protein (APP) (72), ryanodine receptor (RyR) (73) and sarcoplasmic/ endoplasmic reticulum Ca^{2+} ATPase1&2 (SERCA1&2) (73). These targets normally show developmental or tissue-specific regulation of splicing which is lost in DM1 individuals such that the adult splice-form or correct tissue isoform fails to be expressed. Interestingly, splicing mis-regulation can be separated from the co-localisation of MBNL, since MBNL has also been found in association with CAG-repeat containing foci but is not coupled with mis-splicing of MBNL targets in this case (71). Some pathologies associated with DM1 are directly related to mis-spliced MBNL targets; for example, it has been demonstrated that a reduction in CLCN-1 expression equivalent to the loss of expression of the adult isoform observed in DM1 results in myotonia (74) and failure to express the correct muscular isoform of IR results in failure of skeletal muscle cells to respond to insulin (75).

DM2 results from the expansion of a CCTG repeat within an intron of the *ZNF9* gene. In some cases, this expansion can result in a repeat tract which is tens of thousands of repeats long (5). Despite the presence of repeats in completely unrelated genes, DM1 and DM2 have very similar pathologies and therefore a common gain-of-function mechanism of pathogenesis has been proposed. However, there are also a number of differences between the diseases: proximal muscles are most affected in DM2 and distal in DM1, DM1 shows CNS symptoms while DM2 does not and there is no congenital form of DM2 (76). DM1 is also generally associated with more severe pathogenesis (76). These disease features may be related to the specific function of the gene in which the expanded repeat resides. In support of this idea, a role for ZNF9 in regulation of translation has been demonstrated (77). Myoblasts of individuals with DM2 show a reduced rate of translation; an effect which appears to be mediated through decreased expression of ZNF9 resulting from the presence of the CCUG expansion in the mRNA (78). This evidence – along with the observation that

heterozygous *ZNF9* mutant mice display muscle phenotypes, cardiac defects, cataracts and mRNA mis-regulation similar to what is seen in DM2 (79) – suggests that disruption to the normal function of the ZNF9 protein is likely to play a role in DM2 pathogenesis. Nevertheless, the inability of overexpression of ZNF9 to entirely mitigate translation defects in DM2 myoblasts suggests that there is also likely to be a dominant toxic effect mediated through the expanded repeat RNA itself (80). Since MBNL has also been shown to co-localise to CCUG repeat-containing foci (81) resulting in splicing alterations including changes to CLCN-1, TNNT3 and IR splicing (82-83), it has been suggested that sequestration of this RNA binding protein is a common mechanism of pathogenesis in DM1 and DM2.

A second splicing factor CUG-binding protein-1 (CUG-BP1) has also been suggested to play a role in DM1 and DM2 pathogenesis. While CUG-BP1 does not associate with repeat containing foci, MBNL and CUG-BP1 have antagonistic roles in regulating splicing such that the sequestration of MBNL results in an increase in the levels of CUG-BP1 spliceforms, perhaps through an increase in CUG-BP1 activity (84). Several lines of evidence support this hypothesis: expression of human CUG-BP1 in a mouse model is sufficient to induce splicing changes in skeletal muscle and the heart, as well as muscular defects reminiscent of those observed in DM1 (84-85). In a *Drosophila* model, neurodegeneration in the eye caused by expression of RNA containing a large CUG repeat can be suppressed by human MBNL overexpression and is enhanced when human CUG-BP1 is overexpressed (86). This result provides further evidence that the expanded repeat RNA itself can explain a large proportion of the pathology associated with DM1, since CUG repeat RNA alone is sufficient to cause neurodegeneration which can be modified by altering MBNL and CUG-BP1 protein levels.

Fragile X tremor-ataxia syndrome (FXTAS) is caused by a CGG expansion in the 5'UTR of the *Fragile X mental retardation 1* (*FMR1)* gene within the pre-mutation range (55-200 repeats) for Fragile X syndrome. Unlike Fragile X - which gives characteristic mental retardation and anxiety disorders - FXTAS does not result from loss-of-function of the *FMR1* gene. Clinical manifestations of FXTAS - including late-onset neurodegeneration presenting as gait instability, cognitive decline and tremors - cannot be explained simply by disruption of the *FMR1* gene; indeed levels of *FMR1*

transcripts are reported to be elevated by up to 8 times normal levels in FXTAS individuals and protein levels are reported to be normal (87-88).

Individuals with FXTAS show ubiquitin-positive inclusions in neurons and astrocytes throughout the cerebrum and brain-stem which also contain CGG repeat RNAs (89). The composition of these inclusions has been investigated and a number of proteins identified including the splicing factor MBNL previously implicated in the myotonic dystrophies, several intermediate filament proteins including Lamins A/C and Internexin and Heterogeneous nuclear ribonucleoprotein A2 (hnRNP A2) (89). The FMR1 protein is not present in the inclusions. These observations support a similar pathogenic pathway in FXTAS to that described for DM1 and DM2, where sequestration of proteins into inclusions with the expanded repeat-containing RNA results in loss of the normal function of those proteins which is responsible for disease pathogenesis. Expression of ectopic CGG repeat RNA in a mouse model has been shown to be sufficient to induce nuclear inclusions and death in Purkinje cells which strongly supports this hypothesis (90).

SCA8 is characterised by the slow progressive cerebellar ataxia typical of the SCAs (91), but with quite variable degrees of pathology in individual families (92). The mechanism of pathogenesis in SCA8 is a source of some debate since the expanded repeat region has recently been discovered to be transcribed in a bi-directional manner, resulting in the production of both a CUG repeat within the non-coding *Ataxin8OS* gene (also called *SCA8)* and a nearly pure polyglutamine tract encoded by the CAG repeat transcribed from the opposite strand (called *Ataxin-8)* (93). *Ataxin8OS* also overlaps with another gene on the opposite strand, *KLHL1*, which does not contain the repeat tract. The KLHL1 protein has been shown to be involved in regulation of neurite outgrowth via an actin-binding domain and a role in calcium influx regulation through P/Q-type calcium channels has also been demonstrated (94). It has been suggested that the transcription of the CUG repeat-containing transcript from the opposite strand to *KLHL1* may be involved in regulation of mRNA levels of KLHL1 since the two strands both show expression in cells which are involved in processes affected in SCA8 individuals (95). Furthermore, deletion of *KLHL1* in a mouse model has also been shown to result in gait abnormalities and loss of motor control, an effect which was also reproduced by targeted deletion in Purkinje cells alone (96), suggesting that a reduction in KLHL1 function may play a major role in SCA8

pathology. Polyglutamine-containing inclusions have also been detected in both a mouse model of SCA8 and patient tissue, raising the possibility that a mixture of RNA-mediated and polyglutamine-mediated toxicity may be at play in the disease situation (97).

Expression of the non-coding Ataxin8OS transcript alone in the *Drosophila* eye has been demonstrated to elicit a neurodegenerative phenotype, either in the presence or absence of an expansion of the CUG repeat to within the human pathogenic range (98). A number of modifiers of the phenotype caused by expression of the Ataxin8OS transcript were identified in this study, with the majority of these being RNA splicing factors, RNA-binding proteins, RNA helicases, translational regulators and transcription factors (98). Several of these modifiers also demonstrated a change in the strength of interaction dependent upon the size of the CUG repeat tract, including the *Drosophila* orthologue of MBNL and the double-stranded RNA-binding protein Staufen (98). These results support a role for RNA toxicity in SCA8 pathogenesis and suggest that this effect may be partly mediated through changes in interactions with RNA binding proteins caused by repeat expansion.

SCA10 is the only one of the expanded repeat disorders to be caused by a pentanucleotide repeat – ATTCT – of which there can be up to 4500 repeats within intron 9 of the *Ataxin-10* gene (6). The disease manifests as cerebellar dysfunction often involving seizures, with cognitive and neuro-psychiatric impairment (99). AUUCU repeat-containing RNA has been shown to form foci when overexpressed in cell culture (99) and the repeat tract itself has been demonstrated to have the potential to form a hairpin secondary structure under physiological conditions (100). Ataxin-10 protein has been shown to be essential for the survival of cerebellar neurons, however since the disease shows dominant inheritance and *Ataxin-10* heterozygous mutant mice do not recapitulate features of SCA10, a loss-of-function mechanism cannot explain disease pathology (101).

SCA12 is caused by a CAG expansion at the 5’ end of the gene encoding the brain-specific regulatory subunit of the protein phosphatase PP2A holoenzyme (PPP2R2B). PP2A has been shown to be involved in the DNA repair checkpoint (102) and to play a role in induction of neuronal apoptosis via translocation to mitochondria (103). Since the repeat can fall either in the 5’UTR or an upstream promoter of

PP2R2B depending upon alternative splicing, it has been suggested that the expansion may cause upregulation of PPP2R2B resulting in altered regulation of the PP2A enzyme and therefore altered phosphorylation of down-stream targets (104).

Huntington's disease-like-2 (HDL-2) is one of the clearest examples of phenotypic overlap between the translated and untranslated repeat diseases. It is caused by a CTG/CAG expansion in a variably spliced exon of *Junctophilin-3 (JPH3)* which results in a disease which is commonly misdiagnosed as HD (105). Characteristics of HDL-2 include striatal and cortical neurodegeneration coupled with formation of nuclear inclusions such as those typical of HD (106). Alternative transcripts contain the repeat either in the 3'UTR or translated as a polyalanine or polyleucine tract. The repeat is never translated as polyglutamine and therefore it is unclear how the repeat causes a phenotype indistinguishable from HD (105). RNA foci have been detected in frontal cortex from HDL-2 brains and expression of an untranslated CUG repeat-containing form of JPH3 in HEK293 and HT22 cells also resulted in formation of RNA foci which co-localised with MBNL, supporting an RNA pathogenesis model for HDL-2 (107).

1.3 One pathogenic pathway or many?

There are a number of characteristics of the polyglutamine diseases which cannot be explained by loss-of-function of the gene in which the repeat falls. The striking overlap in phenotypes associated with the polyglutamine diseases, which has led to misdiagnosis of SCA17 and DRPLA as HD, indicates that there are pathogenic mechanisms involved which are not gene-specific (108). The most parsimonious explanation for such a phenotypic overlap appears to be a common pathogenic pathway for the polyglutamine diseases. Several groups have demonstrated that expanded polyglutamine peptides are intrinsically toxic both in *Drosophila* models and transfected cells (61, 109-111) and there is evidence to suggest that in HD, DRPLA, SBMA and SCA3 at least, caspase cleavage can release the polyglutamine tract from the disease protein (112). However, it is likely that the functional properties of the expanded polyglutamine-containing proteins also contribute to the pathology, since there are unique clinical features associated with each disease.

The overlap between the polyglutamine and untranslated repeat phenotypes is also striking. Not least, the fact that there are SCAs caused by both polyglutamine and untranslated repeats which show the typical set of phenotypes for this group of diseases – cerebellar neurodegeneration and progressive ataxia - suggests some common causal link between the two sets of diseases. Either there is a common pathogenic agent in the two sets of diseases or an entirely separate pathogenic pathway is involved in each case, resulting in largely overlapping phenotypes. Since the presence of repeat-containing RNA is a common factor between polyglutamine and untranslated repeat diseases, it has been suggested that RNA may be a common pathogenic agent in the two groups of diseases.

A common feature of all disease-associated expanded repeat RNAs is their predicted ability to form stable hairpin secondary structures which increase in stability the longer the repeat region grows. This property has been demonstrated for CUG (113), CAG (113), CGG (113-115), CCTG (113) and AUUCU (100) repeats and appears to be integral to pathogenicity in the untranslated repeat diseases (116). Since the expanded CAG repeats which code for polyglutamine in the polyglutamine diseases can also form a stable hairpin structure at the RNA level with very similar stability to CUG repeats *in vitro* (113), it seems likely that these repeats may be similarly toxic to cells. General acceptance that the polyglutamine proteins themselves are the pathogenic agent in the polyglutamine diseases has meant that other pathogenic mechanisms have not been thoroughly investigated. There is a large pool of evidence for polyglutamine being a pathogenic agent in the translated repeat diseases, however there is equally strong evidence of a role for RNA gain-of-function pathogenesis in DM1 & 2 (81) and FXTAS (90) which demonstrates the ability of RNA to act as a pathogenic agent.

Recent evidence also supports a role for expanded repeat RNA in the translated repeat diseases: upregulating expression of the *Drosophila* RNA binding protein Mbl or overexpressing the human orthologue MBNL1 results in the enhancement of a neurodegenerative eye phenotype, as well as a decrease in life-span associated with expression of human SCA3 containing a CAG-encoded polyglutamine tract (117). This effect was not seen when the same experiment was performed using a polyglutamine tract encoded by a mixed CAG/CAA repeat, which is unable to form a hairpin in the same manner as a pure CAG repeat tract, suggesting

that the interaction is occurring at the RNA level and is sequence-dependent (117). This result supports the hypothesis that RNA toxicity may at least be one component of the pathogenic mechanism in the translated repeat diseases. It seems plausible that both polyglutamine proteins and repeat RNA may play a pathogenic role in the translated repeat diseases, while repeat RNA alone may be the pathogenic agent in the untranslated repeat diseases.

1.4 A *Drosophila* model of polyglutamine and RNA toxicity

Drosophila has been extensively used in studies of neurodegenerative disease because of the techniques available to express toxic proteins in a time- and tissue-specific manner, as well as their amenability for use in genetic and chemical screens. There are a large number of articles detailing the use of *Drosophila* for investigation of expanded repeat disease pathology, including models of polyglutamine diseases – HD (13-14, 118-121), SBMA (122-123), DRPLA (124), SCA1 (125-127), SCA3/MJD (117, 128) and SCA7 (129) – as well as untranslated repeat diseases – FXTAS (130-132), DM1 (86) and SCA8 (98) – which have revealed key features of these diseases. Expanded polyglutamine tracts have also been expressed alone, by us and others (61, 109, 133), to demonstrate the intrinsic, length dependent toxicity of these peptides.

The *Drosophila* eye is a convenient system in which to investigate neurodegenerative phenotypes, as the ordered structure makes disruptions easily observed externally and potentially toxic species can be expressed without generally resulting in lethality. The *Drosophila* eye consists of around 800 ommatidia, each of which contains 8 photoreceptors (R1-R8) and a mechanosensory bristle which are neuronal, along with cone and pigment cells which are non-neuronal (Figure 1.1). The photoreceptors consist of a rhabdomere – a membrane stack containing photopigment – and an axon which projects through the optic stalk to the brain. The rhabdomeres of photoreceptors R1-R6 are arranged as a trapezoid structure, while R7 and R8 are in the middle with R8 directly below R7. Due to this arrangement, tangential sectioning of the eye will only ever reveal 7 of the 8 rhabdomeres, R1-R6 and either R7 or R8 (Figure 1.1 C and D). Each set of 8 photoreceptors is surrounded by pigment cells which contain pigment granules and assist in optical insulation of the ommatidia.

Arranged above the photoreceptors are 4 cone cells which secrete the lens. Each ommatidium also contains a single mechanosensory bristle which projects an axon to a unique part of the brain (Figure 1.1 D). The ommatidia are arranged in an ordered array (Figure 1.1 A&B) making external visualisation of a disruption in patterning of the eye relatively simple.

NOTE:
These figures are included on page 18 of the print copy of the thesis held in the University of Adelaide Library.

Figure 1.1: The structure of the *Drosophila* eye. A) Exterior appearance of a wild-type *Drosophila* eye. Reproduced from (133). **B)** A scanning electron microscope image of the *Drosophila* eye reveals the ordered ommatidial and bristle arrays. **C)** Tangential section of the *Drosophila* eye showing individual ommatidia, each consisting of 7 visible photoreceptors surrounded by pigment granules. **D)** Illustration of an ommatidial unit. Hexagons show cross-sections through the ommatidia at different levels, revealing the arrangement of the cone cells, and the photoreceptor cells. B=bristle, L=lens, C= liquid-filled pseudocone, PP=primary pigment cell, CC=cone cells, SP=secondary pigment cells, TP=tertiary pigment cells, Rh=rhabdomeres, A=axons of photoreceptor cells, M=basal membrane, AC=anterior cone cell, PC=posterior cone cell, PLC=polar cone cell, EQC=equatorial cone cell, 1-8=photoreceptor cells 1-8. B-D are reproduced from (134).

Tissue-specific expression in *Drosophila* can be achieved using the UAS-GAL4 system (135). This system incorporates the yeast GAL4 protein; a transcription factor which binds specifically to an upstream activation sequence (UAS). There is no orthologue of GAL4 in *Drosophila* and therefore introduction of ectopically expressed GAL4 can specifically drive expression of transgenes under the control of UAS sites. The UAS-GAL4 system utilises a “driver” - a promoter region with known expression pattern – to temporally and/or spatially control expression of the GAL4 protein which is then able to specifically induce expression of a transgene (depicted in Figure 1.2 A). Specific expression in the eye can be achieved by using the *GMR-GAL4* driver, which drives expression in both neuronal and non-neuronal cells. Expression specifically in neuronal cells can be achieved using the *elav-GAL4* driver, however since this gives

pan-neuronal expression there is a tendency for lethality when it is used for expression of toxic species.

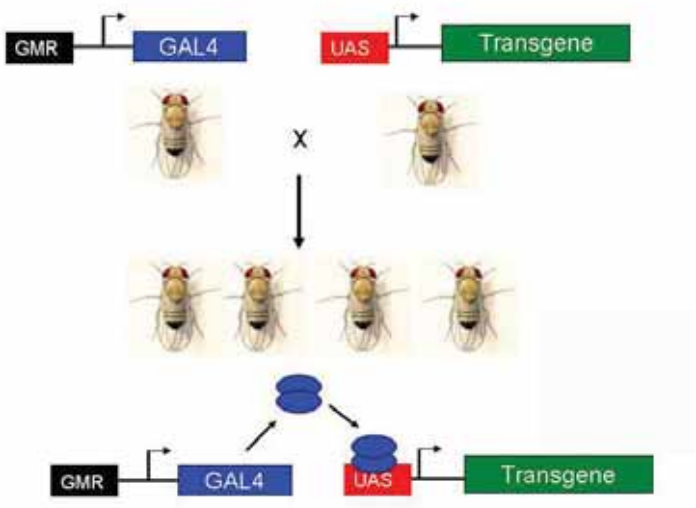


Figure 1.2: Tissue-specific expression in *Drosophila* using the UAS-GAL4 expression system. GAL4-UAS expression system utilises a tissue- and/or developmental time-specific promoter sequence which drives expression of exogenous GAL4 protein. A fly line containing this promoter, in this case the eye-specific GMR-GAL4 promoter, is crossed to a fly line containing the transgene to be expressed. In the progeny, GAL4 binds to the upstream activation sequence (UAS) of the transgene, inducing expression in a tissue-specific manner.

1.4.1 Expanded CAG and CAA-encoded polyglutamine tracts are toxic

In order to address the question of relative toxicity of polyglutamine peptides and expanded repeat RNA, transgenic flies were previously generated (described in 61, 133) containing a polyglutamine tract encoded by either a CAG repeat, which is predicted to form a hairpin at the RNA level (Figure 1.3 A), or a CAA repeat, which still encodes polyglutamine but is not predicted to form an RNA secondary structure (Figure 1.3 B). Crossing these transgenic lines to *GMR-GAL4* results in expression of the polyglutamine tract specifically in the eye. Repeats either above or below the pathogenic threshold in human disease were used in each case. The constructs used are represented in Figure 1.3 D&E. In each case, the resulting polyglutamine tract is encoded within a short peptide and is tagged with a Myc and FLAG tag. A number of independent insertion lines were generated for each construct.

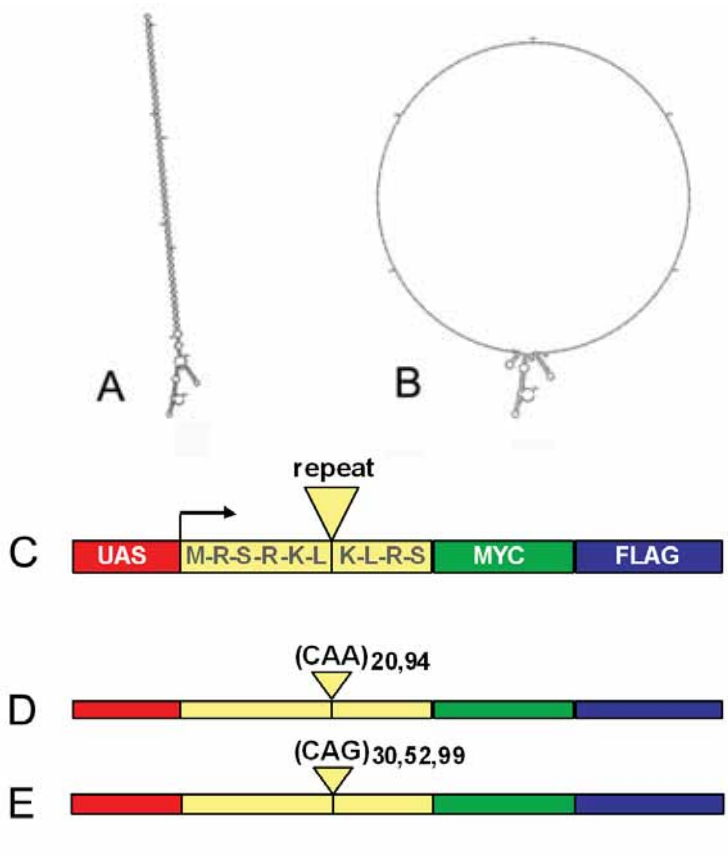


Figure 1.3: A *Drosophila* system to investigate RNA toxicity as a component of polyglutamine pathogenesis.
A&B) Representation of the secondary structures formed by CAG repeat RNA and CAA repeat RNA under physiological conditions, as predicted by Mfold. In each case 50bp of surrounding sequence from the expanded repeat constructs used in this study was included. CAG repeats form a hairpin structure by complementary base-pairing between G and C residues **(A)**. No stable secondary structure is predicted for CAA repeat-containing mRNA (represented by a circle **(B)**). Reproduced from (133).**C)** Design of polyglutamine-encoding constructs: repeats are inserted downstream of a series of UAS binding sites which control transcription and are encoded within a short peptide sequence with downstream Myc and FLAG epitope tags to allow detection of the resultant peptide. **D)** CAA repeat tracts of 20 (below the pathogenic threshold) and 94 (above the pathogenic threshold) and **E)** CAG repeat tracts of 30 (below the pathogenic threshold), 52 and 99 (above the pathogenic threshold) were generated and inserted into this construct. Previously described in (61, 133).

Expression of polyglutamine encoded by either a CAG or CAA pure repeat tract with repeat number above the human pathogenic threshold results in severe and indistinguishable disruption of the eye (Figure 1.3 C&D). Longitudinal sections of eyes from these flies stained with anti-Myc to detect the polyglutamine-containing peptides reveal the formation of aggregates in both CAG- and CAA-encoded expanded polyglutamine expressing flies (Figure 1.4 C&D). Expression of polyglutamine tracts encoded by either CAG or CAA with repeat number below the pathogenic threshold does not disrupt the exterior appearance of the eye (Figure 1.3 A&B) or cause the formation of aggregates in this system (Figure 1.4 A&B). Therefore it was concluded that polyglutamine is toxic in this system irrespective of whether hairpin-forming RNA is also present and that this toxicity is repeat length-dependent, as is seen in the polyglutamine diseases (61). This result is not consistent with a recent *Drosophila* model of SCA3 pathogenesis in which interruption of the CAG repeat tract with CAA

repeats was observed to abrogate toxicity, resulting in a much milder range of phenotypes than expression of the pure CAG repeat when similar levels of protein were expressed (117). It is likely that context of the repeat tract plays a role in mediating pathogenesis in the disease situation, a property which is not tested in our model of polyglutamine toxicity.

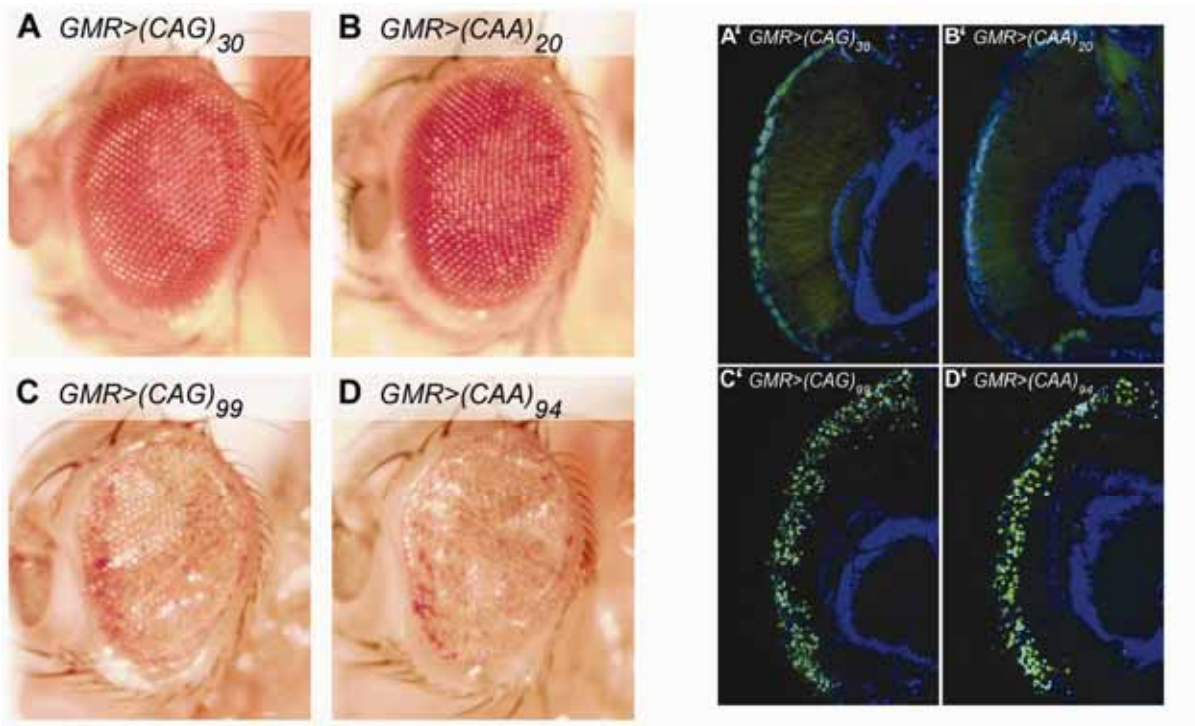


Figure 1.4: Investigation of the effects of expressing polyglutamine in the *Drosophila* eye. All crosses were performed at 25 ℃ and constructs driven by *GMR-GAL4.* **A&B)** Expression of a polyglutamine peptide encoded by either CAG or CAA with repeat number below the pathogenic threshold did not result in any alteration to the ordered structure of the eye. **C&D)** Expression of an expanded polyglutamine tract encoded by either CAG or CAA with repeat number above the pathogenic threshold results in indistinguishable and severe degeneration in the eye. This phenotype is variable between transgenic lines, but consistently includes loss of pigment and collapse of the eye. Previously reported in (61, 133). In order to investigate the ability of polyglutamine tracts encoded by CAG and CAA repeats to form aggregates, horizontal cryosections were double-stained with Hoechst (blue) to show DNA and anti-Myc (green) which detects the epitope tag of the polyglutamine proteins. **A'&B')** Polyglutamine tracts encoded by either CAG or CAA with repeat number below the pathogenic threshold show a diffuse pattern. **C'&D')** Expression of polyglutamine tracts with repeat number above the pathogenic threshold results in tight aggregates, whether encoded by CAG or CAA repeats. Previously reported in (61, 133).

The polyglutamine phenotypes observed when pathogenic length repeats are expressed involve severe degeneration of the eye, with pigmentation largely lost and a general collapse of structure. These phenotypes are so severe that is difficult to determine whether there is any cell-type specificity, however it is clear that non-

neuronal cells, such as the pigment cells, are affected. Since the effect of polyglutamine expression in the eye is so severe, it is unlikely that any difference in phenotype resulting from the presence of repeat-containing RNA would actually be observable.

1.4.2 Polyleucine peptides show distinct toxicity in the *Drosophila* eye

During the generation of the polyglutamine constructs, an expanded CTG_{86} repeat construct, encoding a polyleucine tract was made by inverting an expanded CAG repeat (depicted in Figure 1.6 A). While there are no diseases associated with expanded polyleucine, studies of expression of polyleucine in cultured cells suggest that it is more toxic than polyglutamine (136). Expression of this construct in the *Drosophila* eye resulted in a rough eye phenotype in 5 out of 13 independent lines generated, with variable severity (Figure 1.6 B) (133). Interestingly, this phenotype was quite distinct from those observed when polyglutamine was expressed in the eye: there was no loss of pigment but significant disruption to the ordered arrangement of the ommatidia, suggesting that the effect may have different cell-specificity to the polyglutamine toxicity (133).

Detection of the polyleucine peptide by staining for the Myc tag showed small speckles in the eye, suggesting that the polyleucine peptide is also able to aggregate (Figure 1.6 C). Since both the polyleucine and CAG-encoded polyglutamine expressing flies are also expressing hairpin-forming CUG or CAG repeat RNA yet have very different phenotypes, this result supports the conclusion that the majority of the phenotype associated with expression of polyglutamine encoded by either a CAG or CAA repeat is likely to be due to the polyglutamine protein and not the hairpin RNA. The phenotypes seen in the polyglutamine expressing flies are indistinguishable from each other and distinct from the phenotype seen in the flies expressing polyleucine. It is also possible that the phenotype seen in the polyleucine-expressing flies is a result of the expression of CUG repeat RNA, like that which causes SCA8 and DM1, and not the polyleucine peptide.

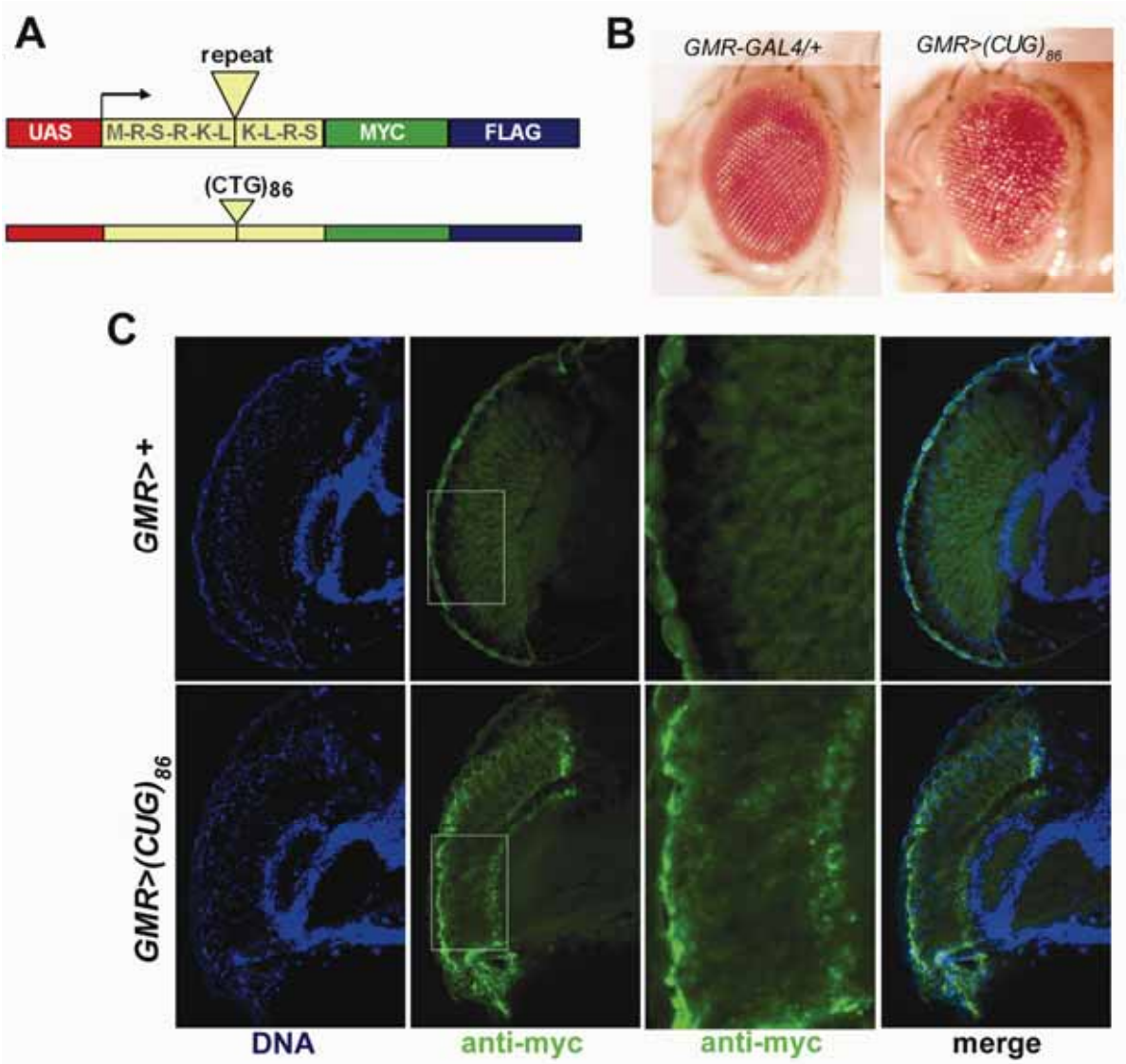


Figure 1.5: Investigation of the effects of expressing polyleucine in the *Drosophila* eye. Expression of all constructs was driven by *GMR-GAL4.* **A)** The construct used to express expanded polyleucine in *Drosophila* is based upon those shown in Figure 1.2. **B)** Expression of polyleucine primarily results in disruption of the ordered ommatidial arrays seen in the wild-type eye. **C)** Horizontal cryosections stained with Hoechst (blue) to show DNA and anti-Myc (green) which detects the epitope tag of the polyleucine peptide. Small speckles containing the polyleucine peptide can be seen, while only diffuse staining is observed in control flies. Reproduced from (133).

1.4.3 A closer look at RNA pathogenesis

Since the severity of the phenotypes associated with expression of expanded polyglutamine tracts made the contribution of RNA to toxicity difficult to determine, another set of constructs were generated in which a termination codon was inserted upstream of the repeat such that the repeat is effectively shifted into the 3' untranslated region of the product (Figure 1.7 A). Driving expression of this construct

therefore results in the presence of repeat RNA which is not able to be translated. Constructs containing pathogenic length untranslated CAG, CUG and CAA repeats were generated (called rCAG, rCUG and rCAA respectively) and multiple transgenic lines generated for each repeat. There was no consistent phenotype observed when any of these constructs were expressed in the eye using *GMR-GAL4*, however expression of 2/14 rCUG insertions did cause a mild disorganization of the eye (133). Interestingly, the phenotype associated with polyleucine expression in the eye is a similar effect to that observed in the 2 rCUG lines which showed a phenotype which may indicate that the phenotypes are a result of the expanded CUG RNA in both cases. Expression of a single rCAG transgene was also later characterised to result in a very mild rough eye phenotype which was associated with its insertion into the 5'UTR of the *filamin* gene (K. Lawlor, unpublished data). Only one of the 2 rCUG lines which showed a phenotype is associated with a gene, *suppressor 2 of zeste*. In this case, the insertion appears to be within an intron and may not have an effect on gene function and it is therefore unclear whether a positional effect is also responsible for the phenotype in this line. Expression of an expanded CUG repeat RNA has previously been demonstrated to result in degeneration in the *Drosophila* eye using a construct containing 480 interrupted CUG repeats (86), however the RNA encoded by this construct is likely to have a different structure to the pure CUG repeat tract tested in this study.

In a recent study, expression of 5 transgene insertions of an untranslated CAG repeat tagged with dsRED was demonstrated to cause mild degeneration in the brain and reduction in life-span when driven by *elav-GAL4* and mild internal degeneration in the eye when driven by *GMR-GAL4* (117). The requirement for expression of multiple insertions before a phenotype is observed in this study suggests that there is a dosage-dependent effect of expanded repeat RNA. *Drosophila* lines were subsequently generated which contained up to 4 transgene insertions of the rCAG, rCUG and rCAA constructs by recombination, in order to increase the amount of expanded repeat RNA being expressed in this system. In this experiment, a control carrying four transgene insertions of the same P-element used to generate the rCAG, rCUG and rCAA lines but lacking any insert (called "UAS") was also expressed in the *Drosophila* eye. Free GAL4 has been previously shown to induce apoptosis in neuronal cells (137) and therefore this control is important to reveal any effect of introducing multiple UAS sites on GAL4-mediated toxicity in the

eye. In this case, the P-element was inserted using the site-specific ΦC31 integration system (138). Expression of transgenes generated by this system using *GMR-GAL4* consistently results in flies with distinctly different eye colour compared to transgenics generated by random P-element mediated insertion, however the eye retains wild-type patterning and organisation (Figure 3.2 A). This effect is likely to be a result of the particular sites into which the transgenes have been inserted in this case. No phenotype was observed in the eye when four transgene insertions of rCAG, rCUG or rCAA repeats were expressed either in the eye with *GMR-GAL4* (Figure 1.7 B-E) or in the nervous system with *elav-GAL4*, however ubiquitous expression (driven by *da-GAL4*) of either the rCAG or rCUG construct, but not the rCAA construct, caused a decrease in viability (K. Lawlor, unpublished data). This result supports the idea that expanded repeat RNA can bring about toxicity in a sequence-dependent manner in our *Drosophila* model.

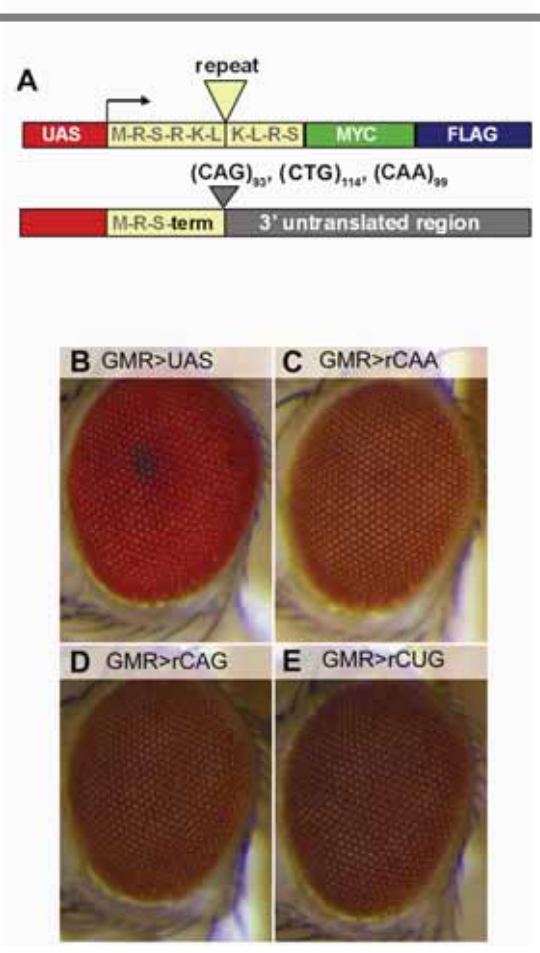


Figure 1.6: Expression of four transgene insertions of rCAG and rCUG repeats does not disrupt the external structure of the *Drosophila* eye. Expression of all constructs is driven by *GMR-GAL4.* Photos were taken by S. Samaraweera. **A)** The construct used to express expanded untranslated CAG, CUG and CAA repeats (rCAG, rCUG and rCAA) in *Drosophila* is based on the polyglutamine constructs described in Figure 1.2 except that a termination codon was inserted upstream of the repeat tract. Reproduced from (133). **B&C)** Control flies either containing four insertions of the UAS transgene without any repeat tract or expressing four transgene insertions of the rCAA repeat construct, which encodes an RNA lacking a secondary structure, have no disruption to the appearance of the eye. **D&E)** Expression of four transgene insertions of either rCUG or rCAG does not cause a disruption to the appearance of the eye, with wild-type arrangement of ommatidia apparent.

While this model does not recapitulate the neuronal phenotypes associated with expression of CAG repeat RNA observed by others (117), there are differences in the design of the constructs – including the tag used – which could alter stability and localisation of the RNAs. These constructs also differ in the location of the repeat tract: in our model the repeats are inserted in the 3'UTR of a short peptide whilst in the model which showed degeneration, an insertion of the repeat tract into the 5'UTR of the construct was used. Nevertheless, the fact that others observe degeneration as a result of expression of similar rCAG (117) and rCUG (86) constructs supports the idea that pathways involved in RNA toxicity are present in *Drosophila* and that both CAG and CUG repeat RNAs are able to elicit toxic effects.

Since the expanded repeat diseases are generally late-onset and characterised by cellular dysfunction preceding cell death, the ability to express these untranslated repeats without causing a large amount of cell death makes this *Drosophila* model ideal for investigating early events in RNA toxicity and progression of the expanded repeat diseases. Both the translated and untranslated expanded repeat diseases are associated with neurodegeneration of particular subsets of neurons. This study focuses on identifying early cellular changes resulting from expression of expanded repeat RNA, which is a common agent in both the polyglutamine diseases and the untranslated repeat diseases, in the neurons of *Drosophila*. Using this approach allows the identification of cellular processes which are specifically disrupted by hairpin-forming RNA expression and, therefore, may contribute to the onset of neurodegeneration in the expanded repeat diseases.

Chapter 2: Materials and Methods

2.1 Materials

Enzymes

Restriction endonucleases: New England Biolabs (NEB)

Pfu DNA polymerase – Roche

Taq DNA polymerase- Invitrogen

Pfu Turbo® DNA polymerase – Roche

Superscript® III reverse transcriptase – Invitrogen

RNase H – Invitrogen

RNAse out™ RNase inhibitor - Invitrogen

DNAse I - Invitrogen

Proteinase K – Sigma Aldrich

SYBR Green® PCR master mix – Applied Biosystems (ABI)

BigDye® terminator mix – ABI

T4 DNA ligase – Roche

T4 polynucleotide kinase, 3' phosphatase free – Roche

Shrimp alkaline phosphatase (SAP) – USB

LR clonase® - Invitrogen

Kits

QIAquick® gel extraction kit - Qiagen

QIAquick® PCR clean-up kit - Qiagen

GenElute® plasmid miniprep kit - Sigma

QIAprep® spin miniprep kit - Qiagen

RNeasy® mini kit – Qiagen

Expand Long Template® PCR kit – Roche

Qiagen DNeasy® tissue kit – Qiagen

2D sample cleanup kit – Bio-Rad

EZQ® protein quantitation assay – Invitrogen

Vectors

pENTR/D-TOPO® - Invitrogen

pGEM®-T – Invitrogen

pBluescript KS+ - Stratagene

pDEST-UAST – generated from pUAST by insertion of the Gateway cloning cassette (H. Dalton)

pUAST-marsh IVM – generated from pUAST by insertion of a short peptide followed by a stop codon (C. McLeod)

pBD1010 – pUAST vector with GFP inserted between the Asp718 and XbaI restriction sites (B.J. Dickson)

Cell lines

ONE SHOT® Top 10 cells – Invitrogen

SURE2® cells – Stratagene

DH5-α *E. coli* – S. Dayan

Antibiotics

Ampicillin: Sigma-Aldrich

Kanamycin: Sigma-Aldrich

Molecular Weight Marker:

1kb+ DNA ladder (Invitrogen)

Oligonucleotides

Oligonucleotides are all standard PCR grade obtained from Geneworks (Adelaide, Australia). Sequences are given 5' to 3'.

Primers for PCR and sequencing:

pUAST-Fw GAAGAGAACTCTGAATAGGG

pUAST-Rv GTCACACCACAGAAGTAAGG

UAS477-F CCTTAGCATGTCCGTGG

M13-Fw GTAAAACGACGGCCAG

M13-Rv CAGGAAACAGCTATGAC

T7 TAATACGACTCACTATAGGG

SP6 CTATTTAGGTGACACTATAG

SCA10-Fw TGGAAGAGCGCGTCTATGC
GFP repeat-Rv ATTGGGACAACTCCAGTG
Marsh TOPO-Fw CACCATGAGGAGCCG
pBKS-Fw GCGTAATACGACTCAC
pBKS-Rv GTTAATTGCGCGCTTG

Primers to amplify and sequence known editing sites in *Drosophila* (taken from (139)).
Para-dS3 Fw TCATGCACACGACGAGGATATACT
Para-dS3 Rv GCTGAATTCACCCACGTGTAGTTC
Para-dS4 Fw GAACTACACGTGGGTGAATTCAGC
Para-dS4 Rv GTCTAGGACCGCGTTATACGTGTC
GluCl Fw AACATGGGCAGCGGACACTATTT
GluCl Rv GACCAGGTTGAACAGGGCGAAGAC

Primers to amplify and sequence the rCAG construct from flies to determine editing status.
MarshATG-Fw ATGAGGAGCTGAAAGCTTCAG
Myc-Rv CTCTTCAGAGATGAGTTTCTGC

Primers for quantitative real-time PCR (Q-PCR) designed using the Primer Express software (ABI):
Rp49-Fw ATCGATATGCTAAGCTGTCGCAC
Rp49-Rv TGTCGATACCCTTGGGCTTG
Repeat-Fw TGTGGTGTGACATAATTGGACAA
Repeat-Rv TGTCGATACCCTTGGGCTTG
Mef2-Fw CATCACCGATGAACGCAATC
Mef2-Rv CGCCGAACTTGCGCTT
Insc-Fw CACGATAGCCCCGGCA
Insc-Rv CATTGTCCGAGAAGCCCG
Ctp-Fw CGACACAGGCCCTCGAGA
Ctp-Rv AACTCCTTCTTGATGTAGGCCG
CG5669-Fw CACCGAAAAGCATTAGCACG
CG5669-Rv CCGGAATCTCAATCGTCACA
Hts-Fw CACGTCCCGCGGATATTG

Hts-Rv GCTTGCGACGCTCCATTT
mGluRA-Fw GTGGTGATAACTACCTCTCTGATCG
mGluRA-Rv TCCAAATCATTGTGATTAGCACCT
Mbl-Fw CTGCTACGACAGCATCAAGGG
Mbl-Rv TTGCACGGCGGTTTATCAC
Mod(mdg4)-Fw GGGCAACACAGAGGCTCAA
Mod(mdg4)-Rv CTTCTGGCTGACAACGTACTCG
Nup62 Fw-AATTCGTTAATCTCCCGAGGG
Nup62 Rv-TCAGGTAGGTTTGGCTGCGT
DPx-2540-1-Fw-TGTCCACTGGTCGCAACGT
DPx-2540-1-Rv-GCAGGGAGTCAATGGTCCTC

Primers to generate SCA10 constructs:
SCA10Intron9-Fw AGAAAACAGATGGCAGAATGA
SCA10Intron9-Rv GCCTGGGCAACATAGAGAGA
To amplify the SCA10 repeat region within intron 9 from human DNA.

SCA10 repeat Fw ATTCTATTCTATTCTATTCT
SCA10 repeat Rv TAAGATAAGATAAGATAAGA
To expand SCA10 repeat tract *in vitro.*

SCA10 TOPO-Fw CACCTGGAAGAGCGCGTCTATGC
SCA10-Rv GCCTGGGCAACATAGAGAGA
Primers used for PCRIII in the expansion protocol. SCA10 TOPO-Fw adds a CACC non-complementary sequence to the 5' end of the product to allow direct cloning into pENTR/D-TOPO.

SP6-HindIII GATTTAGGTGACACTATAGAACTCGAAGCTTGGCC
T7-HindIII GGCCAAGCTTTAATACGACTCACTATAGGCAGACC
To add HindIII restriction sites by PCR from pGEM-T.

Nup62 ORF Rw CACCATGGTATTCCAGTTGCC
Nup62 ORF Rv CTGTGGTTACAATGGAACCATC

Primers to amplify nup62 from *Drosophila* DNA or cDNA. Nup62 ORF Fw adds a CACC non-complementary sequence to the 5' end of the product to allow direct cloning into pENTR/D-TOPO.

Cy3 labelled probes for *in situ* analysis of repeat-GFP constructs

This Cy3-labelled oligonucleotide was generated and tagged at the 5' end by Geneworks (Adelaide, Australia)

GFP1 CCTTCACCCTCTCCACTGACAGAAAATTTGTGCCC

CyDyes

Cy2, Cy3 and Cy5 – GE healthcare

Bacterial media:

All media were prepared with distilled and deionised water and autoclaved or filter sterilised, depending on heat lability. Antibiotics were added from sterile stock solutions.

L-Broth (LB): 1% (w/v) amine A, 0.5% yeast extract, 1% NaCl, pH 7.0.

SOC: 2% bactotryptone, 0.5% yeast extract, 10 mM NaCl, 2.5 mM KCl, 10 mM $MgCl_2$, 10 mM $MgSO_4$, 20mM glucose.

Plates: L-Broth with 1.5% (w/v) bactoagar supplemented with ampicillin (100 mg/mL) or Kanamycin (100 mg/mL) as appropriate.

***Drosophila* media**

Fortified (F1) *Drosophila* medium: 1% (w/v) agar, 18.75% compressed yeast, 10% treacle, 10% polenta, 1.5% acid mix (47% propionic acid, 4.7% orthophosphoric acid), 2.5% tegosept (10% para-hydroxybenzoate in ethanol).

Grape juice agar plates: 0.3% agar, 25% grape juice, 0.3% sucrose, 0.03% tegosept.

Buffers and solutions

2D sample buffer: 7M urea, 2M thiourea, 30mM tris-Cl, 4% CHAPS.

3-11 NL IPG buffer: Immobilised pH gradient buffer, non-linear, pH 3-11 GE Healthcare.

Agarose gel loading dye (6x): 30% glycerol, 0.2% (w/v) bromophenol blue, 0.2% (w/v) xylene cyanol.

Hybridisation buffer (*in situ):* 4x SSC, 0.2g/mL Dextran sulphate, 50% formamide, 0.25 mg/mL polyA, 0.25 mg/mL ssDNA, 0.25 mg/mL tRNA, 0.1 M DTT, 0.5 x Denhardt's reagent.

PBS: 7.5 mM Na_2HPO_4 , 2.5 mM NaH_2PO_4, 145 mM NaCl.

PBST: 1x PBS, 0.1% Tween.

Rehydration buffer (for IPG strips): 7M urea, 2M thiourea, 2% CHAPS, 0.5% 3-11 NL IPG buffer (GE Healthcare), 1.2% DeStreak™ (GE Healthcare).

SDS Equilibration buffer: 50 mM Tris-HCl, pH 8.8, 6 M urea, 30% glycerol, 2% SDS, trace bromophenol blue.

Squishing buffer: 25 mM NaCl, 10 mM Tris-Cl pH 8.2, 1 mM EDTA, 200 μg/mL Proteinase K.

1x SSC: 150mM NaCl, 15 mM Na citrate.

TAE: 40 mM Tris-acetate, 20 mM sodium acetate, 1 mM EDTA, pH 8.2.

TBE: 90 mM Tris, 90 mM boric acid, 2.5 mM EDTA, pH 8.3.

Drosophila stocks

Driver name	Bloomington Stock number	Insertion chromosome	Expression pattern
elav-GAL4^{c155}	458	X	pan-neuronal
elav-GAL4	8765	II	pan-neuronal
GMR-GAL4	9146	II	eye
da-GAL4	8641	III	ubiquitous

Table 2.1: Lines to drive expression using the UAS-GAL4 system.

Gene	Stock number	Source	Description
adar	-	Palladino et al. (2000) (140)	Allele 1F1, derived by imprecise excision of P-element from adar^{HD57}. Null allele.
adh	v34628	VDRC	RNAi construct
akt	v103703	VDRC	RNAi construct
CG15862	v34936/v34937	VDRC	RNAi construct
CG5669	v45300	VDRC	RNAi construct
DPx2540-1	23738	Bloomington	MB01457 Minos insertion •100bp upstream of annotated region.
hts	10989	Bloomington	Allele 01103, P{PZ} insertion into intron. Hypomorph.
	14150	Bloomington	Allele KG06777, P{SUPor-P} insertion. Hypomorph.
insc	v31488	VDRC	RNAi construct
mod(mdg4)	v52268	VDRC	RNAi construct
mef2	v15550	VDRC	RNAi construct
mbl	v28731	VDRC	RNAi construct
	7318	Bloomington	Allele E27, derived by imprecise excision of P-element from mbl^{k05507b}. Removes exons 1&2.
mGluRA	v1793/v1794	VDRC	RNAi construct (2 different insertion sites)
MBNL1	-	de Haro et al. (2006) (86)	Human MBNL isoform 1 overexpression construct.
nup62	v44806/v44808	VDRC	RNAi construct (2 different insertion sites)
	-	This study	Open reading frame, cloned from cDNA. Overexpression construct.
sgg	5435	Bloomington	Overexpression construct.
	v7005/v101538	VDRC	RNAi constructs.

Table 2.2: Candidate gene lines used in this study. VDRC is the Vienna *Drosophila* RNAi Centre.

2.2 Methods

DNA manipulation

Standard molecular genetic techniques were performed as described in (141).

Restriction enzyme digestion of DNA
Digests were carried out according to manufacturer's instructions. Where possible, enzymes were heat inactivated prior to use of digested product in further cloning. Enzymes that could not be heat inactivated were removed by agarose gel electrophoresis and gel extraction with the QIAquick® gel extraction kit according to manufacturer's instructions before further use.

Dephosphorylation of restriction enzyme digested vector DNA
Vector DNA to be used for cloning was dephosphorylated to prevent self-ligation. Following restriction enzyme digestion, 1-2 units of SAP were added directly to the reaction and it was incubated at 37 °C for at least 1 hr. The enzyme was then inactivated by incubating at 65 °C for 15 minutes.

Ligation of DNA fragments
Ligations of PCR products into pGEM®-T were performed according to manufacturer's instructions. Other ligations were generally carried out in a volume of 10 μL with 1 unit of T4 DNA ligase and 1x ligation buffer. Ligations were incubated overnight at room temperature (approximately 22°C).

Agarose gel electrophoresis
Molten 1% agarose dissolved in either 0.5% TBE or 1% TAE was supplemented with ethidium bromide and poured into a plastic gel-cast and allowed to set with well combs in place. The gel was submerged in the appropriate buffer and DNA samples mixed with agarose loading buffer were loaded into wells with one well loaded with 1 kb+ DNA markers. DNA was size-separated by applying 80-120 V to the tank. The gel was then visualised by UV light exposure using Gel-Doc™ apparatus (Bio-Rad).

Purification of DNA from agarose gels

DNA bands were excised from agarose gels and purified using the QIAquick® gel extraction kit according to manufacturer's instructions except that purified DNA was eluted in MQ rather than buffer.

Bacterial manipulation

Transformation of plasmids into bacteria and plating on selective media

Transformation was carried out by heat shock of DH5-α chemically competent cells for standard transformations, of ONE SHOT® Top 10 cells when greater efficiency was required or of SURE2® cells to prevent recombination of large repeat constructs.

Chemically competent cells stored at -80 °C were thawed on ice and 50 μL added to 2-10 μL of each ligation reaction and the mixture incubated on ice for 20-30 minutes before heat shocking at 42 °C for 45-50 seconds. The mixture was then returned to ice for 2 minutes before 2-300 μL of SOC + 0.8 % glucose was added. The tube was inverted and incubated at 37 °C for at least 1 hour, pelleted at 600 g for 5 minutes and 200 μL of SOC removed. The cells were then re-suspended in the remaining SOC and plated on LB media supplemented with Kanamycin or Ampicillin as appropriate (see materials). Where selection for β-galactosidase activity (blue-white colour selection) was required, 56 μL of 100mM IPTG and 16 μL of 50 mg/mL X-gal per plate were plated along with bacteria. Plates were allowed to dry at room temperature before incubation at 37 °C overnight.

Isolation of plasmid DNA from bacteria

Preparation of plasmids was performed using the Sigma GenElute™ kit or Qiagen spin miniprep kit according to manufacturer's instructions except that elution was performed using 50 μL of MQ water rather than buffer.

Genomic preps from *Drosophila*

Single female flies were collected and incubated at -20 °C for at least 1 hour. Flies were then squashed with a 200 µL pipette tip, 50 µL of squishing buffer was added and the reaction was incubated at 37 °C for 30 minutes. The proteinase K was then inactivated by heating to 95 °C for 2 minutes. 2 µL of the prep was used as a PCR template to amplify transgenes.

PCR amplification of DNA

PCR reactions were cycled in an MJ Research PTC-200 Peltier Thermal cycler.

Colony PCR

Selected colonies were tested for presence of the recombinant plasmid by PCR. The colony was transferred with a sterile toothpick to a master plate with appropriate selection then the toothpick was swirled in 10 µL of PCR mix (0.25 units Taq polymerase, 1x supplied buffer, 0.2 mM dNTPs, 2 mM $MgCl_2$ and 2.5 ng/µL of each primer) to release a small number of bacteria. Cycling conditions were 10 cycles of 94 °C for 30 seconds, 60 °C for 30 seconds, 72 °C for 90 seconds then 25 cycles with annealing temperature dropped to 55 °C followed by 72 °C for 10 minutes.

PCR from Drosophila genomic DNA

For verifying insertion of transgenes in SCA10 flies, PCR was performed using the expand long template PCR kit with buffer 3 according to the manufacturer's instructions except that cycling conditions were 94 °C for 2 minutes then 29 cycles of 94 °C for 20 seconds, 45 °C for 30 seconds, 60 °C for 2 minutes followed by 60 °C for 7 minutes. The very low annealing and extension temperatures were necessary since the ATTCT repeat in these constructs makes them very AT rich. Primers were either pUAST-Fw and Rv or pUAST-Fw and GFP-Rv depending on the construct.

Sequencing

DNA was sequenced using the ABI Prism™ Big Dye Terminator v3.1 Cycle Sequencing Ready Reaction Mix (Perkin-Elmer) as described in the manufacturer's protocol except that half the described amount of reaction mix was used. Generally 20 μL reactions were performed with 400-800 ng of double-stranded DNA used as a template and approximately 100 ng of primer. Reactions were performed using an MJ Research PTC-200 Peltier Thermal cycler. Cycling conditions were: 25 cycles of 96 °C for 30 seconds, 50 °C for 15 seconds and 60 °C for 4 minutes. Samples were then precipitated with 80 μL 75% isopropanol for at least 15 minutes at room temperature (approximately 22 °C) before the sequencing product was pelleted by centrifugation for 20 minutes at 13,000 rpm and the supernatant removed. The pellet was then washed in 250 μL of 75 % isopropanol and pelleted for another 10 minutes before the supernatant was removed and the pellet dried on a 95 °C heating block for approximately 5 minutes. Sequencing analysis was performed at the Institute of Medical and Veterinary Science (IMVS) Frome Road, Adelaide.

Generating SCA10 repeat constructs

Cloning the repeat tract of human ataxin-10

The original PCR product used for the SCA10 expansion contained 13 ATTCT repeats and was kindly donated by S. Dayan. The SCA10 repeat region was amplified with Taq polymerase using SCA10Intron9-Fw and Rv from HeLa DNA purified using the Qiagen DNeasy tissue kit according to manufacturer's instructions. PCR conditions were 94 °C for 3 minutes followed by 12 cycles of 94 °C for 30 seconds, 66 °C – 1 °C per cycle for 45 seconds then 23 more cycles with annealing temperature of 55 °C followed by 72 °C for 10 minutes. The product from this PCR was gel purified and sequenced using the SCA10Intron9-Fw primer. The product was ligated into pGEM-T according to manufacturer's instructions.

Expansion of the SCA10 repeat tract

Expansion was adapted from methods outlined in (142). Approximately 100 ng of the pGEM-T vector containing the SCA10 repeat region was linearised with *ApaI* or *NotI* in separate reactions. Digested vector was gel purified to ensure that no uncut vector

remained and the purified products diluted 1/20 and 1 μL used as a template for PCR.

PCR I and II

The *ApaI* digested vector was used as a template in PCRI with T7 and SCA10-repeat-Rv as primers and *NotI* digested vector was used in PCRII with SP6 and SCA10-repeat-Fw as primers. PCRI and II were carried out with the expand long template PCR kit according to manufacturer's instructions except that annealing temperature was lowered to 45 °C and extension temperature was lowered to 60 °C. The products of PCRI and II were then gel purified to remove any non-expanded product or template to ensure that these were not preferentially amplified in PCRIII.

PCRIII

2 μL of gel purified PCRI and PCR II were mixed and heated to 94 °C for 5 minutes then incubated at 65 °C for 2 mins to allow products from PCR I and II to anneal. 1 μL of this mix was then used as a template for PCRIII without further dilution. Primers for PCRIII were SCA10 TOPO-Fw, which adds a CACC to the 5' end of the product to allow direct cloning into the pENTR-D/TOPO® vector, and SCA10 Rv. PCR conditions were identical to PCRI and II. This product was used directly for ligation into the pGEM®-T vector, since the Expand Long Template™ enzyme mix also adds the A-overhang required for this ligation. SURE2® cells were used for transformation of the expanded SCA10 repeats to avoid recombination. Positive clones were identified by restriction digest. During the expansion procedure it appears that several interruptions were introduced to the repeat tract, despite the use of an enzyme mix which had proof-reading capabilities.

A number of clones were sequenced using the SCA10-Fw primer and a clone of 65 repeats was chosen for sub-cloning into the pBD1010 vector and the pUAST-marsh vector previously generated by C. McLeod (133).

Ligation into pBD1010 to generate SCA10-GFP lines

The expanded SCA10 repeat tract was re-amplified using the Expand Long Template™ kit according to manufacturer's instructions except that annealing temperature was reduced to 45 °C and extension temperature to 60 °C. Primers were SP6 and T7-*HindIII* to introduce a *HindIII* site to the 5' end of the product. The

product was then restriction enzyme digested with *NotI* and *HindIII* to generate sticky ends that could be used to ligate into pBD1010 vector digested with *NotI* and *HindIII*. The digested vector was de-phosphorylated prior to ligation and the insert was column purified using the QIAquick® PCR purification kit to remove the fragments cleaved off by the restriction enzymes. Cloning into pBD1010 in this way results in insertion of the repeat tract into the 5'UTR of the GFP transcript. During this process the repeat tract changed size, generating a variety of different repeat lengths. Clones with 67 and over 100 repeats were chosen for microinjection to generate transgenic *Drosophila.*

Ligation into pUAST-Marsh IVM to generate SCA10-Marsh lines

This vector, based on a set of constructs originally described in (143), contains a short peptide sequence followed by a stop codon such that insertion of a repeat tract downstream of the stop codon results in an untranslated 3' repeat tract. The expanded SCA10 repeat tract was inserted into a *HindIII* site within this short peptide sequence. To achieve this, the expanded product was re-amplified from pGEM®-T using the Expand Long Template™ kit and SP6 and T7 primers with *HindIII* sites added. This results in a small amount of extra sequence surrounding the repeat tract including the MCS from the pGEM®-T vector. The product of this PCR was then digested with *HindIII* before purification with the QIAquick® PCR purification kit. The pUAST-marsh IVM vector was digested with *HindIII* and de-phosphorylated before ligation with the PCR product. Colonies obtained in this way were screened for presence and direction of the insert by diagnostic restriction enzyme digestion and sequencing with the Marsh TOPO-Fw primer.

P-element mediated transformation of *Drosophila*

DNA for microinjections was prepared using the GenElute™ plasmid miniprep kit or the Qiagen spin miniprep kit according to manufacturer's instructions.

Microinjection

Microinjections were kindly performed by J. Milverton. An injection mix with 0.5 – 1 mg/mL transformation vector and 0.3 mg/mL delta 2-3 transposase plasmid (pp25.7wc) was prepared in 1x embryo injecting buffer. A drawn out capillary was

used to back-fill the injection needle with 2 mL of this mix, which had been centrifuged to remove any particulate matter. W^{1118} embryos were collected from 30 minute lays on grape juice agar plates at 25 °C, dechorionated for 3 minutes in 50% bleach then rinsed thoroughly in MQ water. The embryos were then aligned on non-toxic rubber glue and a drop of liquid paraffin was placed on them. A micromanipulator was used to position the needle and the microscope stage moved to bring the embryos to the needle for injection such that a very small amount of DNA was injected into the posterior cytoplasm.

Identification of transformants

Injected embryos were grown at 25 °C on the injection slide in a petri dish containing moist paper towel with some yeast paste. After 2 days, larvae were collected onto Whatman paper and placed into vials containing F1 medium and allowed to develop to adulthood at 25 °C. Eclosed adults were crossed to w^{1118} flies and transformants identified amongst the progeny of these crosses on the basis of the presence of colour in the eye bestowed by the *white* mini-gene present in the construct. A number of independent transformants for each construct were then mapped to determine the chromosome of insertion using the CyO and Tm6B dominantly marked balancer chromosomes present in the *Bl/CyO; Tm2/Tm6B* stock. Balanced stocks were then generated for each transformant.

Drosophila **cultures**

Flies were generally raised at either 18 °C or 25 °C with 70% humidity on F1 medium. Crosses were performed at 25 °C unless otherwise stated.

Fly crosses and strains

To generate flies carrying two independent insertions on the same chromosome, two lines of flies carrying balanced independent insertions were crossed and trans-heterozygous female virgins selected, since recombination between chromosomes only occurs frequently in female *Drosophila*. These females were then crossed to male w^{1118} flies and progeny with the two insertions recombined onto the same

chromosome were selected based on darker eye colour than either one of the single insertion lines. The recombinant chromosomes were then rebalanced with either Cyo or Tm6B dominantly marked balancer chromosomes. Flies carrying two independent insertions on the 2nd chromosome could then be crossed to flies carrying insertions on the 3rd chromosome to generate flies with four independent insertions, which could again be selected by eye colour. This method was used to generate two and four copy lines of the SCA10 constructs and to generate other recombinants described.

RNA extraction and purification

For microarray analysis and Q-PCR of microarray candidates

Approximately 100 male *Drosophila* heads were collected for each genotype and stored at -80 °C until extraction. 100 µL of Trizol™ (Invitrogen) were then added and the heads homogenised with a pestle before a further 900 µL of Trizol™ was added. The homogenate was passed through a 20 gauge needle several times and centrifuged at 13, 000 rpm for 10 minutes at 4 °C to remove cellular material. Supernatant was decanted into a sterile, RNAse free tube and incubated at room temperature (approximately 22 °C) for 10 – 15 minutes. An additional 300 µL of Trizol™ was then added and the mixture was vortexed for 1 minute before centrifugation for 15 minutes at 13, 000 rpm. The upper aqueous phase was collected (approximately 500 µL, taking care to avoid the interphase) and transferred to a sterile RNAse free tube. An equal volume of 100% ethanol was added and the mix vortexed briefly to precipitate DNA. The resultant mix was then loaded onto an RNeasy column and the remainder of the purification carried out according to the RNeasy mini kit instructions except that elution was with 50 µL of 0.1% DEPC treated de-ionised water. Preps were stored at -80 °C until use.

Transportation of RNA for microarray analysis

RNA to be used for microarrays was precipitated by adding 0.1 volumes of 3 M Na acetate, pH 5.2 and 2.5 volumes of ice cold 100% ethanol and incubating at -20 °C for at least 10 minutes before centrifuging at 10,000 rpm for 10 minutes at 4 °C. The resulting RNA pellet was then washed twice with 250 µL of 75% ethanol before

storage in 250 μL of 75% ethanol at -80 °C. Samples were shipped under ethanol on wet ice.

Microarrays

Microarrays were performed by Dr Gareth Price and in collaboration with Professor Deon Venter at Mater Hospital, South Brisbane. Affymetrix *Drosophila* GeneChip® 2.0 arrays were used and preparation, hybridisation and detection were performed according to manufacturer's instructions.

Preparation of cDNA

cDNA was prepared from 1 μg of total RNA using Superscript® III RNase H$^-$ according to manufacturer's instructions. 500 ng oligo(dT)18 was generally used per 20 μL reverse transcription reaction.

Quantitative real-time PCR (QPCR)

cDNA was diluted 1/5 and 5 μL used as a template for each 25 μL reaction with 1.26 pmol of each primer and SYBR® green master mix diluted to 1x. A standard curve was prepared by further serial dilution of cDNA (1/2, 1/5, 1/10) and used for each primer set. Each reaction was performed in triplicate in a 96 well plate. Cycling conditions were 50 °C for 2 minutes, 95 °C for 10 minutes then 40 cycles of 95 °C for 15 seconds and 60 °C for one minute on an ABI Prism® 7000 sequence detection system (Applied Biosciences). A dissociation curve was produced for each primer set to ensure that only one product was amplified in each reaction. The ABI Prism® 7000 SDS program was used to analyse data and produce the standard curve from the serial dilutions of cDNA to which each sample was then compared to determine the relative amount of product in each. Data was then exported to Microsoft Excel for further statistical analysis. The quantity of product for each sample with each primer pair was normalised to the quantity of product with Rp49 primers for the same sample to give an idea of relative expression levels between samples.

Editing Assays

RNA for editing assays was prepared as described for microarray and QPCR except that 5 adult flies were used. RNA to be used in the reverse transcription was treated with DNase I according to manufacturer's instructions before reverse transcription was performed using Superscript® III RNase H^- also according to manufacturer's instructions. 1 μL of the 20 μL reverse transcription reaction was then used as template for PCR with the Expand Long Template® kit. PCRs which gave a single product were selected for sequencing analysis, with 1 μL of the PCR product used as a template in the sequencing reaction. In each case, the forward primer for the PCR was used for sequencing.

Proteomic analysis

Protein samples were prepared from 30 male *Drosophila* heads for each genotype. Protein preparation, 2D gel electrophoresis and mass spectrometry were all performed by the Adelaide Proteomics Centre. The following is an abbreviated version of the methods used.

Sample preparation

Samples were homogenised with a micro-pestle on liquid nitrogen before addition of 100 μL of 2D sample buffer and incubation for 1 hour on ice. Samples were then centrifuged for 30 minutes at 13,000 rpm and supernatant collected. This supernatant was purified using a 2D sample clean-up kit. Pellets obtained were washed in 1mL cold acetone, centrifuged again for 30 minutes at 13,000 rpm and supernatant was removed. Purified protein pellets for each sample were pooled and dissolved in 30 μL 2D sample buffer. Protein concentrations were determined using an EZQ® protein quantitation assay against an ovalbumin standard curve according to manufacturer's instructions.

DIGE labelling

Powdered CyDyes (Cy2, Cy3 and Cy5) were dissolved in anhydrous dimethylformamide (DMF) to generate 200 pmol/μL solutions which were stored at -80°c under argon until required. 100 μg of total protein was labelled with 1 μL of Cy3

or Cy5 for each sample. An internal standard was prepared by pooling 50 μg of protein from each sample. The resulting 800 μg total protein was labelled with 8 μL of Cy2. Samples were incubated for 30 minutes on ice in darkness before the labelling reaction was stopped by addition of 1 μL of 10 mM lysine per 100 μg protein. Sample volumes were made up to 93 μL with 2D sample buffer before addition of DTT to final concentration of 65 mM and carrier ampholytes (3-11 NL IPG buffer) to a final concentration of 0.5%.

Isoelectric Focusing

24cm pH 3-11 non-linear immobilised pH gradient strips (GE Healthcare) were re-hydrated overnight in 450 μL rehydration buffer. Samples were applied to the strips by cup-loading. Isoelectric focusing was performed on an IPGphor™ II (GE Healthcare) at 20°C using a 6 step program (300 V for 2 hours, 500 V for 2 hours, 1000 V for 2 hours, a gradient of 1000-8000 V for 5 hours, 8000 V for 40,000 Vhours and 500 V for 10 hours) with the current limited to 50 μA per strip.

SDS-PAGE (2nd dimension)

Following isoelectric focusing, strips were equilibrated in equilibration buffer containing 100 mg/mL DTT for 15 minutes, then equilibration buffer containing 250 mg/mL of idoacetamide in place of DTT. SDS-PAGE was carried out using 12.5% polyacrylamide gels and an EttanDalt 12 unit in Tris-gly buffer at 15 °C at 95V for approximately 21 hours.

DIGE imaging and analysis

Gels were scanned using an Ettan™ DIGE imager (GE Healthcare) and cropped to show relevant regions. Image analysis was performed using the Differential In-Gel Analysis (DIA) module of the DeCyder™ 2D software (Version 6.5, GE Healthcare). Exclusion filters were set to reject spots with a slope of >1.1, an area of <600, a volume of <10,000 and a peak height of <80 and >65,000. The resulting spot maps were then inspected manually and poorly resolved areas excluded. Spot matching and comparative analysis were performed using Biological Variation Analysis (BVA).

Liquid chromatography-ESI mass spectrometry (MS & MS/MS)

Spots of interest were excised from the gels and digested with 100 ng of trypsin per sample. The sample was chromatographed using an Agilent Protein ID Chip column

assembly housed in an Agilent HPLC-Chip Cube Interface and connected to an HCT ultra 3D-Ion-Trap mass spectrometer (Bruker Daltonik GmbH). The column was equilibrated with 4% acetonitrile/0.1% FA and eluted with an acetonitrile gradient (4%-31%). Ionisable species were trapped and the most intense ions eluting were fragmented by collision-induced dissociation (CID) and electron-transfer dissociation (ETD).

MS & MS/MS Data analysis

Spectra were subjected to peak detection using DataAnalysis (version 3.4, Bruker Daltonik GmbH) then imported into BioTools (Version 3.1, Bruker Daltonik GmbH). An in-house Mascot database-search engine (Version 2.2, Matrix Science) was then used to identify proteins present within the sample.

In situ hybridisation of AUUCU repeat RNA

Cryosections of Drosophila larvae

Whole larvae were positioned in optimal cutting temperature (OCT) medium, frozen on dry ice and stored at -80 ºC until cutting. Sectioning was performed by K. Lawlor using a Leica CM1900 cryostat, with both the chuck and the chamber set to between -16 and -19 ℃. 10μm sections were cut and collected on poly-lysine slides to be stored again at -80 ºC until staining.

In situ hybridisation

Sections were fixed in fresh 4% paraformaldehyde for 15 mins, washed 3x 5 mins in PBS at room temperature and then briefly rinsed in 100% ethanol. After drying, slides were incubated with 0.5 ng/μL probe in hybridisation solution for at least 2 hours at 37 ºC in a humid chamber. Slides were then washed twice in 2x SSC and twice in 0.5x SSC for 15 mins each at 37 ºC. Slides were then mounted with vectashield™ (Vector Laboratories) and 1 ng/μL DAPI.

Microscopy

Image preparation was performed using Adobe Photoshop 6.0.

Light microscopy

Light photos were taken with an Olympus SZX7 dissection microscope fitted with an SZX-AS aperture. Images were captured with a Colorview IIIu camera and AnalysisRuler image acquisition software. In all cases, anterior is to the left.

Fluorescent microscopy

Fluorescent microscopy was performed using a Zeiss Axioplan 2 upright microscope with 63x PlanApo objective. Images were captured with an Axiocam MRm camera and AxioVision 4.5 image acquisition software.

Chapter 3 – The RNA editing hypothesis

3.0 Roles for RNA as a pathogenic agent

The mechanisms proposed for expanded repeat RNA toxicity have been informed by the investigation of pathogenic pathways in DM1. In DM1, the presence of an expanded CUG repeat within the DMPK transcript causes an inappropriate interaction with the MBNL1 splicing factor which results in mis-splicing of a number of downstream targets (3). Many of the pathologies associated with DM1 can be directly attributed to these splicing changes (74-75). The ability to form double-stranded regions of RNA is a common characteristic of all of the expanded repeats associated with disease to date (100, 113) and therefore it has been predicted that altered interactions with double-stranded RNA binding proteins, like those seen in DM1, may also be a general disease feature. This sort of phenomenon has been demonstrated for both CAG and CUG repeats with the double-stranded RNA regulated protein kinase (PKR) and with MBNL splicing factor (71, 76, 144-145), although the interaction of CAG repeats with MBNL1 does not appear to have the same effects on splicing as CUG repeats (71).

Many RNA binding proteins play roles in regulation of gene expression and therefore can have profound effects on cell survival. For example, PKR is able to inhibit translation in the presence of dsRNA via phosphorylation of the eukaryotic initiation factor 2 (eIF2) (146). This effect demonstrates how RNA hairpins formed by repeats could alter gene expression at a global level in the cell. PKR has been shown to preferentially bind large CAG repeats in mutant HTT RNA *in vitro* and is activated in affected regions of HD brains (144). It has been predicted that the presence of basal levels of activated PKR in neurons and axons of regions which are highly affected in HD could result in sensitivity to the presence of CAG repeats and therefore may explain the specificity of cell death in HD individuals (144). CUG repeat RNA-containing foci in DM1 also contain PKR (147) and PKR is also activated in the presence of expanded CUG repeat RNA (145), however it is unclear whether this is a component of pathogenesis since DM1 model mice lacking PKR still show phenotypes of DM1(76).

Another important family of dsRNA binding proteins is the RNAse III family which includes the Dicer and Drosha ribonucleases, best characterised for their role in RNAi and miRNA pathways. These ribonucleases specifically bind duplex regions of RNA and play a role in regulation of RNA turn-over and protect the cell against viral and transposon insults (148). In the RNAi pathway, long perfect duplexes of RNA are cleaved to 21-24bp small interfering RNAs (siRNAs) and incorporated into the RNA-induced silencing complex (RISC) resulting in sequence-specific cleavage of target RNAs. MicroRNAs (miRNAs) are of similar length but regulate translation of transcripts with imperfect complementarity. In each case Dicer and Drosha are involved in the cleavage of larger RNAs to the smaller siRNAs or miRNAs (148). These pathways play a major role in the regulation of gene expression in the cell. The presence of long dsRNAs in the cell could feed into these regulatory pathways, possibly resulting in altered expression of a large number of down-stream targets. The ability of Dicer to cleave both CUG (149) and CCG (150) repeats has been demonstrated and may indicate a role for this pathway in pathogenesis of at least some of the expanded repeat diseases.

3.1 RNA editing: roles and consequences

RNA editing is the post-transcriptional modification of bases within mRNA and is used primarily to increase the coding-power of the genome by creating alternative transcripts which may result in effects such as altered splice-site choice (151-152), increased degradation rate, altered protein binding sites (153) or incorporation of an alternative amino acid at the protein level (154). One particular RNA-editing enzyme, adenosine deaminase acting on RNA (ADAR), is specifically involved in the de-amination of adenosine (A) residues to inosine (I), which is in turn recognized by the translational machinery as guanosine (G) (154). The majority of specific targets of ADAR in mammals are neuronally-expressed ion-gated channels or receptors involved in neurotransmission; for example the serotonin and glutamate receptors (155) .

ADAR has both site-specific and promiscuous de-aminating ability, depending on the structure of the mRNA being edited. Site-specific de-amination is thought to be achieved through a very short tract (9-15 bases) of complementary RNA embedded in an intron adjacent to the exon in which editing is required, such that an

imperfect duplex is formed around the adenosine which is to be edited (154). ADAR also plays a role in viral defense where perfectly double-stranded RNAs greater than 100 base pairs long are promiscuously hyper-edited, with up to 50% of A residues being edited in some cases (156). Hyper-edited RNA associates with a number of I-RNA binding proteins which play roles in gene-regulation (see Figure 3.1 B). ADAR itself has been reported to co-localise with splicing factors and to sites of transcription, further supporting a role for editing in regulation of gene expression (156).

There are three fates for edited RNAs in the cell, depending upon their level of editing. Firstly, RNAs which are selectively edited are simply exported to the cytoplasm where they are translated and a protein incorporating amino acid changes is produced (see Figure 3.1 A). Alternatively, hyper-edited transcripts can be bound by either Vigilin or the p54nrb complex (see Figure 3.1 B). The p54nrb protein (NonA in *Drosophila*) has been shown to associate with a range of other RNA and DNA binding factors including poly-pyrimidine-tract associated splicing factor (PSF), which is a negative regulator of transcription, and Matrin 3, which is thought to be a transcriptional enhancer. In complex, p54nrb and PSF have been demonstrated to interact with the C-terminus of the large RNA polymerase II subunit co-transcriptionally (157). Binding by p54nrb may alter export of some transcripts from the nucleus via association with Matrin 3, a component of the nuclear matrix (156), however it is unclear whether this is a general effect of hyper-editing. Vigilin has been predicted to play a role in regulation of gene expression via an association with heterochromatin and also plays a role in cytoplasmic stability of mRNA (156). Furthermore, Vigilin is found in complex with ADAR itself as well as RNA helicase A. This complex is thought to recruit kinases which phosphorylate targets including RNA helicase itself and histones and is therefore predicted to be involved in gene silencing (158).

Hyper-edited RNA is also cleaved by Tudor Staphylococcal Nuclease (TudorSN), a component of the RISC complex, suggesting that promiscuous RNA editing feeds into RNAi pathways (159). The ability of TudorSN to bind and cleave edited RNAs is dependent upon a high percentage of IU and UI base pairs since these are less stable than the normal AU and UA pairings and result in localized distortions to RNA (159). Since the majority of large, perfectly double-stranded RNAs

in the cell are viral in origin, this is likely to be a protective mechanism to shut down viral gene expression. There is some evidence for an antagonistic role for RNA editing and miRNA processing by Dicer. Promiscuous editing of double-stranded RNAs may prevent further processing to produce miRNAs – either by structural changes to the RNA or because other proteins such as TudorSN bind to transcripts containing a high proportion of inosine residues - and, conversely, slicing by Dicer may result in fragments too short to be editing substrates (153, 160). Editing of specific sites on miRNA precursors has also been demonstrated to play a regulatory role by preventing Dicer processing, suggesting a more elegant role for RNA editing in gene regulation (161-162).

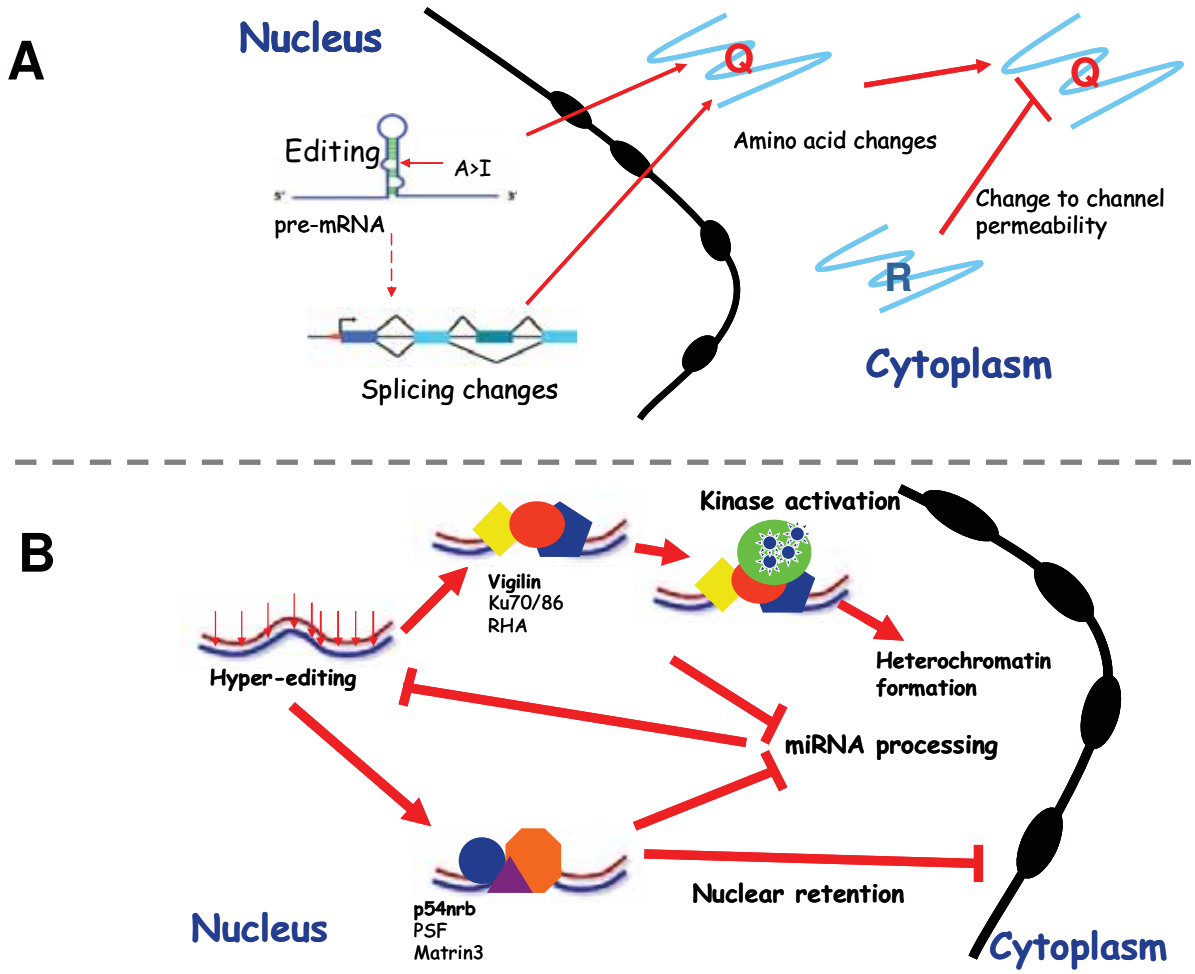


Figure 3.1: Proposed outcomes of site-specific and promiscuous RNA editing. A) The downstream effects of site-specific RNA editing by ADAR can include splice-site changes or single amino acid changes in the resultant protein. Since the majority of ADAR targets are channels and receptors, this has implications for the efficacy of channel and receptor function and therefore can be used as a regulatory mechanism by the cell. **B)** The downstream effects of promiscuous RNA editing by ADAR can include nuclear retention of the transcript resulting in translational silencing or heterochromatin formation resulting in transcriptional silencing. There is also thought to be an antagonistic relationship between miRNA pathways and RNA editing.

3.2 ADAR editing and disease

Although ADAR is fairly ubiquitous, expression in the CNS is much higher than elsewhere and expression in muscle is quite low; approximately 1/8 of that of the MBNL1 splicing factor which plays a role in DM1 and 2 (76). Site selective A to I editing has been widely studied in mammalian systems, where editing plays a vital role in the regulation of neurotransmission pathways through targets which include the glutamate receptor family and G-protein-coupled serotonin receptors. ADAR activity has been shown to be developmentally regulated in mammalian brains (163) producing different forms of a number of ion channels at different developmental stages: for example the $GABA_A$-α3 channel is edited only around 40% of the time in newborn mice compared to 100% in adults (152). This editing is linked to the switch from an excitatory role for GABA in the developing brain, to an inhibitory role in the adult brain via alteration of the chloride permeability of the channel on GABA binding. Loss of this editing would be predicted to result in a continuous large chloride influx, which could be damaging to the neuron over a long period of time.

The editing of the AMPA (α-amino-3-hydroxy-5-methyl-4-isoxazole propionateglutamate) receptor B subunit (GluR-B) at a single position, resulting in a substitution of arginine for glutamine, controls the Ca^{2+} permeability of the channel and this position has been found to be edited 99.9% of the time in humans. Mutant mice expressing unedited GluR-B at even low levels with the edited form are prone to epileptic seizures and premature death due to toxic effects of increased Ca^{2+} influx into neurons (164). Substitution of the normal allele for one encoding the edited form of GluR-B rescues this phenotype (165). It is unclear why such a vital amino acid substitution is not simply encoded but requires editing; however there may be some degree of spatial and temporal specificity of editing which simply altering genomic sequence does not allow. Mice mutant for *ADAR2*, the isoform responsible for the editing of this GluR-B site, show similar phenotypes to mice expressing the unedited form and can also be rescued by the substitution of an allele encoding the edited form of GluR-B (165). A reduction of editing of this site has also been observed in the motor neurons of amyotrophic lateral sclerosis (ALS) patients, suggesting that this may be sufficient to result in neuronal death (166).

Amongst neurons which express the GluR-B receptor are the medium spiny neurons of the striatum which are most vulnerable in Huntington's disease. Loss of

these neurons has been shown to result in movement control defects in animal models (167). It has also been demonstrated that there is a reduction of GluR-B Q/R site editing in the striatum from 99.5% in controls (normal and schizophrenic) to approximately 95% in HD brains (168). Intriguingly, ADAR has been shown to edit the cytoplasmic FMR1 interacting protein 2 (CYFIP2) in a mouse model (169). CYFIP2 has been proposed to play a role in neuronal path-finding and axonal growth (170) and interacts with and regulates expression of the FMR1 protein, suggesting that there is also a possible link to pathogenesis in both Fragile X mental retardation and FXTAS.

3.3 *Drosophila* Adar

One *ADAR* orthologue has been identified in *Drosophila* by sequence comparison with mammalian and *C.elegans ADAR*. Editing is highly developmentally regulated – via use of alternative promoters and splice sites such that activity is low in embryonic stages and high in adults – and seems to occur at a much higher rate in the CNS than in other tissues, since the inosine content of RNA in the brain is much higher (171). *Adar* loss-of-function mutants show complete loss of editing of the sodium channel paralytic (para), the Dmca1A calcium channel (cac) and glutamate-gated chloride channel-α (DrosGluCl-α) all of which are vital for neuronal function in *Drosophila* (172). The mutant flies display a number of interesting phenotypes including age-dependent neurodegeneration, reduced neuronal tolerance to anoxia (172), an overall reduction in life-span under competitive conditions and temperature sensitive seizures or paralysis (140).

The specific targets of Adar currently known in *Drosophila* are not orthologues of the known targets in mammals, although they play functionally similar roles. It is probable that there are more targets in both mammals and *Drosophila* which have not yet been identified. A recent study comparing sequences in *Adar* mutants to wild-type *Drosophila* revealed that a number of transcripts which are edited play roles in neurotransmission and synaptic growth, including AP-50 which plays a role in secretory pathways for neurotransmission and the Boss glutamate receptor (173). This suggests that there is a common regulatory role for mammalian ADARs and *Drosophila* Adar in neurotransmission. *Adar* mutant flies also show statistically significant upregulation of reactive oxygen species (ROS) scavengers including the

thioredoxin homologue deadhead (Dhd) and the cytochrome p450 family member Cyp4g1 (174). It is therefore suggested that Adar plays a role in oxidation pathways and regulates ROS scavengers, although it is not clear how this fits with the neurotransmission regulation roles already established for Adar.

3.4 A role for RNA editing in the dominant expanded repeat diseases?

The downregulation of RNA editing of GluR-B specifically in the striatum seen in individuals with HD suggests an intriguing link between RNA editing and expanded repeat disease. Furthermore, the observation that the majority of specific targets of RNA editing are neurotransmitter receptors and voltage and ligand-gated ion channels appears consistent with a model where long-term disruption to editing could result in progressive cellular dysfunction. We therefore hypothesise that expanded CAG repeats may be sequestering ADAR, whether or not they are edited themselves, resulting in a reduction in editing of ADAR targets. This could represent a common pathogenic mechanism in both the untranslated CAG repeat and polyglutamine diseases similar to the sequestration of MBNL1 observed in DM1 and DM2. Both the high expression of ADAR in the CNS and the dependence of neural circuits on the normal function of these targets suggest that neural systems would be highly affected by a reduction in ADAR editing activity.

There are various mechanisms by which ADAR could influence disease progression in the dominant expanded repeat diseases. Firstly, there could be a direct interaction between ADAR and the expanded repeat tract, either involving editing by or simply sequestration of the ADAR enzyme. It has been shown that the location of mammalian ADAR2 within the nucleus has a profound effect on the overall editing level of endogenous transcripts (175), therefore the presence of ADAR in repeat-containing foci could prevent the enzyme from performing its normal function. While studies using tissue from DM individuals revealed that ADAR does not co-localise with either CUG or CCUG repeats in muscle (76), the localization of ADAR with CAG repeats has not been investigated.

It is also possible that either the expansion of the repeat tract results in it being edited in a promiscuous manner – a process which could have regulatory outcomes for the entire repeat-containing transcript – or that the presence of an

expanded repeat tract disrupts the normal editing that occurs in the non-expanded transcript, resulting in aberrant processing. The repeat regions involved in the expanded repeat diseases are all predicted to form large, imperfect hairpin structures at the RNA level. In the case of CAG repeat RNAs, this hairpin contains a mis-match every third base between two adenosine residues (as depicted in Figure 1.3 A). It is known that ADAR is able to promiscuously edit long, perfectly double-stranded RNAs (145, 176) and that ADAR editing in the human brain and in *C. elegans* frequently occurs in hairpin-forming non-coding regions of the RNA (177), however it is not clear whether structures such as those predicted to be formed by CAG repeats are likely targets of ADAR editing. Alternatively, there could be an antagonistic effect on the RNA editing pathway by repeat-mediated over-activation of other RNA processing systems such as the miRNA pathway. Our *Drosophila* model is an ideal system to investigate the contribution of Adar to repeat pathogenesis, since it enables the use of rapid genetic techniques to look for both direct and indirect interactions between Adar and expanded CAG repeat RNA.

3.5 Investigation of the effects of altering Adar expression in *Drosophila* expressing expanded repeat RNA

Since there is no phenotype in *Drosophila* expressing up to four transgene insertions of rCAG or rCUG repeats under the control of either a pan-neuronal driver (*elav-GAL4*) or an eye specific driver (*GMR-GAL4*), it is not possible to test for modification of an RNA-induced phenotype. *Adar* mutant *Drosophila* are viable and fertile, with temperature-dependent locomotion defects (140). They do not show a disruption to the exterior organisation of the eye, either when an RNAi construct targeting Adar is driven specifically in the eye by *GMR-GAL4* or when a null allele of *Adar* (*Adar*$^{\Delta}$*)* is introduced, resulting in a 50% reduction in Adar expression in all tissues. Expression of up to four transgene insertions of either rCAG, rCUG or rCAA repeats in the eye in a heterozygous *Adar*$^{\Delta}$ background similarly does not result in a disruption to the exterior organisation of the eye (Figure 3.2 B-D).

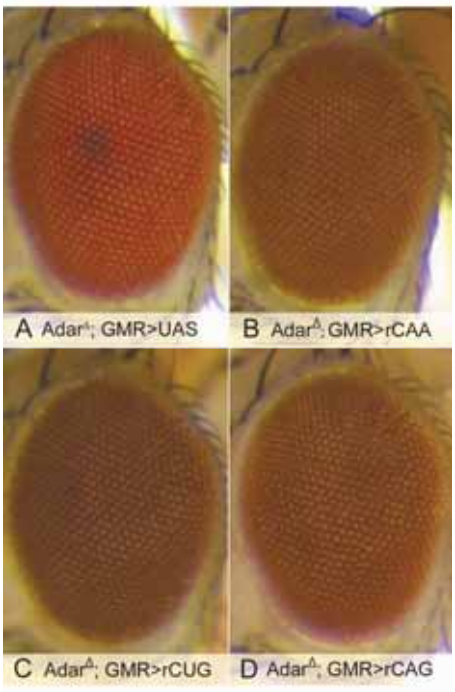


Figure 3.2: Effect of expression of rCAA, rCUG or rCAG repeats in a heterozygous Adar null background. Expression of untranslated repeat tracts is driven by *GMR-GAL4*. In each case, female flies heterozygous for *Adar*$^{\Delta}$ were photographed. **A)** Eyes of heterozygous *Adar*$^{\Delta}$ females are wild-type in appearance. In this case, the flies also carry *GMR-GAL4* and four transgene insertions of the UAS region from the repeat constructs, but with no repeat tract. **B-D)** Expression of four transgene insertions of the rCAA, rCUG or rCAG repeat constructs in the eye of *Adar*$^{\Delta}$ heterozygous female flies does not alter the exterior appearance of the eye.

Expression of a polyglutamine tract encoded by either CAG or CAA repeats in the *Drosophila* eye has been shown to cause a severe disruption to the eye, resulting in loss of pigment and in some cases necrotic patches (61) (described in 1.4.1). While these polyglutamine phenotypes are visually indistinguishable, only the flies expressing the CAG-encoded polyglutamine tract are also expressing CAG repeat RNA. Therefore the identification of genes which modify the eye phenotype in the CAG repeat expressing flies but not the CAA repeat expressing flies could indicate an interaction with the CAG hairpin RNA. Using this reasoning, the effect of a reduction in Adar levels on the polyglutamine eye phenotypes was investigated. Introducing one null allele of *Adar (Adar*$^{\Delta}$*)* into flies expressing polyglutamine encoded by a CAA repeat appears to slightly suppress the eye phenotype: there is a slight increase in the size of the eye and a slight reduction in the extent of the loss of pigment (Figure 3.3 B compared to E). This effect is not seen in flies expressing polyglutamine encoded by a CAG repeat (Figure 3.3 A compared to D).

In support of a sequence-dependent interaction between Adar and CAG repeat RNA, the mild disruption to patterning of the eye caused by expression of a translated CUG repeat encoding polyleucine (shown in Figure 3.3 C and described in section 1.4.2) is also suppressed by loss of one copy of *Adar*, resulting in a marked reduction in the area of roughness and an increase in the size of the eye (Figure 3.3 F). Since CAG and CUG repeat tracts are both predicted to form RNA hairpins, this

result indicates that the difference in secondary structure between the CAG and CAA repeat RNAs is not solely responsible for the difference in observed interaction. It is unclear how the suppression of the CUG-encoded polyleucine and CAA-encoded polyglutamine phenotypes is mediated, however the inability of a reduction in Adar to modify the CAG-encoded polyglutamine phenotype may suggest that this repeat sequence is unique.

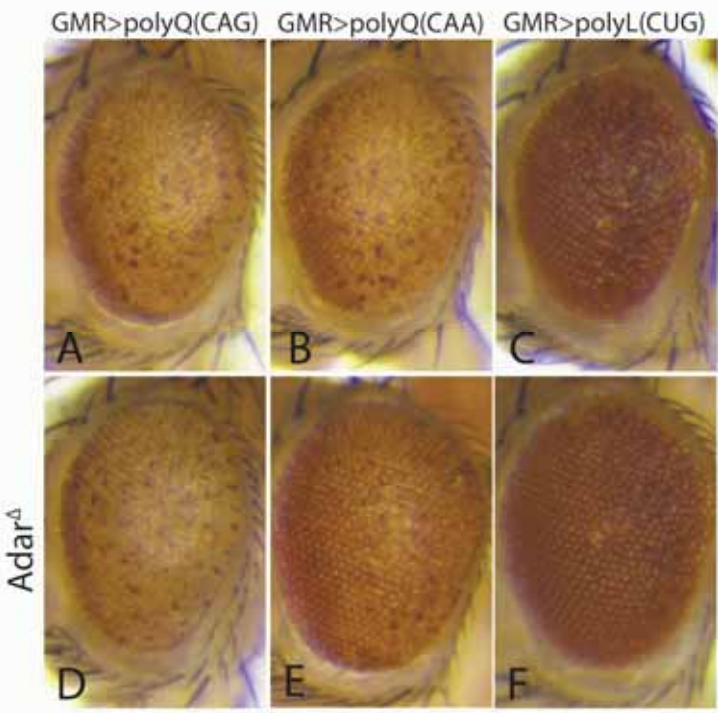


Figure 3.3: Effect of reducing Adar levels on *Drosophila* eye phenotypes elicited by expression of CAG or CAA-encoded polyglutamine or CUG-encoded polyleucine. Flies were grown at 25°C and repeat expression driven by *GMR-GAL4*. *Adar*$^{\Delta}$ flies are heterozygous female flies in all cases. **(A&B)** Expression of either CAG or CAA-encoded polyglutamine results in a loss of pigment eye phenotype. **(C)** Expression of CUG-encoded polyleucine results in roughening of the surface of the eye without loss of pigment. **(D)** The exterior appearance of the eye does not change when a heterozygous *Adar*$^{\Delta}$ mutation is introduced into flies expressing CAG-encoded polyglutamine. **(E)** Suppression of the loss of pigment eye phenotype and an increase in the size of the eye is observed when a heterozygous *Adar*$^{\Delta}$ mutation is introduced into flies expressing CAA-encoded polyglutamine. **(F)** Suppression of the disorganisation of patterning and an increase in the size of the eye is observed when a heterozygous *Adar*$^{\Delta}$ mutation is introduced into flies expressing CUG-encoded polyleucine.

3.6 Investigation of the editing status of ectopically expressed CAG and CAA repeat tracts in *Drosophila*

One difference between CAG repeat RNA and either CUG or CAA repeat RNA is its ability to form a hairpin secondary structure containing multiple mis-matched adenosine residues. It is possible that these residues could themselves be targets of Adar editing or could result in the sequestration of Adar, in a similar manner to the sequestration of MBNL splicing factor by expanded CUG repeats in DM1, without themselves being edited. To determine whether expanded CAG repeats are edited in

Drosophila, translated and untranslated CAG and CAA repeats were expressed in the nervous system (*elav>GAL4)*, total RNA was extracted and the ectopically expressed repeat tracts were reverse transcribed and sequenced. There was no detectable A to I editing in either CAG or CAA repeat tracts in this model (Figure 3.4). While this result does not support a role for Adar in editing expanded CAG repeat transcripts in *Drosophila*, it does not rule out the possibility that Adar is sequestered by the presence of expanded CAG repeat RNA.

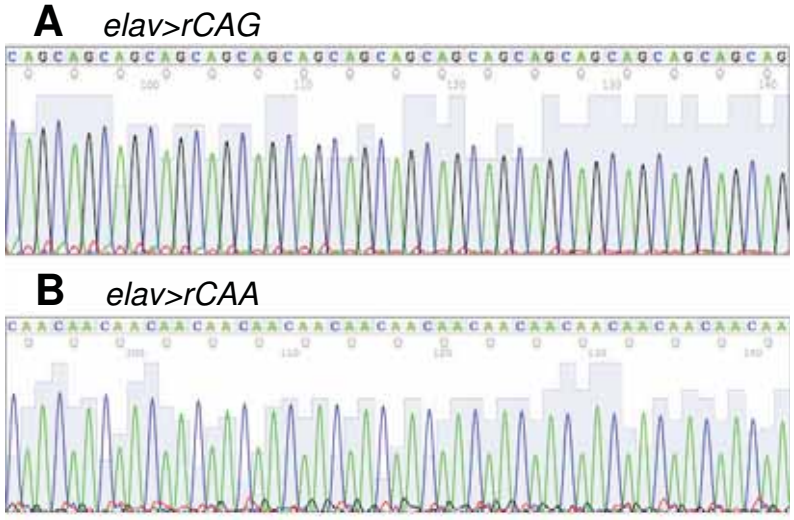


Figure 3.4: There is no detectable editing of pure CAG or CAA repeats expressed pan-neuronally in *Drosophila*. Untranslated rCAG **(A)** and rCAA (**B)** repeats were expressed specifically in the nervous system with *elav-GAL4* and total RNA from approximately 100 heads was purified for reverse transcription and sequencing.
A&B) Sequenced rCAG and rCAA repeat tracts show no sign of editing, which can be observed as a mixed A/G signal at the edited residue. Translated CAG or CAA repeats expressed pan-neuronally were also sequenced and similarly showed no evidence of A to I editing (data not shown).

3.7 Investigation of the effect of expression of CAG repeat RNA on editing of endogenous Adar editing targets in *Drosophila*

A situation where Adar is sequestered through a direct interaction with repeat-containing transcripts would be expected to result in a reduction in observed editing of the normal targets of the enzyme, as is observed when Adar levels are reduced by genetic means. In order to test whether a reduction of Adar activity occurs in CAG repeat-expressing *Drosophila*, four transgene insertions of the rCAG repeat construct were expressed pan-neuronally with *elav>GAL4* and the editing status of a set of sites known to be highly edited in *Drosophila* was determined by reverse transcription of the mRNA and sequencing. Two adenosine residues for the GluCl-∝ channel (Figure 3.5 A) and five sites for the para sodium channel (Figure 3.5 B & C) were

tested in this manner, chosen because these sites are known to be normally edited at levels detectable by sequencing and this editing is reduced or completely lost in *Adar* mutant *Drosophila* (172). A significant decrease or complete loss of editing was seen at all sites in *Adar*$^{\Delta}$ flies, while no consistent decrease in editing was observed in rCAG repeat expressing flies. It therefore does not appear that the expression of CAG repeat RNA is able to reduce *Adar* editing levels in *Drosophila*, at least not at a level detectable by sequencing.

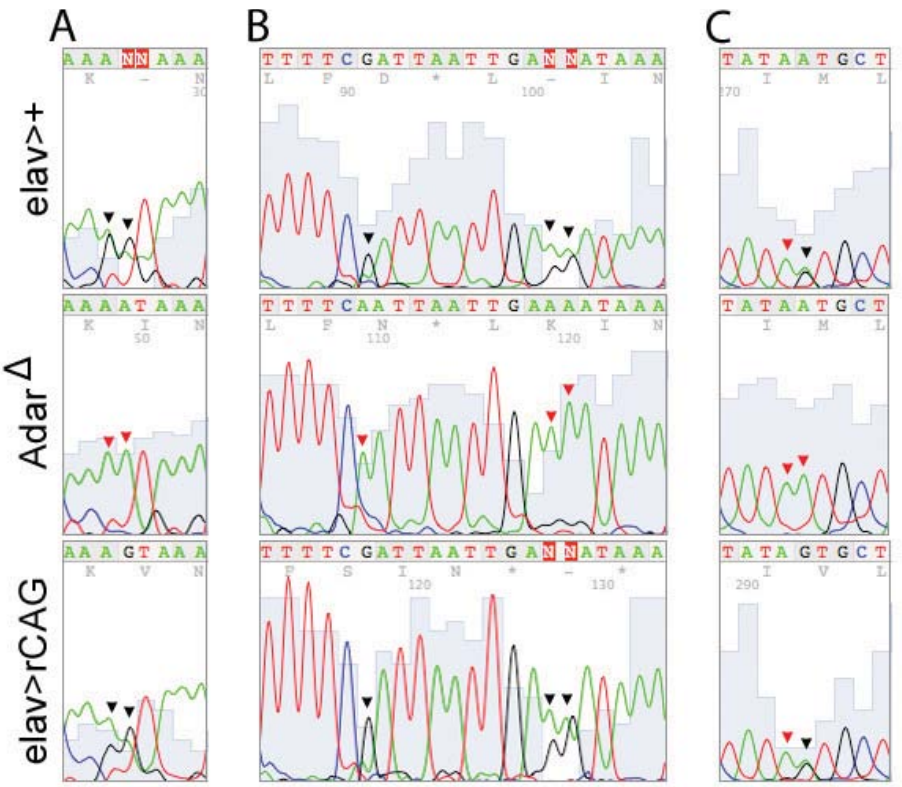


Figure 3.5: There is no detectable decrease in editing of normal Adar targets when rCAG repeat transcripts are expressed throughout the nervous system of *Drosophila*. Sites known to be edited with high frequency in *Drosophila* were selected and sequenced in male *Drosophila.* Since *Adar* is on the X chromosome in *Drosophila*, this means that Adar$^{\Delta}$ flies tested do not have a functional *Adar* allele. A clear reduction in the level of RNA editing at known editing sites is observed in *ADAR* null flies (***Adar***$^{\Delta}$) compared to either wild-type flies (not shown) or the out-crossed *elav*-GAL4 driver line (***elav>+***). There is no reduction in editing when four transgene insertions of the rCAG construct are expressed throughout the nervous system (***elav>rCAG***). **A)** Sites known to be edited in the *Drosophila* GluCl-∝ channel mRNA. **B&C)** Sites known to be edited in the para channel mRNA. Black arrow heads indicate sites that show a substantial proportion of editing, observed as a mixture of A and G signals in the sequencing chromatograph, while red arrow heads show sites that appear to have a reduction or loss of editing. Interestingly, there is one site reported to be highly edited in *Drosophila* (172) which does not show evidence of editing in any genotype in this assay (**C**- red arrow heads).

3.8 Summary of investigation of RNA editing as a component of CAG repeat RNA pathogenesis.

There is a large amount of data demonstrating the pathogenic nature of CUG repeat RNA in DM1, however the possibility of CAG repeat RNA being pathogenic and the mechanisms by which this might occur have not been extensively explored. We propose a mechanism whereby the expansion of a CAG repeat tract might result in sequestration of the RNA editing enzyme ADAR in much the same manner as MBNL is sequestered in DM1. ADAR appears to be a good candidate for a role in at least some of the expanded repeat diseases because of its demonstrated importance in maintaining function in a specific subset of neurons which includes those most affected in HD.

There are several mechanisms by which expansion of CAG repeat tracts might be envisaged to disrupt ADAR activity, two of which were investigated in this study. Firstly, the expanded repeat tracts themselves might be editing targets for the enzyme and therefore expansion beyond a particular repeat number may have a rate-limiting effect on editing of normal ADAR targets in the disease situation. Alternatively, expanded CAG repeat tracts may sequester ADAR without themselves being edited. Despite observing a difference in interaction between *Adar* and either CAA-encoded polyglutamine or CUG-encoded polyleucine compared to CAG-encoded polyglutamine in the eye, no evidence was obtained that Adar is able to edit neuronally expressed CAG repeat RNA in *Drosophila*, or that the presence of CAG repeat RNA in neurons results in a reduction in editing of the normal targets of *Drosophila* Adar; an effect which could be indicative of sequestration of the enzyme. It was therefore concluded that Adar is unlikely to be binding to the expanded CAG repeat tract in our *Drosophila* model and that the difference in interaction between *Adar* and expanded CUG or CAA repeats and CAG repeats in the *Drosophila* eye is not the result of a direct physical interaction between Adar and the repeat tract, but is most likely mediated through an indirect mechanism.

The expanded repeat disease model investigated in this study examines only the intrinsic toxicity of expanded repeat tracts and therefore the possibility that expanded CAG repeats are able to be edited and/or sequester ADAR in the context of the disease-associated transcripts cannot be ruled out. The *Drosophila* Adar

enzyme may also have different binding preferences than the human enzymes; a possibility which could be tested in this model by generating *Drosophila* lines containing insertions of the human *ADAR* genes. Therefore, while the results presented in this study do not support a role for Adar in pathogenesis in this *Drosophila* model, a role for RNA editing in the expanded repeat diseases may warrant further investigation.

Chapter 4: Identifying pathogenic pathways of expanded repeat disease by proteomic analysis

The experiments described in Chapter 3 of this thesis investigated a role for the editing enzyme Adar in pathogenesis in flies expressing expanded CAG repeat RNA. While this kind of directed investigation of pathogenic pathways can provide biologically relevant information, one advantage of exploring disease pathways in *Drosophila* is the ability to perform large scale screens of candidate genes with relative ease. This property makes genetic validation of candidates identified by techniques such as microarray and proteomic analysis, which typically produce large amounts of data, a viable option.

There are indications from both mouse and *Drosophila* models of DM1 that CUG repeat RNA alone is intrinsically toxic and can elicit many of the pathological features observed in the disease (70, 86). More recently, the toxicity of untranslated CAG repeats has also been demonstrated in *Drosophila* (117), however the primary cellular changes responsible for neurodegeneration are not yet clear. In order to investigate the pathogenicity of repeat-containing RNA specifically in neurons, transgene insertions of the rCAA, rCAG and rCUG repeat constructs were expressed under the control of the pan-neuronal *elav-GAL4* driver (as depicted in Figure 4.1). Flies expressing up to four transgene insertions of these constructs are viable and show no obvious phenotypes (K. Lawlor, unpublished data). Expression of these untranslated repeats in the nervous system therefore allows the investigation of the effects of expression of hairpin repeat RNA (rCUG and rCAG) in cells which are not dying, but are likely to demonstrate early hallmarks of expanded repeat pathogenesis. Given that CAG and CUG repeat RNAs form structurally similar hairpin structures, it has been proposed that similar mechanisms may be involved in pathogenesis of disease caused by each of these repeats. Therefore common changes observed in both CAG and CUG repeat expressing flies are of particular interest in this study. The rationale for this methodology is summarised in Figure 4.2.

Using this *Drosophila* system, *in vivo* perturbations to neuronal pathways caused by expression of expanded untranslated repeats have been investigated by looking at both proteomic and transcriptional changes. In order to identify early events in expanded repeat pathology – which are more likely to represent causative

changes rather than the down-stream effects of cellular perturbation – newly eclosed flies were used in these analyses. By investigating changes occurring at both the protein (using 2D-DIGE followed by MS) and transcript (using microarrays) levels, a broader view of the sorts of pathways which are disrupted by repeat expression can be obtained. It is predicted that amongst the proteins and transcripts altered as a result of hairpin repeat expression in this model will be mediators of pathogenesis and therefore these analyses should provide information about the primary steps in pathogenic progression. In order to validate results obtained from these studies, the *Drosophila* eye was then used as a tool to look for modification of phenotypes associated with expanded repeat expression. The results of these experiments are described in Chapters 4-6 of this thesis.

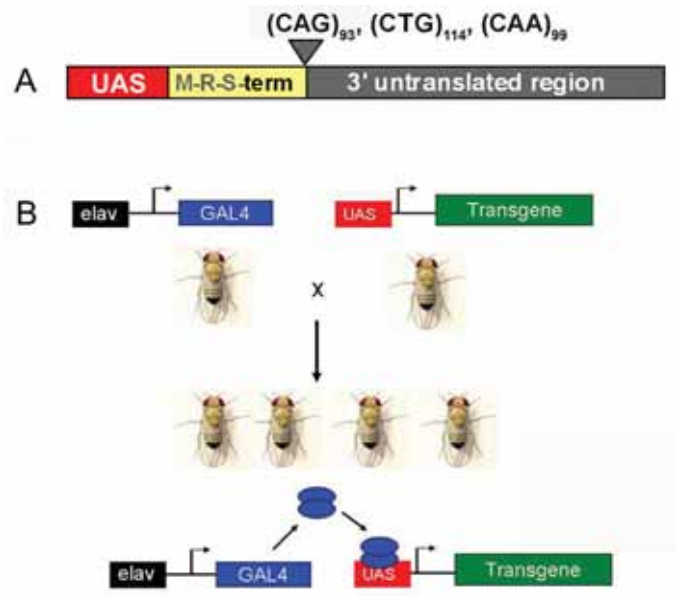


Figure 4.1: Obtaining *Drosophila* expressing untranslated repeats pan-neuronally for microarray and proteomic analysis. A) Constructs used to express expanded untranslated CAG, CUG and CAA repeats pan-neuronally in *Drosophila.* CAG, CUG and CAA repeat tracts are inserted into the 3' untranslated region of a short peptide (called rCAG. rCUG and rCAA respectively). **B)** Expression can be induced by GAL4 protein binding to the upstream activation sequence (UAS) region. In this case, pan-neuronal expression of each of the untranslated repeat constructs was achieved by driving expression with *elav-GAL4*. Crossing *elav-GAL4* flies to flies carrying the repeat insertions results in progeny expressing the expanded untranslated CAG, CUG or CAA repeats specifically in the nervous system.

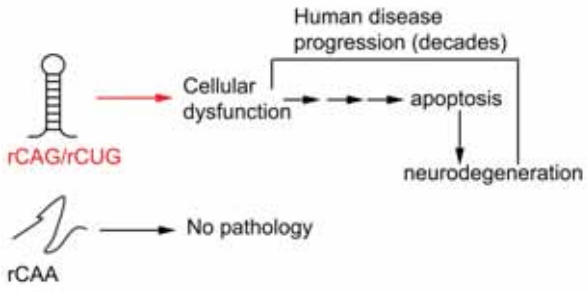


Figure 4.2: Rationale for use of this *Drosophila* model for investigation of early changes in expanded repeat disease. Expression of rCAG/rCUG repeats in this model should allow the identification of early changes caused by expression of hairpin repeats (red arrow). In the disease situation, there are likely to be many steps between the initial dysfunction and cellular death which take place over many decades. Many of these processes are unlikely to occur within the life-span of *Drosophila*, however by comparing the transcriptomes and proteomes of flies expressing rCAG/rCUG hairpin RNAs with those of flies expressing rCAA RNA, which cannot form a hairpin, or flies not expressing any ectopic RNA species (*elav>+)*, early changes that represent causative pathways leading to this long-term pathogenesis should be identified.

4.1 Identification of proteomic changes in neuronal cells expressing expanded repeat tracts

Proteomic analysis has not been extensively used in models of expanded repeat pathogenesis to date, with emphasis in many of the existing studies being on proteins found to aggregate with polyglutamine (178-179). We used a proteomic approach to identify changes caused by expression of rCAG and rCUG repeat RNAs in neurons of *Drosophila* with the aim of distinguishing hallmarks of expanded repeat pathogenesis. Proteomic analysis has the advantage that it can detect changes to both the abundance and post-translational state of proteins, therefore giving insight into mechanisms such as oxidative damage of proteins which have been proposed to play a role in polyglutamine disease (180).

Proteomic analysis was performed by the Adelaide Proteomics Centre (University of Adelaide, Australia). For each sample, total protein was extracted from heads of male flies. Samples were prepared from four biological replicates for each genotype and spread across a total of 8 2D mini gels, such that each genotype was represented on 4 separate gels as shown in Figure 4.3 A. Details of lines used can be found in Appendix A, Figure A1. Differential in-gel electrophoresis (DIGE) analysis was used to identify protein spots with a difference in abundance between genotypes. The comparisons performed between genotypes are represented in Figure 4.3 B&C. Two-tailed Student's t-tests were initially performed to identify spots with a significant difference in protein abundance in rCAG, rCUG and rCAA repeat expressing flies compared to *elav>+* control flies (Figure 4.4 A). A single spot which showed a significant change in abundance in rCAG and rCUG repeat expressing flies compared to the *elav>+* control but not in rCAA repeat expressing flies compared to the *elav>+* control was then selected from this list (Figure 4.4 B). Since there is some evidence to suggest that GAL4 is itself toxic to cells when it accumulates and can trigger apoptosis (137), the *elav>rCAA* comparison provides a control where the GAL4 present in the cells of the nervous system should be able to bind to the UAS sites of the rCAA construct, preventing accumulation. We therefore predict that driving expression of this CAA repeat RNA, which is unable to form a hairpin structure like the CAG and CUG repeats, should provide a control for the effects of GAL4 toxicity. The *elav>+* sample is included as a control for any effects that the presence of this untranslated CAA RNA may be having on the cells.

A

Gel Number	Cy2	Cy3	Cy5
1	IPS	elav>+ (A)	elav>rCAG (C)
2	IPS	elav>rCAA (B)	elav>rCUG (D)
3	IPS	elav>rCUG (B)	elav>+ (D)
4	IPS	elav>rCAA (D)	elav>rCAG (B)
5	IPS	elav>+ (C)	elav>rCUG (A)
6	IPS	elav>rCUG (C)	elav>rCAA (C)
7	IPS	elav>rCAG (A)	elav>+ (B)
8	IPS	elav>rCAG (D)	elav>rCAA (A)

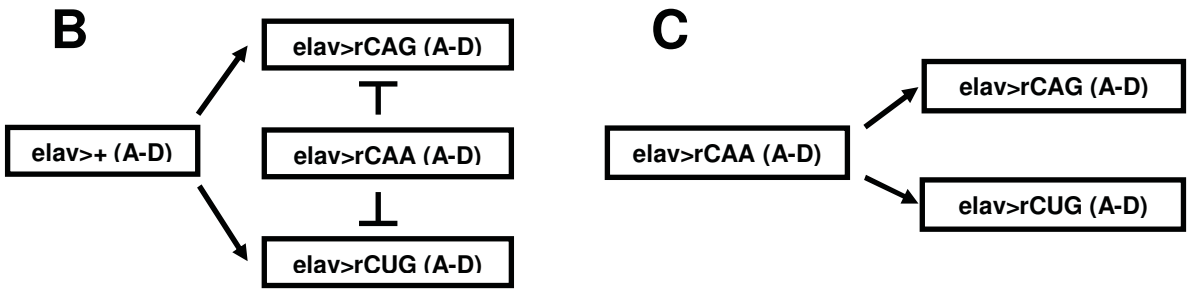
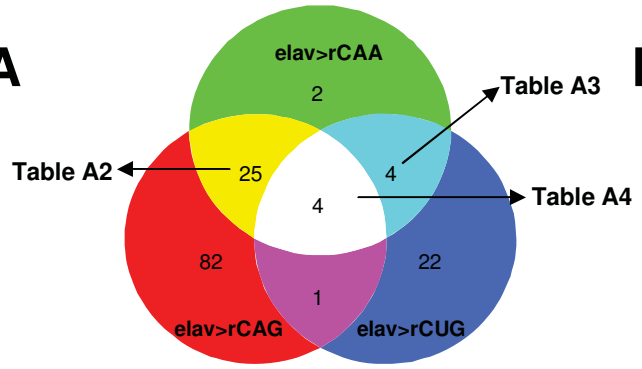


Figure 4.3: Overview of 2D-DIGE experiment procedure and analyses performed. **A)** IPS= Internal pooled standard. This sample is an equal mix of proteins from each of the 16 samples analysed on the gels (elav>+ A-D, elav>rCAA A-D, elav>rCAG A-D and elav>rCUG A-D) labelled with Cy2 and was run on every gel to allow comparisons and spot-matching to be performed between gels. Protein was extracted from 4 biological replicates for each genotype and 2 samples were labelled with Cy3 and 2 with Cy5. These samples were then spread across a total of 8 gels as shown. Gel 4 was selected as the master gel for this experiment, since it showed the largest number (2753) of properly resolved protein spots. Protein spots on all other gels were spot matched to the master gel. **B)** Initially, average spot ratio calculations and two-tailed Student's t-tests were performed for the four replicates for each of elav>rCAA, elav>rCUG and elav>rCAG compared to *elav>+*. A single spot that showed a significant change in abundance in both elav>rCUG and elav>rCAG but not elav>rCAA flies compared to *elav>+* was then identified by MS/MS (Figure 4.4). **C)** Average spot ratio calculations and two-tailed Student's t-tests were also performed directly comparing the *elav>rCUG* and *elav>rCAG* samples to *elav>rCAA*. This analysis also identified a single spot with a significant change in abundance which was identified by MS/MS (Figure 4.6).

A



B

genotype	Spot 1978	
	T-test	Fold Change
elav>rCAG	0.005	1.22
elav>rCUG	0.002	1.27

C

Spot number	Protein ID	Human orthologue	Accession Number	emPAI	% sequence MS/MS	Combined ion Score/Cut-off
1978	1-cys peroxiredoxin DPx-2540-1 (CG12405)	PRDX6	gi|12044363	0.87	37.7	434/58
	Alcohol dehydrogenase (CG3481)		gi|8282	0.30	18.8	225/58

Figure 4.4: Summary of changes in protein abundance detected in flies expressing rCAG, rCUG and rCAA RNA when compared to elav>+ control flies. A) Number of spots with a detected change in abundance when compared to the *elav>+* control. In each case, two-tailed students t-tests were performed on biological replicates and spots chosen with $P<0.05$. Spot IDs are listed in Appendix A as indicated. **B)** Spot 1978 showed a change in abundance in flies expressing rCAG and rCUG repeats but not rCAA repeats. Fold change was calculated from the average change in spot intensity compared to *elav>+* across the four gels for each genotype. Spots were selected for $P<0.05$. **C)** Spectra obtained from MS/MS were submitted to a MASCOT database search-engine. Two proteins, DPx-2540-1 and Alcohol dehydrogenase, returned combined ion scores above the cut-off, indicating 95% confidence in these proteins matching the MS/MS spectra and therefore suggesting that both proteins are likely to be present as a mix in this spot. Combined ion scores are calculated by summing the statistical score for each individual peptide match and excluding any redundant matches. DPx-2540-1 returned a combined ion score and % sequence coverage nearly twice that for Alcohol dehydrogenase and was also predicted to be nearly three times more abundant, as indicated by the exponentially modified protein abundance index (emPAI) scores. This suggests that DPx-2540-1 is more likely to be responsible for the observed change in spot abundance in this case, however neither protein can be ruled out.

4.2 Identification of proteins altered in *Drosophila* expressing rCAG or rCUG repeats pan-neuronally

Only one detected spot was significantly altered in flies expressing rCAG and rCUG repeats, but not rCAA repeats, compared to *elav>+* (Figure 4.4 B). This spot was excised from the master gel and identified by MS/MS (Figure 4.4 C). Two protein matches were found for the MS/MS spectra obtained for this spot by MASCOT search – Dpx-2540-1 (CG12405) and Alcohol dehydrogenase (CG3481) – both of which had combined ion scores that greatly exceeded the cut-off score of 58 required for 95% confidence in the match. This suggests that both proteins are likely to be present in spot 1978. The combined ion score is calculated by summing the statistical probability associated with all of the non-redundant peptide queries assigned to that protein match and therefore is a measure of the total level of support for that match. The sequence coverage and exponentially modified protein abundance index (emPAI) score give a measure of the percentage of the matched protein sequence covered by the peptides identified in MS/MS and a measure of the relative abundance of the protein predicted from the relative amount of each peptide identified by MS/MS respectively. The relative abundance score (emPAI) for DPx-2540-1 was more than double that of Alcohol dehydrogenase, suggesting that this protein is more likely to be responsible for the observed increase in spot intensity.

PRDX6, the human orthologue of DPx-2540-1, has been previously shown to protect against apoptosis (181), as well as assisting in maintenance of Ca^{2+} homeostasis (182). Induction has been demonstrated in both Alzheimer's patient brains – although in this case induction is seen only in astrocytes and not neurons (183) – and HD patient brains (180). It is a particularly unique member of the peroxiredoxin family in that it has both peroxidase and phospholipase activities; functions which may be particularly important in the brain because of its high lipid content (184). While regulation of human PRDX6 has been demonstrated to occur at the transcriptional level (185), there was no increase in DPx-2540-1 mRNA expression in flies expressing rCAG or rCUG repeats by quantitative real time PCR (data not shown), suggesting that transcriptional regulation is not responsible for the observed change in protein abundance. It is therefore likely that stabilisation or modification of the Dpx-2540-1 protein is responsible for any change in abundance in expanded repeat expressing flies.

In order to confirm a role for DPx-2540-1 in expanded repeat disease pathogenesis, we obtained a fly line which contains an insertion within the region of *DPx-2540-1*. This insertion, called MB01457, is a Minos element insertion from *Drosophila hydei* (186) located approximately 100 bp upstream of the annotated *DPx-2540-1* gene. The ability of the MB01457 insertion to reduce *DPx-2540-1* expression has not been demonstrated, however there is currently no RNAi line and no other insertion associated with *DPx-2540-1* available in *Drosophila*. Since expression of up to four transgene insertions of the untranslated repeat constructs (rCAG, rCAA and rCUG) in the eye does not elicit a phenotype, we examined the ability of the MB01457 allele to alter the phenotypes seen in the eye when a polyglutamine tract encoded by either a CAG or CAA repeat or a polyleucine tract encoded by a CUG repeat are expressed using the eye-specific *GMR-GAL4* driver (described in Chapter 1.4.1 and 1.4.2). Where altering expression of a candidate gene results in a modification of the phenotypes caused by all three of these repeats, we predict that this interaction is unlikely to be due to a specific interaction with hairpin repeat RNA since the CAA RNA is unable to form a secondary structure. A similar methodology has been successfully used to identify *mbl* as a candidate gene which is able to modify the phenotype caused by a polyglutamine tract encoded by a pure CAG repeat but not a mixed CAG/CAA repeat (117).

The phenotypes associated with expression of CAG or CAA-encoded polyglutamine or CUG-encoded polyleucine did not appear to be dramatically altered in flies heterozygous for the MB01457 insertion (Figure 4.5). Subsequent analysis of *DPx-2540-1* expression levels in the MB01457 insertion line revealed that levels of DPx-2540-1 RNA are not reduced in this stock compared to the wild-type w^{1118} stock used to generate the expanded repeat lines; in fact expression is consistently higher in the insertion line (Appendix A, Figure A5). A hypomorphic or null allele of *DPx-2540-1* could be generated by remobilising the Minos element to excise surrounding DNA (as described in (186)) however this was not pursued further in this study. Upregulation or modification of DPx-2540-1 may play a protective role in cells expressing expanded repeat RNA however, given that induction of PRDX6 is also seen in Alzheimer's patients, it seems unlikely that this is a primary change elicited specifically by expanded repeat RNA.

The effect of altering levels of the other protein identified in spot 1978, Alcohol dehydrogenase, on the expanded repeat phenotypes was also tested (Figure 4.5 G-I). In this case, an RNAi construct targeting Adh was co-expressed with CUG, CAG or CAA repeats in the eye. A slight change in the phenotype resulting from expression of polyglutamine encoded by either CAG or CAA was seen when these repeats were co-expressed with the Adh RNAi construct in the eye, consisting of a slight improvement in the ordered structure of the ommatidia but a decrease in pigmentation. A slight suppression of the rough eye phenotype resulting from expression of polyleucine encoded by CUG repeat was also seen. This effect is unlikely to be mediated through an interaction with hairpin RNA, since it is observed in flies expressing CAA repeat RNA which is unable to form a secondary structure.

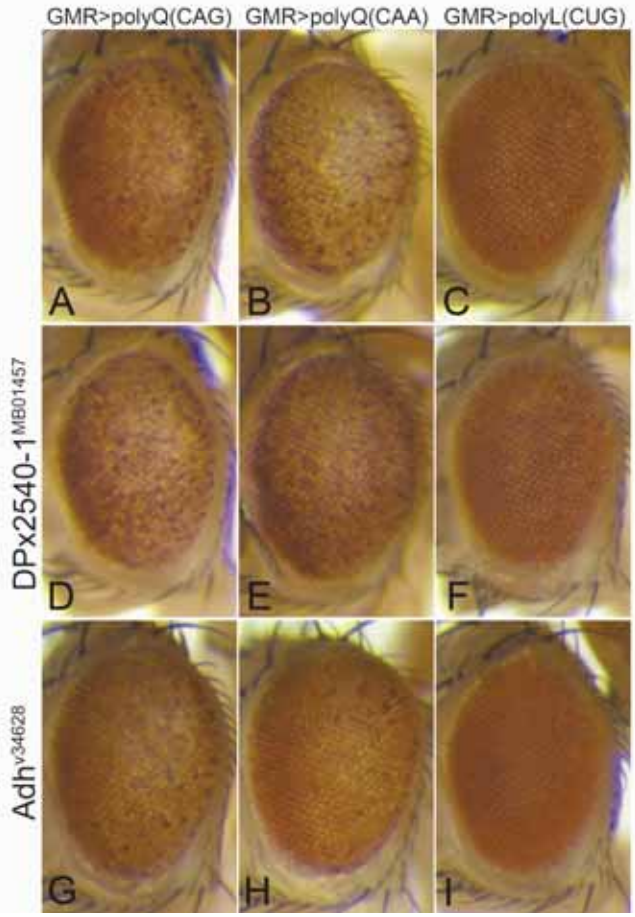


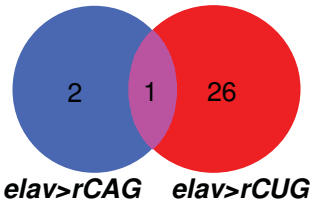
Figure 4.5: Modification of phenotypes resulting from expression of translated CUG, CAG and CAA repeats by insertion of a Minos element upstream of DPxr2540-1 or knocking down expression of Alcohol dehydrogenase. In all cases, repeat expression was driven by *GMR*-GAL4. **A&B)** Expression of polyglutamine encoded by either CAG or CAA results in loss of pigment in the eye. **C)** Expression of polyleucine results in a very mild rough eye phenotype. **D-F)** Expression of translated CAG, CUG or CAA repeats in a heterozygous MB01457 insertion background does not result in a dramatic alteration to the exterior appearance of the eye. **G-H)** A slight change in the polyglutamine phenotype is seen when an RNAi construct targeting Adh is co-expressed with either CAG or CAA repeats. **I)** Co-expression of an RNAi construct targeting Adh slightly suppresses the ommatidial disorganisation in flies expressing CUG-encoded polyleucine.

4.3 Identification of proteins altered in flies expressing rCAG or rCUG repeats compared to rCAA repeats.

Since analysis of spots altered in both *elav>rCAG* and *elav>rCUG* flies compared to the *elav>+* control resulted in detection of only 1 spot with a change in abundance which was not also altered in *elav>rCAA* flies, the 2D-DIGE data was re-

analysed comparing rCAG and rCUG repeat expressing flies directly to rCAA repeat expressing flies (comparison depicted in Figure 4.3 C). The resulting number of spots for each genotype compared is shown in Figure 4.6 A. This comparison also identified only one spot showing a change in abundance common to flies expressing rCAG and rCUG (Figure 4.6 A&B). The low number of changes identified common to flies expressing rCAG and rCUG repeats compared to either *elav>+* or *elav>rCAA* flies was unexpected. However, since these 2D mini-gels were only able to resolve between 2513 and 2753 spots per gel and *Drosophila* are predicted to have somewhere in the vicinity of 50,000 protein variants, the detected changes in spot abundance are likely to represent only a small proportion of changes to the proteome resulting from rCAG and rCUG repeat expression.

A



B

genotype	Spot 1101	
	T-test	Fold Change
elav>rCAG	0.001	-1.51
elav>rCUG	3.5E-4	-1.51

C

Spot number	Protein ID	Human orthologue	Accession Number	emPAI	% sequence MS/MS	Combined ion score/Cut-off
1101	nup62 (CG6251)	NUP62	gi|20130049	0.26	8.4	183/48
	cAMP-dependent protein kinase R2 (CG15862)	PRKAR2A	gi|17647815	0.08	7.7	128/48

Figure 4.6: Summary of changes in protein abundance detected when rCAG and rCUG repeat expressing flies were compared directly to *elav>rCAA* flies. A) Number of spots altered in *elav>rCAG* and *elav>rCUG* compared to *elav>*rCAA. In each case, two-tailed Student's t-tests were performed on replicates and spots chosen with P<0.05. All spots meeting these criteria are listed in Appendix A, Table A6-A7. **B)** Only one spot with changed abundance in both *elav>rCAG* and *elav>rCUG* flies compared to *elav>rCAA* was detected, which showed significantly decreased abundance. **C)** Summary of MASCOT search results for spectra obtained by MS/MS for spot 1101. Two proteins, Nup62 and cAMP-dependent protein kinase R2, returned combined ion scores above the cut-off, indicating >95% confidence in the protein match. Combined ion scores are calculated by summing the statistical score obtained for each individual peptide match and excluding any redundant matches. Nup62 was predicted to have greater relative abundance based on the frequency of the peptides matched to it in MS/MS as indicated by the emPAI score.

The single spot detected which showed a common change in abundance in both rCAG and rCUG expressing flies compared to rCAA expressing flies was also excised from the master gel and identified by MS/MS (Figure 4.6 B&C). There was a 1.51 fold decrease in abundance of this spot in both *elav>rCAG* and *elav>rCUG* flies. Two protein matches were found for the MS/MS spectra by MASCOT search – Nucleoporin 62 (Nup62) and cAMP-dependent protein kinase R2 (CG15862) – both of which greatly exceeded the cut-off score of 48 required for 95% confidence in the match. This suggests that spot 1101 contains a mixture of both proteins, however the emPAI score for Nup62 was much higher than for cAMP-dependent protein kinase R2 (0.26 compared to 0.08) indicating that this protein is likely to be more abundant in the spot. Given the higher relative amount of Nup62 in spot 1101, a decrease in abundance of this protein would be predicted to be more likely to be the cause of the observed change in spot intensity.

4.4 Evidence for involvement of nuclear transport in expanded repeat disease pathogenesis

NUP62 is a central component of the nuclear pore complex (NPC), the protein channel which spans the nuclear envelope and regulates transport in and out of the nucleus. Mutations in components of the NPC have been implicated in a number of diseases, known as the laminopathies, which have diverse symptoms depending upon the affected tissues (reviewed in (187)). Mutations in *NUP62* have been demonstrated to result in infantile bilateral striatal necrosis; a condition caused by degeneration of the basal ganglia and characterised by symptoms including involuntary movements, mental retardation, seizure and abnormal eye movements (188). NUP62 has also been reported to localise to protein aggregates in Alzheimer's disease (189) and polyglutamine expressing cells (190), suggesting that altered nuclear transport pathways play a role in some neurological diseases. Irregular localisation of the NPC, including NUP62, has also been described in anterior horn cells from spinal cords of patients with sporadic or familial ALS (191), a neurodegenerative disorder characterised by death of motor neurons. This observation supports a role for the NPC and nuclear transport in neuronal survival.

The nuclear pore also plays a role in stress response through extensive modifications and degradation of components of the NPC including NUP62 (192).

These modifications, which include phosphorylation and O-glycosylation, alter the interactions of the NPC with proteins both within the nucleus and the cytoplasm and thus inhibit nuclear export (192). In *Drosophila*, Nup62 has been shown to co-localise with the RNA binding protein Staufen2 (193) which has been implicated in nuclear export of mRNA (194). Data from *Xenopus* showing a direct interaction between mRNA and Nup62 during export (195) further supports a central role for Nup62 in mRNA export from the nucleus. A role for Nup62 in transcriptional regulation has also been suggested through an interaction with the transcription factor SP1 (196). Recent evidence from *Drosophila* suggests a central role for nucleoporins, including Nup62, in gene activation and silencing through interactions with chromatin. While these interactions can occur at the NPC, they have also been found to occur in the nucleoplasm (197-198) suggesting that nucleoporin function is not limited to the nuclear envelope.

In order to verify an interaction between Nup62 and CAG and CUG repeat RNA, the ability of an RNAi construct targeting nup62 or an overexpression construct encoding the *Drosophila* nup62 cDNA to alter the appearance of the eye when introduced into flies expressing a polyglutamine tract encoded by either a CAG or CAA repeat or a polyleucine tract encoded by a CUG repeat was tested. Expression of an RNAi construct targeting nup62 with *GMR-GAL4* did not cause a phenotype in the eye alone, but consistently caused a mild enhancement of the phenotypes resulting from expression of CAG-encoded polyglutamine and CUG-encoded polyleucine (Figure 4.7 A&C compared to D&F). No significant change to the exterior appearance of the eye was observed when the same RNAi construct was co-expressed with polyglutamine encoded by a CAA repeat (Figure 4.7 B compared to E). This suggests that Nup62 is able to modify expanded repeat pathology in our *Drosophila* model and that, since this interaction is not observed when the non-hairpin forming CAA repeat is expressed, this is likely to be sequence-dependent effect occurring at the RNA level. In further support of a role for Nup62 in expanded repeat pathogenesis, expression of an RNAi construct targeting the transcript encoding the other protein identified in spot 1101, cAMP-dependent protein kinase R2 (CG15862), did not cause a significant change to the appearance of the eye in flies co-expressing polyglutamine encoded by a CAG or CAA repeat or polyleucine encoded by a CUG repeat (Figure 4.7 G-I). This suggests that while there may be a change in expression or modification of both Nup62 and cAMP-

dependent protein kinase R2 in flies expressing expanded repeat RNA, Nup62 is likely to be involved in a rate-limiting step in pathogenesis in this model.

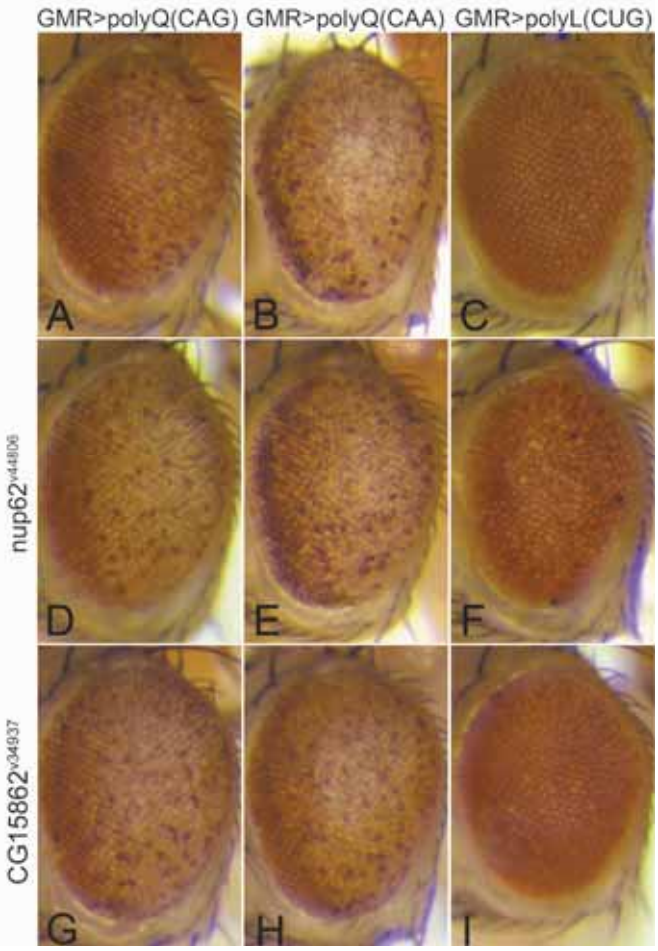


Figure 4.7: Co-expression of an RNAi construct targeting nup62 enhances CAG and CUG but not CAA eye phenotypes. Both RNAi construct expression and repeat construct expression are driven by *GMR-GAL4*. Two independent insertion lines were tested for this RNAi construct and the same result was observed in both cases. **A&B)** Expression of polyglutamine encoded by either CAG or CAA results in a loss of pigment eye phenotype. **C)** Expression of polyleucine results in a mild rough eye phenotype. **D)** Co-expression of an RNAi construct targeting nup62 with polyglutamine encoded by a CAG repeat results in a consistent increase in the area of the eye displaying loss of pigment. **E)** Co-expression of the same RNAi with polyglutamine encoded by a CAA repeat does not cause a change to the exterior appearance of the eye. **F)** Co-expression of the nup62 RNAi construct enhances the polyleucine eye phenotype, resulting in an increased area and severity of roughness, a decrease in the size of the eye and the appearance of necrotic patches. **G&H)** Co-expression of an RNAi construct targeting CG15862 (cAMP-dependent protein kinase R2) with polyglutamine encoded by either a CAG or CAA repeat results in indistinguishable eye phenotypes. **I)** Co-expression of an RNAi construct targeting CG15862 with polyleucine encoded by a CUG repeat does not significantly alter the exterior appearance of the eye.

Since Nup62 is involved in RNA export from the nucleus, the modification of the eye phenotype in *Drosophila* may be mediated via a direct interaction between Nup62 and the CAG and CUG repeat RNA. Nuclear retention of CUG repeat-containing RNA has been described in models of DM1 as well as in patient tissue (72, 76, 81) and is thought to be necessary for pathogenesis. In support of this, the formation of cytoplasmic foci has been shown to be insufficient to elicit toxicity in a mouse model (199) and forcing RNA out of the nucleus in myoblasts in a cellular DM1 model is able to reduce some of the muscle differentiation phenotypes associated with CUG repeat expression (200). It is unclear whether nuclear retention is the cause of sequestration

of key RNA binding proteins including MBNL1, or a result of these interactions. One key piece of evidence to support a role for RNA binding proteins in preventing CUG repeat RNA from leaving the nucleus is that reducing levels of one component of these foci, the RNA binding protein HnRNP-H, rescues nuclear retention in DM1 cells (201). Interestingly, HnRNP-H itself has been shown to play a role in shuttling between the nucleus and cytoplasm (202). The finding that altering Nup62 levels modifies phenotypes associated with expanded repeat expression in *Drosophila* supports an important role for nuclear retention in expanded repeat disease pathogenesis.

In order to further investigate the sequence-dependent interaction of Nup62 and CAG and CUG repeat RNA, an overexpression Nup62 construct was generated. *GMR-GAL4* driven overexpression of Nup62 alone did not result in any disruption to the *Drosophila* eye. Interestingly, overexpression of *Drosophila* Nup62 in flies co-expressing CUG, CAG or CAA repeats resulted in a reduction in the severity of the eye phenotype in all cases (Figure 4.8 D-F). This result is not consistent with the sequence-dependent effect seen when levels of Nup62 were reduced. To confirm that the absence of a phenotypic modification in flies co-expressing polyglutamine encoded by a CAA repeat with the nup62 RNAi construct was not a result of an insertional effect in a single transgenic line, an independent CAA repeat line was subsequently tested. Expression of the CAA-encoded polyglutamine tract in this line with *GMR-GAL4* gives a strong loss of pigment eye phenotype and no change in the appearance of the eye was seen when a nup62 RNAi construct was co-expressed (data not shown). Similarly, a second candidate gene line with an independent insertion of the nup62 RNAi construct was tested and also slightly enhanced the CAG and CUG repeat phenotypes but not the CAA repeat phenotype. This suggests that the absence of an interaction with CAA-encoded polyglutamine when Nup62 levels are reduced is not due to an insertional effect of a single transgenic line, but a more general property of this repeat tract. The interaction with CAG and CUG repeats is therefore likely to be occurring at the RNA level and is dependent upon the ability of the expanded repeat RNA to form a hairpin secondary structure, a property which CAA repeat RNA lacks.

The localisation of Nup62 to protein aggregates in polyglutamine diseases (190) suggests that there may be a direct physical interaction between polyglutamine

tracts and nuclear pore components. Since increasing levels of Nup62 in *Drosophila* expressing polyglutamine tracts reduces toxicity in the eye, it appears that the localisation of Nup62 to polyglutamine aggregates may play a role in pathogenesis. This does not rule out a separate role for Nup62 in RNA-mediated toxicity; a role which is strongly supported by the fact that the alteration in Nup62 levels was originally observed in *Drosophila* expressing untranslated expanded repeat tracts. In order to further investigate the interaction between expanded repeat RNA and Nup62, the effect of reducing Nup62 levels in flies expressing expanded untranslated rCAG, rCUG and rCAA repeats was also tested. Co-expression of the nup62 RNAi construct with four transgene insertions of the rCAG, rCUG and rCAA constructs did not result in a disruption to the exterior appearance of the eye (data not shown) and therefore the effects of altering Nup62 levels on the polyglutamine component of pathogenesis and RNA-mediated toxicity were not able to be further separated using this approach.

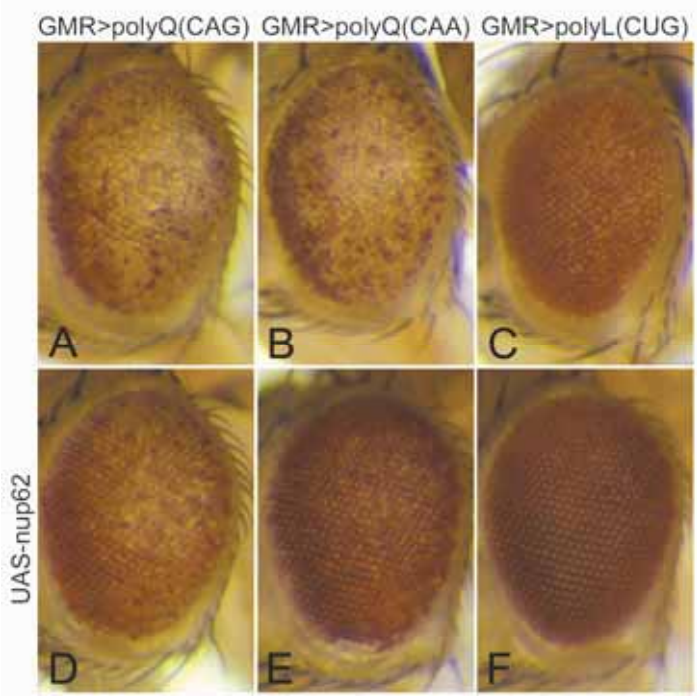


Figure 4.8: Overexpression of Nup62 in the *Drosophila* eye suppresses both polyglutamine and polyleucine eye phenotypes. Both Nup62 overexpression and repeat construct expression are driven by *GMR*-GAL4. **B)** Expression of polyglutamine encoded by either CAG or CAA results in a loss of pigment eye phenotype. **C)** Expression of polyleucine results in a very mild rough eye phenotype. **D-E)** Ectopic expression of Nup62 in the eye resulted in a considerable suppression of both the CAG and CAA-encoded polyglutamine eye phenotypes. In both cases, there is a reduction of the area and severity of loss of pigment. **F)** Ectopic expression of Nup62 with polyleucine completely suppressed the rough eye phenotype and increased the amount of pigment in the eye.

4.5 Summary of proteomic changes elicited by expression of CAG and CUG repeats in neurons of *Drosophila*

A small number of changes in protein abundance were observed which were common to flies expressing rCAG and rCUG repeat RNA compared to rCAA repeat expressing flies and the *elav>+* control. This may be a result of the limited number of proteins properly resolved and of sufficient abundance to be detected on the 2D mini gels used for this experiment. There were also a number of protein spots which showed unique changes in abundance in either rCAG or rCUG repeat expressing flies which may represent sequence-dependent components of RNA pathogenesis, however these proteins were not further investigated in this study.

MS/MS analysis of the 2 spots which showed a common change in abundance in rCAG and rCUG repeat expressing flies identified 2 candidate proteins in each case. Further genetic analysis of one candidate protein, Nup62, showed that a reduction in expression was able to modify phenotypes resulting from expression of translated CAG and CUG repeats but not CAA repeats in *Drosophila*. This suggests that this interaction may be occurring at the RNA level, since both CAA and CAG repeats code for polyglutamine at the protein level but only the CAG repeat RNA is able to form a hairpin secondary structure. The identification of Nup62, a central component of the nuclear pore, as a modifier of expanded repeat pathogenesis in *Drosophila* suggests that nuclear transport may play a central role in pathogenesis. Since Nup62 is known to play a role in mRNA export from the nucleus, reducing levels of Nup62 could modify the toxicity of expanded repeat RNA by altering its localisation and therefore the proteins with which it interacts.

The observation that overexpression of Nup62 is able to suppress both CAG and CAA-encoded polyglutamine phenotypes may indicate that the nuclear pore also assists in clearance of polyglutamine aggregates from the nucleus. This does not appear to be a rate-limiting step in pathogenesis, since the CAA-encoded polyglutamine phenotype was not altered when Nup62 levels were reduced. These results point to a central role for nuclear transport pathways in pathogenesis of the expanded repeat diseases which, in the case of the polyglutamine diseases, may involve components of both polyglutamine and RNA-mediated toxicity. However it is also possible that the effects observed in this assay are specific to the eye, which

consists of a mixture of both neuronal and non-neuronal cells, and therefore a role for nuclear transport in pathogenesis should also be validated in a neuronal assay. Since very few protein spots were detected which showed a change in abundance in both rCAG and rCUG repeat expressing flies, identification of proteins which are altered in either rCAG or rCUG will also be a focus of further experiments.

Chapter 5: Identifying pathogenic pathways of expanded repeat disease by microarray analysis

A number of studies have attempted to establish common and unique pathways of expanded repeat pathogenesis through microarray analysis of various disease models (203-208). In most cases, these studies have aimed to identify disease-specific transcriptional changes through the use of repeat tracts in the context of the different expanded repeat disease genes. There is evidence that in both the polyglutamine diseases (109-110) and untranslated repeat diseases (117), the repeat-encoded peptides or expanded repeat RNAs respectively are intrinsically toxic, suggesting that there are likely to also be context-independent pathways of pathogenesis in the expanded repeat diseases.

Our model differs from those previously studied by microarray analysis in that the repeat tracts are encoded within a short peptide and are not in the context of the transcripts in which they are normally found. This study tests the outcomes of expression of hairpin-forming repeat sequences (rCAG and rCUG), with the aim of identifying components of expanded repeat pathogenesis which are the result of RNA toxicity of the repeat sequences themselves. This approach also allows direct comparison of the cellular outcomes resulting from expression of different repeat sequences, which will provide information on both sequence-dependent and sequence-independent features of expanded repeat disease pathogenesis. Another difference between this model and many that have been previously analysed is the lack of severe degeneration apparent in this model. This analysis should therefore allow the identification of transcriptional changes which are hallmarks of early cellular dysfunction in disease, rather than the result of induction of apoptotic pathways, and therefore should represent causative components of pathogenesis rather than the downstream effects (as depicted in Figure 4.2).

5.1 Identification of transcriptional changes in neuronal cells expressing expanded repeat tracts: microarray experiment 1

Microarray experiments were performed using Affymetrix *Drosophila* 2.0 arrays on RNA extracted from the heads of newly eclosed male *Drosophila* expressing the rCAG, rCUG and rCAA constructs. Crosses to generate these flies were performed as depicted in Figure 4.1. In microarray experiment 1, repeat constructs were under the control of the same *elav–GAL4* driver used in proteomic analysis (*elav*^{c155}*–GAL4*) which consists of a P-element insertion within the promoter of the endogenous *elav* gene, such that GAL4 is expressed in the same pattern as the endogenous ELAV protein (209). To increase repeat RNA expression levels, recombinant chromosomes were generated each carrying two UAS-repeat transgene insertions. Three fully independent transgenic lines were tested for each of the repeat sequences (rCAG, rCUG and rCAA) to overcome the possibility of insertional effects in individual lines impacting on the results. Details of lines analysed in this experiment are shown in Appendix B, Table B1. All microarray experiments were kindly performed by Gareth Price and in collaboration with Deon Venter (Pathology, Mater Health Services, South Brisbane QLD).

Comparisons of microarray data were performed as represented in Figure 5.1. In each comparison, genotypes were filtered for transcripts which returned a "present" call as determined by the Affymetrix Expression Console™ Software. To return a present call, the hybridised spot must show significantly higher signal than the background measured across the entire chip (P-value <0.001). While these criteria are likely to result in the exclusion of some genes which are very lowly expressed, including genes in analysis which do not return a present call can result in a false representation of the degree of change between samples when fold change is calculated, since the signal is likely to be more variable at the lower end of the detectable scale. Excluding genes which do not give a signal above the Affymetrix detection threshold should decrease the false-positive detection rate and produce a more robust data set for further analysis. Nevertheless, it is likely that performing the analysis in this way will exclude genes which are completely "off" in either genotype being compared. This set of genes may warrant further investigation in a secondary analysis of the data.

Two-tailed Student's t-tests were then performed for each comparison represented in Figure 5.1 and lists compiled of genes which showed significantly altered (P<0.05) expression in rCAG or rCUG repeat expressing flies compared to either *elav>+* (the *elav–GAL4* driver out-crossed to the w^{1118} wild-type line) or *elav>rCAA* flies. Expression ratios of these genes were then calculated based on the average Affymetrix chip signal which was determined from all 3 lines (or 2 biological replicates in the case of *elav>+*) analysed for each genotype. This list was then further filtered for genes which gave a value for log2 of the expression ratio – defined as "log2(ratio)" – >0.5 or <-0.5; that is genes with a fold change greater than approximately ±1.4. The resulting number of genes for each comparison is represented in Figure 5.3 A.

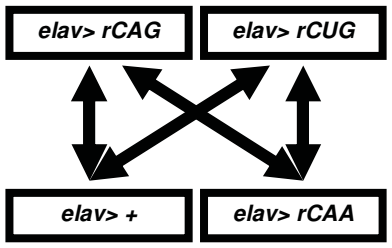


Figure 5.1: Overview of microarray experiments. Fly lines used for each genotype are listed in Appendix B, Table B1. Comparisons were performed between *elav>rCAG* and *elav>rCUG* flies and each of *elav>rCAA* and *elav>+* as shown. In all cases, only probe-sets which were called "present" in both genotypes under comparison were included in further analysis. For each comparison, Log2 of the ratio for each probe-set - log2(ratio) - was calculated from the average chip signal of all lines tested for each genotype. Final lists of genes with a Student's t-test value of P<0.05 and log2(ratio)>0.5 or <-0.5 were then generated for each comparison.

As described for the proteomic analyses (Chapter 4), transcriptional profiles of rCAG and rCUG repeat-expressing flies were compared to both *elav>+* and *elav>rCAA* flies in order to account for both toxicity resulting from accumulation of GAL4 protein in the nervous system and any effects that expression of the rCAA RNA may have. The complete gene lists for each of these analyses can be found in Appendix B (Table B3-B6). Gene ontology analysis was then performed on each of these gene lists to identify cellular pathways which may be disrupted by the expression of rCAG or rCUG hairpin repeats (Figure 5.4 A-D).

5.2 Validation of cellular changes by independent microarray experiment: Microarray experiment 2

A second microarray experiment was performed to provide an independent validation of transcriptional changes in flies expressing expanded repeat RNAs.

Analysis was performed as described for microarray experiment 1, except that in this case flies were expressing four transgene insertions of rCAG, rCUG or rCAA under the control of a different *elav-GAL4* driver. This *elav*-GAL4 line consists of a P-element insertion on the second chromosome encoding GAL4 under the control of the promoter region of the *Drosophila elav* gene (210), and therefore these flies have a normal endogenous copy of *elav* on the X chromosome (represented in Figure 5.2). Two completely independent four transgene insertion lines were available for each repeat (listed in Appendix B, Table B2). Comparisons were made using both this *elav-GAL4* driver out-crossed to the wild-type w^{1118} line as a control (*elav>+*) and *elav-GAL4* driving an untranslated CAA repeat (*elav>rCAA*), as described in 5.1. The number of genes significantly altered in each comparison (log2(ratio)>0.5 or <-0.5, P<0.05) is shown in Figure 5.3 B. Complete lists of genes for each comparison can be found in Appendix B (Table B7-B10). Gene ontology analysis was again performed on each of these gene lists (Figure 5.4 E-H).

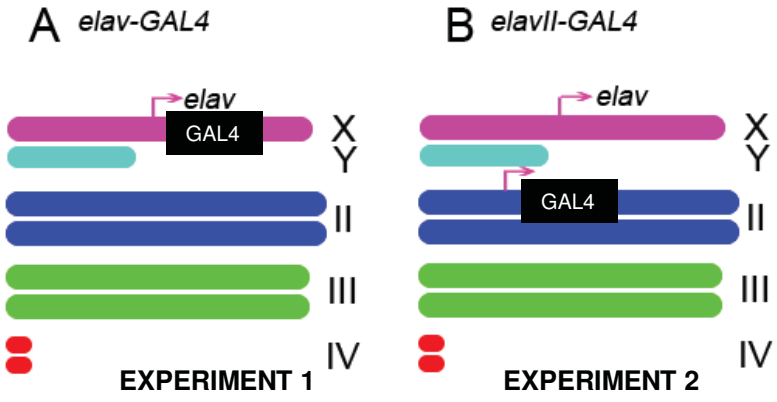


Figure 5.2: comparison of the two *elav-GAL4* driver lines used in this study. Pictures represent the chromosomes of a male *Drosophila*. **A)** The *elav-GAL4* line contains an insertion within the endogenous *elav* gene such that GAL4 is expressed in cells where *elav* would normally be expressed. **B)** The *elavII-GAL4* line has an insertion on chromosome II which contains the *elav* promoter controlling expression of GAL4. In this case, the flies have a normal copy of the *elav* gene.

5.3 Comparison of microarray experiment 1 and 2

A similar number of genes are altered in each comparison in microarray experiment 2 as in microarray experiment 1 (Figure 5.3 A compared to B), however there is no overlap in genes changed when flies expressing the rCUG repeat are compared to both *elav>rCAA* and *elav>+* control flies, as compared to 26 changing genes for the same comparison in the first microarray experiment. The corresponding comparison for *elav>rCAG* results in detection of 8 transcripts in both microarray

experiment 1 and 2, however none of these transcripts are common to both experiments (Table 5.1 and 5.5 respectively). Furthermore, only 9 genes are changed in both *elav>rCAG* and *elav>rCUG* flies compared to *elav>rCAA* in microarray experiment 2 while 38 changing genes were detected for the same comparison in microarray experiment 1. Genes common to microarray experiment 1 and 2 for each comparison are represented in Figure 5.3 C.

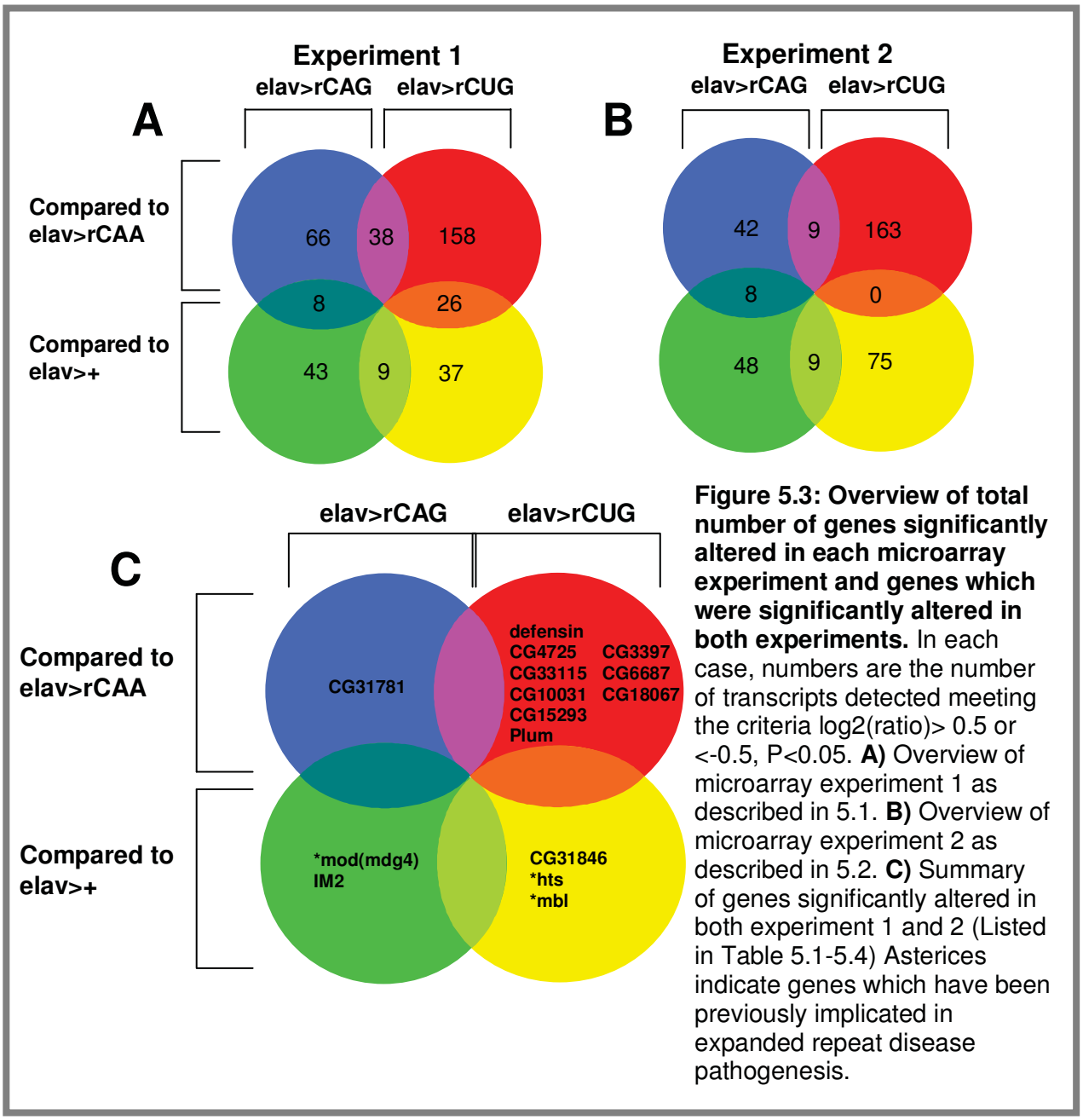


Figure 5.3: Overview of total number of genes significantly altered in each microarray experiment and genes which were significantly altered in both experiments. In each case, numbers are the number of transcripts detected meeting the criteria log2(ratio)> 0.5 or <-0.5, P<0.05. **A)** Overview of microarray experiment 1 as described in 5.1. **B)** Overview of microarray experiment 2 as described in 5.2. **C)** Summary of genes significantly altered in both experiment 1 and 2 (Listed in Table 5.1-5.4) Asterices indicate genes which have been previously implicated in expanded repeat disease pathogenesis.

There are several possible explanations for the differences seen between the two experiments. Firstly, it is likely that there are some transcripts being detected as significantly altered in one experiment which fall below the set detection threshold for the other experiment and are therefore not included in analysis. It should also be noted that the driver line used to express the repeats could have an impact on the changes seen, since we have some evidence to suggest that the two *elav-GAL4* drivers used in this study have slightly different expression patterns (K. Lawlor, unpublished data). Since we are increasing the amount of repeat RNA expression in experiment 2 by driving expression of four insertions of the repeat constructs compared to two insertions in experiment 1, it is also possible that these differences demonstrate real dose-dependent effects of repeat RNA expression. The purpose of these analyses was not to produce a comprehensive list of all changes resulting from expanded repeat expression, but to gain insight into the "categories" of changed transcripts and therefore pathways which may be important in expanded repeat disease pathology.

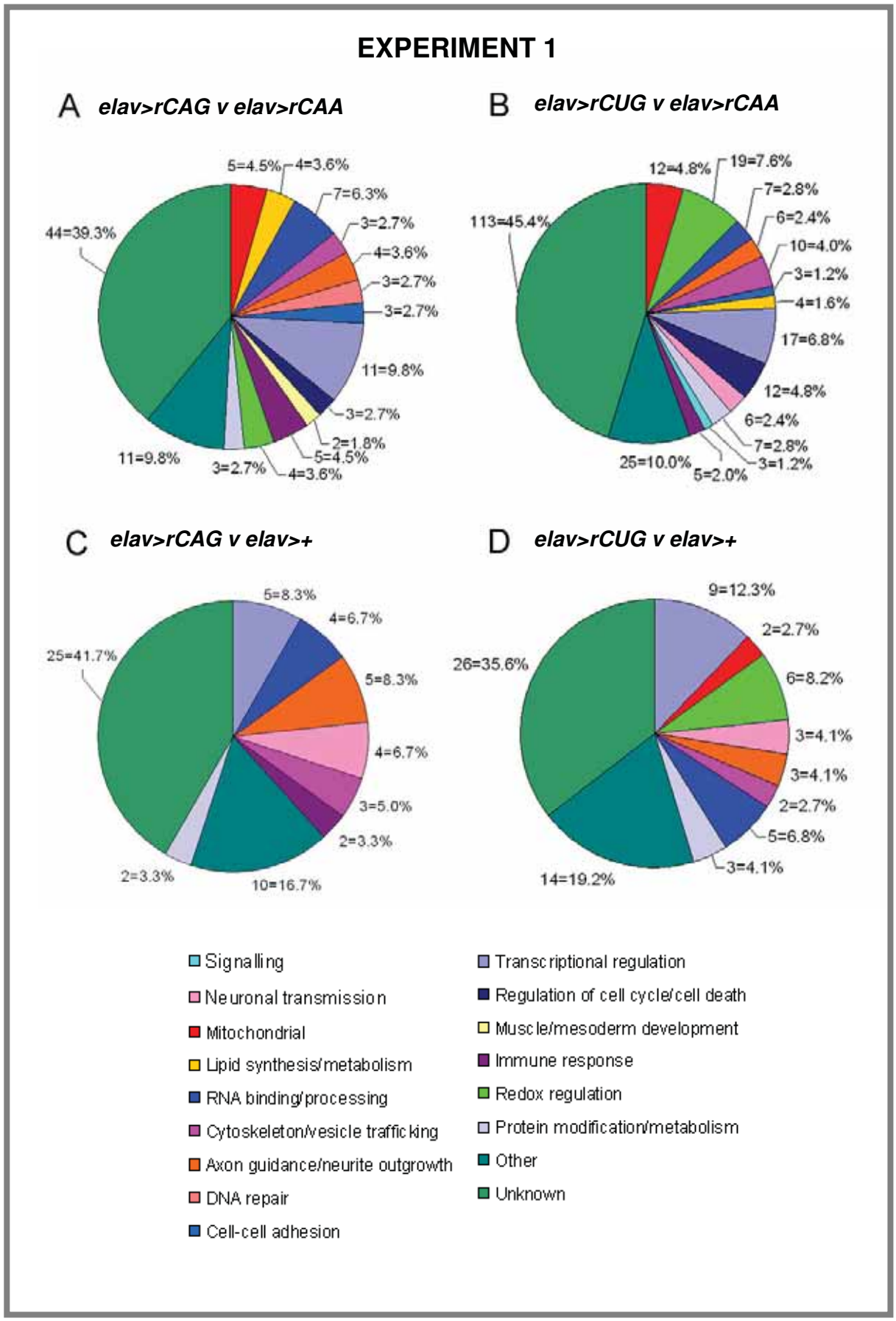
5.4 Gene ontology analysis of genes altered in rCAG and rCUG repeat-expressing flies compared to *elav*>rCAA flies

Lists were compiled of genes with altered expression in flies expressing either rCAG or rCUG repeats compared to rCAA repeats in both experiment 1 and 2 as described in 5.1 (Full gene lists can be found in Appendix B, Table B3-B4 & B7-B8), with the aim of identifying key cellular processes which are specifically perturbed by the expression of hairpin-forming repeat RNA. Genes were then grouped into categories based upon their known or predicted function. Each comparison resulted in lists in which more than 30% of the detected genes have unknown function and therefore do not provide additional information on cellular pathology. There are also a considerable number of genes which were placed in the "other" category, indicating that they do not fit into one of the gene ontology pathways listed. Together, the "unknown" and "other" categories make up nearly 50% of the listed genes when *elav>rCAG* flies are compared to *elav>rCAA* flies and more than 50% of the listed genes for the other comparisons (Figure 5.4 A,B,E&F).

Virtually all gene ontologies listed overlap between comparisons performed on *elav>rCUG* flies in both experiments and *elav>rCAG* flies in experiment 1, suggesting that there are repeat sequence-independent effects of hairpin RNA expression. Analysis of the transcriptional profile of *elav>rCAG* flies in experiment 2 resulted in detection of less than half the number of genes altered as was seen for the *elav>rCUG* genotype in either experiment or *elav>rCAG* in experiment 1. This suggests that there was a large amount of variation in the transcriptional profiles resulting from pan-neuronal expression of the independent four transgene insertion rCAG lines analysed in experiment 2, which may indicate location-dependent effects of the transgene insertion sites. Nevertheless, expression of rCAG RNA did consistently result in altered expression of genes involved in “lipid synthesis/metabolism”, “cytoskeleton/ vesicle trafficking” and “RNA binding/ metabolism” which may indicate that these are components of CAG repeat RNA pathogenesis.

There are a large number of gene ontologies represented in each experiment when transcript levels in *elav>rCUG* flies are compared to *elav>rCAA*, each constituting only a small percentage of the total genes. The largest group of genes changed in each experiment belongs to the “redox regulation” category and of the nine genes common between the two experiments, most are also involved in stress response (Table 5.2) further suggesting that some of the primary cellular changes in these cells may involve a response to cellular stresses caused by the presence of the rCUG RNA. “Transcriptional regulation” is also highly represented in both experiment 1 and 2 when *elav>rCUG* flies are compared to *elav>rCAA*, but only in experiment 1 when *elav>rCAG* flies are compared to *elav>rCAA* (Figure 5.4 A, B & F).

EXPERIMENT 1



EXPERIMENT 2

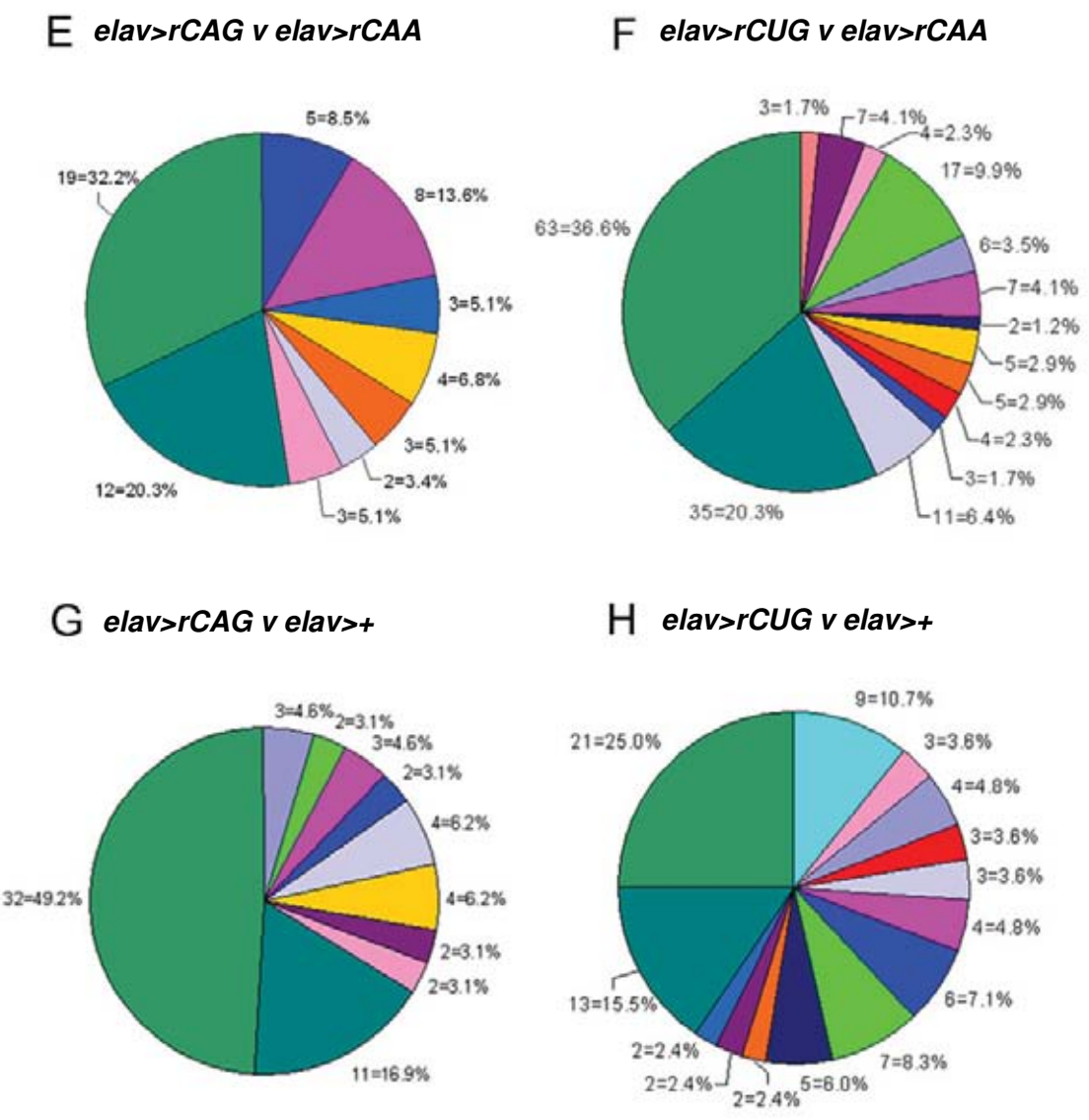


Figure 5.4: Gene ontology analysis of genes which were significantly altered in *Drosophila* expressing rCAG or rCUG RNA pan-neuronally. Genes were selected for Log2(ratio) >0.5 or <-0.5, P<0.05. The number of genes placed in each category and the percentage of the total number of genes which they represent are indicated. Categories are as listed in figure legend. Genes in the "unknown" category have no known function. Gene ontology was determined either from known phenotypic data or homology with other genes of known function. Genes in the "other" category have a known or suspected function that does not fit into one of the listed categories.

5.5 Gene ontology analysis of genes altered in rCAG and rCUG repeat-expressing flies compared to *elav>+* flies

Analysis of changes detected in flies expressing rCAG and rCUG repeats compared to *elav>+* flies was also performed for both experiment 1 and 2 (full lists can be found in Appendix B, Table B5-B6 & B9-B10). This comparison should also identify changes specifically caused by expression of hairpin-forming CAG and CUG repeat RNA, however it is possible that some changes that are detected may also be the result of accumulation of GAL4 protein. In experiment 1, a similar pattern of categories of altered genes was detected when the transcriptional profiles of *elav>rCAG* and *elav>rCUG* flies were compared to *elav>rCAA* flies as was seen in the comparison to *elav>+*, with over 50% of genes falling into the "unknown" or "other" categories (Figure 5.4 C & D). Of the remaining genes, "transcriptional regulation" was again highly represented, with 8.3% of the total genes altered in *elav>rCAG* flies and 12.5% of those altered in *elav>rCUG* flies falling into this category (Figure 5.4). Again, there was nearly complete overlap between categories of genes altered in rCAG and rCUG repeat-expressing flies compared to *elav>+*, supporting the idea that there may be a component of repeat sequence-independent pathogenesis in this model.

A number of studies have identified transcriptional dysregulation as a feature of the polyglutamine diseases including HD (203, 205), SCA7 (206), DRPLA (203) and SCA3 [8, 9]. Other categories of genes identified in this study, including "mitochondrial processes" (211-212), "RNA processing/metabolism" (73, 213), "cytoskeleton/vesicle trafficking" (11, 13, 205, 214-217), "lipid metabolism" (205) and "neuronal transmission" (206-207, 218-219), have also been previously implicated in pathogenesis in polyglutamine models. It therefore seems likely that there is a significant degree of sequence-independent pathogenesis occurring in the expanded repeat diseases, since both rCUG and rCAG repeat RNAs are able to induce alterations to these pathways. This result also suggests that the expanded repeat-containing RNA itself may be involved in inducing at least some of the early changes observed in the polyglutamine diseases. The reported ability of the untranslated CUG repeat RNA in DM1 to induce transcriptional dysregulation (220) further supports the ability of repeat-containing RNA alone to induce significant perturbations to cellular homeostasis.

Analysis of categories of genes altered in *elav>rCAG* flies compared to *elav>+* in experiment 2 provided limited information as nearly 2/3 of the listed genes fall into the "unknown" or "other" categories (Figure 5.4). Of the remainder, "lipid synthesis/ metabolism" and "protein modification/metabolism" are most highly represented, however each of these groups only contains four genes (6.2% of the total number) and therefore this is fairly inconclusive. The corresponding comparison of *elav>rCUG* flies to *elav>+* (Figure 5.3 H) gives a similar set of gene categories to the comparison to *elav>rCAA* (Figure 5.3 F), although there is no overlap in the specific genes detected. The lack of gene overlap between the two comparisons may not necessarily indicate that different cellular processes are disrupted, but merely that the particular genes which exceed the threshold for detection are different. Alterations to "signalling" (10.7%), "redox regulation" (8.3%) and "RNA binding/metabolism" (7.1%) are most highly represented, again suggesting that expression of CUG repeat RNA is able to induce stress responses in neuronal cells. There is also a large degree of overlap in the categories of gene changes detected in microarray experiment 1 and microarray experiment 2 confirming that these are likely to be real effects of expressing CUG repeat RNA in the neurons of *Drosophila*.

While the comparison of categories of genes altered in rCAG and rCUG repeat expressing flies identified a number of processes which have been previously implicated in expanded repeat pathogenesis, this type of analysis proved to be of limited use in ascertaining the primary changes occurring as a result of repeat expression in our *Drosophila* model. There was no clear bias towards perturbation of a particular category of genes, but rather an indication of broad cellular dysfunction in flies expressing either of these repeat RNAs. Analysis of particular genes which were commonly altered in different comparisons was therefore performed in order to identify key transcriptional changes elicited by hairpin RNA expression in *Drosophila*.

5.6 Analysis of genes significantly altered in both microarray experiment 1 and 2

In order to identify transcriptional changes which represent key effects elicited by rCAG and rCUG repeat expression, genes which were consistently altered in both experiment 1 and 2 were more closely investigated. These genes are likely to represent specific effects of repeat RNA expression, since they are altered regardless of which *elav-GAL4* driver and specific transgenic lines are used to express expanded repeat RNA. These genes are listed in Tables 5.1-5.4. Only one gene was found to be commonly altered when the transcriptional profile of rCAG expressing flies was compared to rCAA expressing flies. This gene, CG31781, is a predicted lipase with no other associated functional information (Table 5.1). A similar comparison performed for rCUG expressing flies compared to rCAA expressing flies identified 9 commonly altered genes (Table 5.2). Little functional information is available for these genes, however a number of them are involved in stress response mechanisms which is consistent with what was observed in the gene ontology analysis.

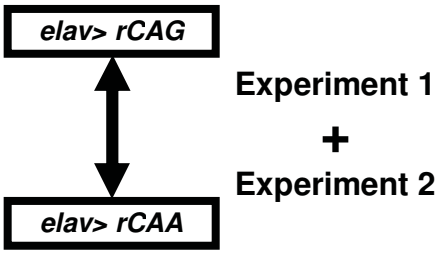
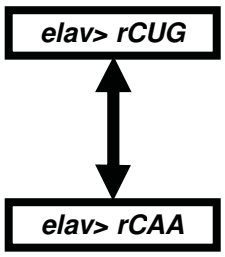


Table 5.1: Common changes for elav>rCAG compared to elav>rCAA in experiment 1 and 2.

Log2(ratio) >0.5 or <-0.5, P<0.05.

Gene Title	Gene Symbol	Ensembl	Experiment 1: log2 (ratio)	Experiment 2: log2 (ratio)	Human Orthologue	Function
CG31781	CG31781	CG31781	0.61 P=0.037	0.62 P=0.011	LIPF-002	triacylglcerol lipase



Experiment 1

+

Experiment 2

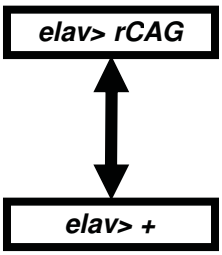
Table 5.2: Common changes for elav>rCUG compared to elav>rCAA in experiment 1 and 2.

Log2(ratio) >0.5 or <-0.5, P<0.05.

Gene Title	Gene Symbol	Ensembl	Experiment 1: log2 (ratio)	Experiment 2: log2 (ratio)	Human orthologue	Function
defensin	Def	CG1385	-1.25 P=0.002	-1.04 P=0.041		immune response
Plum	bw	CG17632	-1.13 P=0.027	-0.51 P=0.018		eye pigment precursor transport activity
CG4725	CG4725	CG4725	0.53 P=0.006	0.83 P=0.027		Metallo-endopeptidase
CG3397	CG3397	CG3397	0.81 P=0.042	0.81 P=0.022		oxidative stress response
CG33115	CG33115	CG33115	0.84 P=0.027	0.78 P=0.020		Nimrod B4
CG6687	CG6687	CG6687	0.91 P=0.035	0.89 P=4.09E-5		hydrogen ion transmembrane transporter
CG10031	CG10031	CG10031	0.91 P=0.044	0.50 P=0.015		peptidase activity
CG18067	CG18067	CG18067	0.95 P=0.014	0.76 P=0.006		
CG15293	CG15293	CG15293	1.26 P=0.018	0.83 P=0.007		

There are only two genes commonly altered in rCAG repeat expressing flies compared to *elav>+* between the two experiments (listed in Table 5.3) and three genes for rCUG repeat flies (listed in Table 5.4). Of the two genes altered in *elav>rCAG* flies compared to *elav>+* in both microarray experiments, one gene, *mod(mdg4)*, was previously identified as a modifier in a P-element screen of a SCA8 *Drosophila* model where the human SCA8 non-coding RNA was expressed in the *Drosophila* eye (98). Similarly, of the three genes altered in *elav>rCUG* flies compared to *elav>+* in both microarray experiments, two genes have been previously implicated in expanded repeat pathogenesis. The first, *Hu li tai shao (hts)*, is an orthologue of mammalian Adducin 1 (ADD1) (221), a protein which has been identified as an HTT interactor by yeast-two-hybrid analysis (121). In *Drosophila,* Hts has been demonstrated to play a role in oogenesis and embryogenesis (221-222) however, following identification of ADD1 as an HTT interactor, mutations in *hts* were shown to modify a model of polyglutamine pathogenesis with 128Q in the

context of exon1 of the human *HTT* gene (121). *Muscleblind* (*mbl*), the *Drosophila* orthologue of the human MBNL splicing factor, was also identified as commonly altered in *elav>rCUG* flies compared to *elav>+*. MBNL has been demonstrated to play a key role in the pathogenesis of DM1 & 2, where it is thought to be sequestered by the presence of CUG or CCUG RNA repeats and therefore cannot perform its normal splicing functions (72, 76, 220).



Experiment 1

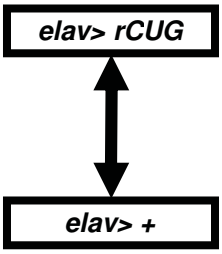
+

Experiment 2

Table 5.3: Common changes for elav>rCAG compared to elav>+ in experiment 1 and 2.

Log2(ratio) >0.5 or <-0.5, P<0.05.

Gene Title	Gene Symbol	Ensembl	Experiment 1: log2(ratio)	Experiment 2: log2(ratio)	Human Orthologue	Function
Modifier67.2	mod(mdg4)	CG32491	0.68 P=0.008	-0.76 P=0.012		Regulation of chromatin assembly
Immune induced molecule 2	IM2	CG18106	1.70 P=0.034	0.75 P=0.017		Immune response



Experiment 1

+

Experiment 2

Table 5.4: Common changes for elav>rCUG compared to elav>+ in experiment 1 and 2.

Log2(ratio) >0.5 or <-0.5, P<0.05.

Gene Title	Gene Symbol	Ensembl	Experiment 1: log2(ratio)	Experiment 2: log2(ratio)	Human Orthologue	Function
CG31846	CG31846	CG31846	0.64 P=0.041	-0.77 P=0.027		
hu-li tai shao	hts	CG9325	0.63 P=0.006	-0.62 P=0.032	ADD1	Actin assembly, ring canal formation
mindmelt	mbl	CG33197	0.66 P=0.016	-0.60 P=0.022	MBNL1	Splicing factor, muscle and nervous system development

While there are a small number of common changes induced by expression of both rCAG and rCUG repeats in the two experiments, in a number of cases there is a lack of concordance in the direction of the change (Table 5.3 and 5.4). Since in many cases there is more than one transcript detected by a probe-set and, in some cases, more than one probe-set for a single gene present on the Affymetrix array, it is possible that this represents a difference in the particular splice-forms of the gene being detected. It is also possible that there is some threshold effect on these pathways, whereby different effects are elicited through the same pathway depending on the amount of stress the cell is experiencing. In this study, alterations were used as an indication that the particular pathway represented was perturbed by repeat expression, irrespective of the direction of the change.

5.7 Analysis of genes significantly altered in each microarray experiment

The identification of altered expression of a number of genes previously implicated in expanded repeat pathogenesis in rCAG and rCUG repeat expressing flies when compared to *elav>+* but not *elav>rCAA* flies may indicate that at least some of the detected changes are not specific outcomes of hairpin RNA expression. These changes may be due to altered gene expression in *elav>+* flies resulting from toxicity associated with neuronal expression of GAL4. It is predicted that a decreased amount of free GAL4 protein should be present in flies expressing the rCAG, rCUG or rCAA repeat constructs due to the presence of an equivalent number of UAS sites in each of these genotypes and, therefore, the impact of GAL4 toxicity should be reduced in these flies. It is also possible that expression of rCAA RNA itself may result in a unique alteration to the transcriptional profile of *Drosophila*. Since the *elav-GAL4* driver line used to express each of the repeat constructs was different between the two microarray experiments, any specific effects elicited by rCAA RNA expression may result in differences in the transcriptional changes observed in each experiment. For this reason, it was also necessary to investigate transcriptional profiles for each of the microarray experiments separately.

5.7.1 Genes changed in rCAG repeat-expressing flies compared to both *elav>rCAA and elav>+*

We predict that changes detected in rCAG or rCUG repeat-expressing flies compared to both the *elav>rCAA* and *elav>+* control lines will be most likely to represent hairpin RNA-dependent pathogenic pathways and not just the result of cellular stress, since this comparison takes into account both the documented toxicity of the GAL4 protein (137) and any possible affect of expressing CAA repeat RNA in cells. Investigation of transcriptional changes elicited by neuronal expression of rCAG and rCUG repeats should also identify both sequence-independent components of hairpin RNA toxicity – that is changes which are seen in both rCAG and rCUG repeat expressing flies – and unique sequence-dependent effects of expression of each of these repeats.

In experiment 1, eight genes showed a significant change in expression when *elav>rCAG* flies are compared to *elav>*rCAA and *elav>+* (listed in Table 5.5). Of these, five have some function ascribed to them: glycogenin is an important metabolic enzyme which plays a role in the synthesis of glycogen and mesoderm development, muscleblind (Mbl) is the *Drosophila* orthologue of the human MBNL splicing factor, which has been implicated in the pathogenesis of DM1 and DM2 (72, 81), Pk is involved in determination of cell polarity (223), Zip3 is a zinc transporter which may play roles in immune response (224) and Lea is an axon guidance receptor which regulates processes in neurogenesis (225). In experiment 2, eight genes showed altered expression in rCAG repeat expressing flies compared to both the *elav>rCAA* and *elav>+* control lines (Table 5.6). These genes play roles in lipid metabolism, myogenesis and neuronal signalling in *Drosophila*, however there is little information available regarding their specific function or the function of their mammalian orthologues. While there is no specific overlap in genes altered as a result of rCAG repeat expression between the two experiments, there appears to be a consistent disruption of processes such as neurogenesis and muscle development and therefore these processes are likely to be important components of CAG repeat-mediated RNA toxicity in this model.

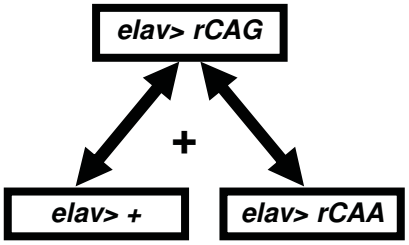


Table 5.5: Genes of particular interest for elav>rCAG identified in microarray experiment 1. Genes are significantly altered in *elav>rCAG* flies compared to both *elav >rCAA* and *elav >+*. Log2(ratio) >0.5 or <-0.5, P<0.05.

Gene Title	Gene Symbol	Ensembl	Log2(ratio) compared to elav>+	Log2(ratio) compared to elav>rCAA	Human orthologue	Function
molting defective	mld	CG34100	-0.62 P=0.040	-0.82 P=0.019		
CG15642	CG15642	CG15642	-0.91 P=0.030	-0.74 P=0.023		
Glycogenin	glycogenin	CG9480	-0.78 P=0.015	-0.69 P=0.023	GYG1	glycogenin
mindmelt	mbl	CG33197	0.88 P=0.022	-0.62 P=0.037	MBNL1	RNA binding
prickle-spiny legs	pk	CG11084	-0.62 P=0.049	-0.62 P=0.041	PRICKLE2	neurite outgrowth
Zinc/iron regulated transporter-related protein 3	zip3	CG6898	-0.81 P=0.019	-0.61 P=0.013	SLC39A2	zinc transporter
CG8925	CG8925	CG8925	-0.56 P=0.048	-0.58 P=0.015		
Robo2	lea	CG5481	-0.53 P=0.049	-0.56 P=0.037	ROBO1	axon guidance receptor

Table 5.6: Genes of particular interest for elav>rCAG identified in microarray experiment 2. Genes are significantly altered in *elav>rCAG* flies compared to both *elav >rCAA* and *elav>+*. Log2(ratio)>0.5 or <-0.5, P<0.05.

Gene Title	Gene Symbol	Ensembl	Log2(ratio) compared to elav>+	Log2(ratio) compared to elav>rCAA	Human Orthologue	Function
CG3823	CG3823	CG3823	-0.8 P=0.005	-0.95 P=0.006		
CG33528	CG33528	CG33528	0.68 P=0.025	0.60 P=0.042	SLC18A2	Neurotransmitter secretion, synaptic vesicle amine transport
CG7714	CG7714	CG7714	0.55 P=0.016	0.68 P=0.034		
CG14528	CG14528	CG14528	0.85 P=0.026	0.72 P=0.048		
CG13062	CG13062	CG13062	1.26 P=0.029	0.72 P=0.025		
yolk protein	yp3	CG11129	1.43 P=0.004	0.79 P=0.022		Fat body protein, putative lipase
CG18641	CG18641	CG18641	0.85 P=0.011	0.81 P=0.007		lipase activity
echinoid	ed	CG12676	0.63 P=0.002	1.05 P=0.034		actin cytoskeleton/cell adhesion, may form a signalling complex with Grip, important in myogenesis

5.7.2 Genes changed in rCUG repeat-expressing flies compared to *elav>rCAA and elav>+*

In experiment 1, 26 genes showed altered expression in *elav>rCUG* flies compared to both *elav>rCAA* and *elav>+* (listed in Table 5.7) and of these, 20 have some functional information attributed to them. Amongst these 20 genes, there are a number involved in transcriptional regulation and RNA metabolism which are downregulated – *sv, dve, lola, msl-2, CG8273, MED24, xl6* - while upregulation of several metabolic genes is also observed. In contrast, there were no common changes detected in *elav>rCUG* flies compared to *elav>rCAA* and *elav>+* in experiment 2, despite a large amount of overlap in categories of detected genes between the two experiments (Figure 5.3 F & H).

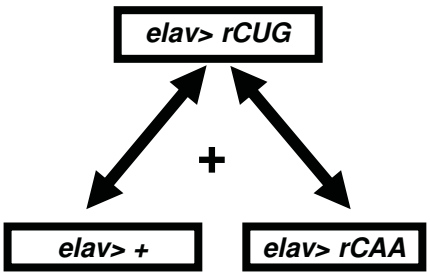


Table 5.7: Genes of particular interest for elav>rCUG identified in microarray experiment 1. Genes are significantly altered in *elav>rCUG* flies compared to both *elav >rCAA* and *elav>+*. Log2(ratio)>0.5 or <-0.5, P<0.05.

Log2(ratio) >0.5 or <-0.5, P<0.05.

Gene Title	Gene Symbol	Ensembl	log2 (elav>rCUG to elav>+)	log2 (elav>rCUG to elav>CAA)	Human orthologue	Function
sparkling	sv	CG11049	-0.84 P=0.038	-0.87 P=0.038	PAX5	B cell specific transcription factor/ midbrain dopaminergic neuron specification
defective proventriculus	dve	CG5799	-0.73 P=0.028	-0.81 P=0.012		Transcription factor activity
longitudinals absent	lola	CG12052	-0.53 P=0.028	-0.74 P=0.019	ZBTB3	Transcription factor activity
CG5514	CG5514	CG5514	-0.57 P=0.003	-0.74 P=0.004	FAM44A	Structural component of cell wall
CG12641	CG12641	CG12641	-0.66 P=0.008	-0.73 P=3.99E-05		
U26	U26	CG13401	-0.54 P=0.042	-0.71 P=0.019	AASDH	2-aminoadipic 6-semialdehyde dehydrogenase, fatty acid metabolism
metabotropic GABA-B receptor subtype 3	GABA-B-R3	CG3022	-0.51 P=0.045	-0.64 P=0.007		neurotransmission
male specifc lethal	msl-2	CG3241	-0.55 P=0.002	-0.62 P=0.011	MSL2L1	H4 histone acetylation
Suppressor of cytokine signaling at 36E	Socs36E	CG15154	-0.52 P=0.024	-0.61 P=0.033	SOCS1	suppressor of cytokine signalling

Gene Title	Gene Symbol	Ensembl	log2 (elav>rCUG to elav>+)	log2 (elav>rCUG to elav>CAA)	Human orthologue	Function
deep orange	dor	CG3093	-0.52 P=0.038	-0.59 P=0.011	VPS18	vacuolar protein sorting
CG33331	CG33331	CG33331	-0.57 P=0.036	-0.59 P=0.021	C3orf31	Mitochondrial import protein
CG8273	CG8273	CG8273	-0.57 P=0.019	-0.59 P=0.048	SON	Double stranded RNA binding, DNA binding
CG13594	CG13594	CG13594	-0.52 P=0.042	-0.57 P=0.033		
Mediator complex subunit 24	MED24	CG7999	-0.56 P=0.039	-0.54 P=0.020	MED24	component of mediator complex, transcriptional co-activator
xl6	xl6	CG10203	-0.52 P=0.022	-0.53 P=0.011	SFRS7	splicing factor, arginine/serine rich
CG6364	CG6364	CG6364	0.58 P=0.045	0.56 P=0.025	UCK2	pyrimidine ribonucleoside kinase, production of UMP and CMP
Cyp12c1	Cyp12c1	CG4120	0.51 P=2.66E-04	0.66 P=0.003	CYP24A1	Mitochondrial enzyme that inactivates metabolites of vitamin D
CG3534	CG3534	CG3534	0.58 P=0.027	0.68 P=0.014	XYLB	Energy metabolism
CG13845	CG34376	CG13845	0.50 P=0.018	0.69 P=0.019		
CG3246	CG3246	CG3246	0.51 P=0.040	0.70 P=0.024		
CG5793	CG5793	CG5793	0.69 P=0.019	0.73 P=0.004	FAHD1	fumarylacetoacetate hydrolase, mitochondrial
CG4716	CG4716	CG4716	0.55 P=0.009	0.83 P=4.74E-04		
CG14629	CG14629	CG14629	0.65 P=0.024	0.87 P=4.50E-04		
CG31370	CG31370	CG31370	0.78 P=0.008	0.98 P=2.51E-05		
hemolectin	hml	CG7002	1.05 P=0.047	1.02 P=0.017		Immune response
PGRP-SB1	PGRP-SB1	CG9681	1.16 P=0.023	1.24 P=0.008		Immune response

5.7.3 Genes changed in both rCAG and rCUG repeat-expressing flies compared to *elav>rCAA*

Since the hairpin structures predicted for rCAG and rCUG RNAs are structurally nearly identical and a number of clinical features of the different expanded repeat diseases overlap, we predict that there are likely to be common effects of expressing these RNAs in the neurons of *Drosophila*. Lists were therefore also compiled of genes which changed in flies expressing both rCAG and rCUG repeats compared to each of the controls. There were no genes altered in both rCAG and rCUG repeat-expressing flies compared to both controls.

Of the 38 genes shown in Table 5.8 which are changed in both *elav>rCAG* and *elav>rCUG* flies compared to *elav>rCAA*, 23 have some functional information. One gene altered in this list is CG5669, an orthologue of the human SP1 transcription factor which has been implicated in transcriptional regulation in HD (226-227). While another gene, *CG1343*, is more closely related at the sequence level to mammalian SP1, regulatory roles for each of the SP1 orthologues in *Drosophila* have not been extensively studied. The mammalian SP1/SP3 family of transcription factors have been implicated in regulation of a broad range of processes including the mTOR signalling pathway (228), mitochondrial biogenesis (229-230), glucose signalling pathways (231), mitosis (232) and glutamate signalling (233-235). Interestingly, Table 5.8 includes changes to genes involved in mTOR signalling – *CG30044* – and mitochondrial biogenesis – *sun, ttm50, CG4306* – as well as the *Drosophila* orthologue of the wolfram syndrome 1 gene, *wfs1*, which is regulated by SP1 in mammals (236), raising the possibility that dysregulation of similar pathways is occurring in flies expressing rCAG and rCUG repeats.

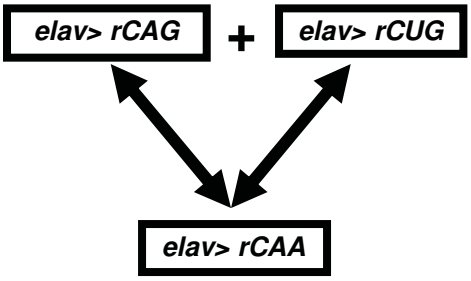


Table 5.8: Common transcriptional changes in elav>rCAG and elav>rCUG compared to elav>rCAA in experiment 1.

Log2(ratio) >0.5 or <-0.5, P<0.05.

Gene Title	Gene Symbol	Ensembl	Log2 (elav>rCAG to elav>rCAA)	Log2 (elav>rCUG to elav>rCAA)	Human orthologue	Function
defensin	def	CG1385	-1.37 P=0.001	-1.25 P=0.002		Immune response
stunted	sun	CG9032	-1.50 P=0.009	-1.13 P=0.033	ATP5E	mitochondrial ATP synthase epsilon chain
CG8297	CG8297	CG8297	-0.97 P=0.021	-1.04 P=0.025	TXNDC15	Thioredoxin
beat-IIIc	beat-IIIc	CG15138	-0.95 P=0.044	-1.01 P=0.035		
CG18437	CG18437	CG18437	-0.70 P=0.039	-0.98 P=0.020	C2orf21	
Cad99C	Cad99C	CG31009	-0.74 P=0.021	-0.89 P=0.012	PCDH15	protocadherin 15 precursor, cell adhesion
CG9264	CG9264	CG9264	-0.61 P=0.038	-0.87 P=0.008		
CG32056	CG32056	CG32056	-0.65 P=0.041	-0.84 P=0.022	PLSCR1	Phospholipid scramblase
CG6695	CG6695	CG6695	-0.59 P=0.039	-0.80 P=0.023	SFRS16	Splicing factor, arginine/serine-rich 16: hSWAP
CG2010	CG2010	CG2010	-0.67 P=0.015	-0.80 P=0.012		
His3:CG31613	His3:CG31613	CG31613	-0.81 P=0.004	-0.74 P=0.014	HIST2H3	H3 histone
CG12641	CG12641	CG12641	-0.63 P=0.031	-0.73 P=3.99E-05		
His1:CG31617	His1:CG31617	CG31617	-0.83 P=0.006	-0.71 P=0.045	HIST1H1	H1 histone
CG30044	CG30044	CG30044	-0.75 P=0.029	-0.71 P=0.032	mLST8	Target of Rapamycin complex subunit
CG8505	Cpr49Ae	CG8505	-0.63 P=0.043	-0.70 P=0.044		
Wolfram syndrome 1	wfs1	CG4917	-0.56 P=0.008	-0.69 P=0.011	WFS1	modulates free calcium in ER, regulated by Sp1
CG2713	ttm50	CG2713	-0.53 P=0.022	-0.67 P=0.029	TIMM50	Translocase of inner mitochondrial membrane
CG5669	CG5669	CG5669	-0.54 P=0.011	-0.62 P=0.007	SP1/SP3	Transcription factor
CG8833	CG8833	CG8833	-0.63 P=0.011	-0.62 P=0.027	GPATC1	RNA processing
CG16857	CG16857	CG16857	-0.57 P=0.047	-0.62 P=0.024		

Gene Title	Gene Symbol	Ensembl	Log2 (elav>rCAG to elav>rCAA)	Log2 (elav>rCUG to elav>rCAA)	Human orthologue	Function
Suppressor of cytokine signaling at 36E	Socs36E	CG15154	-0.52 P=0.047	-0.61 P=0.033	SOCS1	suppressor of cytokine signalling
CG10321	CG10321	CG10321	-0.60 P=0.002	-0.60 P=0.027		
tonalli	tna	CG7958	-0.55 P=0.043	-0.60 P=0.032	ZMIZ1	transcriptional coactivator
Bem46	Bem46	CG18642	-0.51 P=0.041	-0.60 P=0.016	ABHD13	hydrolase activity
CG31638	CG31638	CG31638	-0.70 P=0.002	-0.58 P=0.021	CCDC102A	Tropomyosin
Robo2	lea	CG5481	-0.56 P=0.037	-0.58 P=0.037	ROBO1	axon guidance receptor
CG9213	CG9213	CG9213	-0.51 P=0.044	-0.56 P=0.007	CWF19L2	cell cycle control
CG10362	CG10362	CG10362	-0.54 P=0.034	-0.54 P=0.035	PDZK8	signalling
Odorant-binding protein 19a	Obp19a	CG11748	-0.60 P=0.038	-0.54 P=0.042		
CG12116	CG12116	CG12116	0.73 P=0.002	0.56 P=0.015		
CG11072	mamo	CG11072	0.84 P=0.037	0.68 P=0.024		
CG31781	CG31781	CG31781	0.61 P=0.037	0.70 P=9.21E-05	LIP	lipase
CG15068	CG15068	CG15068	0.97 P=0.012	0.91 P=0.016		
CG10026	CG10026	CG10026	0.52 P=0.011	0.94 P=0.026	TTPA	Vitamin E metabolism, deficiency leads to cerebellar degeneration
CG4306	CG4306	CG4306	0.82 P=0.036	1.23 P=0.020	GGCT	glutathione homeostasis, release of cytochrome c from mitochondria
CG15293	CG15293	CG15293	0.71 P=0.032	1.26 P=0.018		
CG4716	CG4716	CG4716	0.90 P=0.005	1.33 P=0.007		
CG13086	CG13086	CG13086	0.89 P=0.039	1.40 P=0.024		

In experiment 2, only 8 genes were changed in both *elav>rCAG* and *elav>rCUG* flies compared to *elav>rCAA* (Table 5.9) and there is no overlap with the list of genes altered in experiment 1 (Table 5.8). Interesting candidates on this list include *ctp* (*ddlc1*) and *insc*, which are involved in microtubule-based movement of proteins and RNA in the cell. *Ctp* is the *Drosophila* orthologue of mammalian dynein light chain-like 2 (DYNLL2), a component of the dynein complex which has been implicated in processes including axonal transport of mitochondria (237), retrograde transport (238) and stress granule formation (239). In *Drosophila,* Ctp has been shown to be involved in dendritic branching and endosome movement (240), axon path-finding (241) and toxic protein clearance through autophagy and cell death, with mutants showing decreased motor activity (242), gross morphological defects and apoptosis (243). *Ctp* has also been previously identified as downregulated in a microarray study investigating polyglutamine-specific transcriptional changes common to *Drosophila* and human cell lines (244).

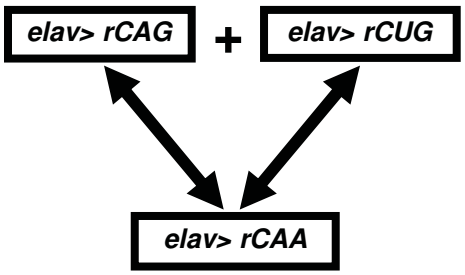


Table 5.9: Common transcriptional changes in elav>rCAG and elav>rCUG flies compared to elav>rCAA in experiment 2.

Log2(ratio) >0.5 or <-0.5, P<0.05.

Gene Title	Gene Symbol	Ensembl	log2 (elav>rCUG to elav>rCAA)	log2 (elav>rCAG to elav>rCAA)	Human Orthologue	Function
Inscuteable	insc	CG11312	0.55 P=0.050	0.89 P=0.013	INSC	cytoskeletal adaptor, protein and RNA localisation, localisation is dynein dependent
dynein	ctp	CG6998	0.55 P=0.002	0.70 P=0.001	DYNLL2	microtubule-based movement, required for proper axon guidance of sensory neurons.
CG18641	CG18641	CG18641	0.67 P=0.031	0.81 P=0.003		lipase
CG12998	CG12998	CG12998	0.71 P=0.023	0.92 P=0.018		
CG34104	CG34104	CG34104	0.80 P=0.004	0.58 P=1.43E-05		microtubule based G-protein coupled signal transduction
CG14528	CG14528	CG14528	0.84 P=0.015	0.72 P=0.024		Metallo-endopeptidase
CG9400	CG9400	CG9400	0.97 P=0.014	0.84 P=0.021		peptidase inhibitor
CG9079	Cpr47Ea	CG9079	1.34 P=0.003	0.51 P=0.011		structural component of cuticle

Expression of mutant HTT in *Drosophila* has been shown to cause visible axonal blockages in motor neuron axons, with aggregation of vesicles containing the mutant protein observed (13). Similarly, mutant androgen receptor has been demonstrated to form aggregates which disrupt transport and cause axonal swelling in immortalized motor neuron cells (245). However in a *Drosophila* model, expression of either polyglutamine alone or in the context of the SCA3 gene was not able to disrupt axonal traffic in the same manner (119). These observations may suggest that transport defects are a context-specific effect related to the function of the polyglutamine-containing protein and not a more general property of polyglutamine expression. A broader role for both dynein and kinesin motor complexes (239) and the microtubule network (246) has also been described in the formation and dynamics of stress granule and p-body formation; a role which is conserved in *Drosophila* (247). Both stress granules and p-bodies are involved in regulation of translation, editing, splicing, degradation and transport of RNAs in the cell and are highly populated by RNA binding proteins. It is therefore possible that dynein transport defects can be elicited through either toxic RNA-dependent or toxic protein-dependent pathways resulting in some overlapping consequences and that, in the polyglutamine diseases, both of these mechanisms may be in action, making cells particularly vulnerable.

5.7.4 Genes changed in both rCAG and rCUG repeat-expressing flies compared to *elav>+*

Analysis of genes changed in both *elav>rCAG* and elav>*rCUG* compared to *elav>+* was also performed for both experiment 1 and 2. Only 9 genes were altered in both *elav>rCAG* and *elav>rCUG* compared to *elav>+* in experiment 1 (Table 5.10) and of these only 4 have a previously characterised function. All 4 of these genes have been previously implicated in expanded repeat diseases and 3 were also altered for either, but not both, rCAG or rCUG in experiment 2 (Table 5.4 & 5.5). As previously described, mutations in *hts* have been shown to modify a model of polyglutamine pathogenesis with 128Q in the context of exon1 of the human *HTT* gene (121), the human orthologue of Mbl, MBNL, has been characterised for a role in DM1 & 2 where it is thought to be sequestered by the presence of CUG or CCUG RNA repeats and therefore cannot perform its normal splicing functions (72, 76, 220)

and *mod(mdg4)* has been previously identified as a modifier in a P-element screen of a SCA8 *Drosophila* model where the human SCA8 non-coding RNA was expressed in the *Drosophila* eye (98). The final gene, metabotropic glutamate receptor A (*mGluRA)*, is the sole metabotropic glutamate receptor in *Drosophila,* which has been shown to have a mutual negative feedback relationship with *dFMRP*, the *Drosophila* orthologue of the Fragile X mental retardation protein (248). This antagonistic relationship is thought to regulate levels of ionotropic glutamate receptors and therefore either dampen synaptic excitability if mGluRA dominates, or sharpen it if FMRP dominates (249).

Analysis of genes altered in *elav>rCAG* and elav>rCUG flies compared to *elav>+* in experiment 2 (Table 5.11) showed no overlap with the same comparison for experiment 1 (Table 5.10). Of the 9 genes identified 5 have functional information associated with them, although none of them have been formerly linked to expanded repeat disease pathogenesis. Amongst these is *hr38*, which encodes a *Drosophila* orthologue of Nerve growth factor I-B (NGFI-B) – also known as Nuclear receptor 77 (NUR77) – which is a nuclear receptor known to play a role in dendritic differentiation and synapse formation (250). In humans, it is highly expressed in the striatum and prefrontal cortex (251), regions which are most affected in HD. NUR77 has also been demonstrated to be a regulator of glucose and lipid metabolism in rat skeletal muscle (252) and therefore the upregulation of Hr38 expression observed in *Drosophila* expressing CAG and CUG repeat RNAs may be indicative of metabolic dysfunction in these flies.

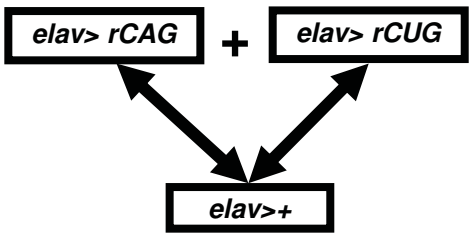


Table 5.10: Common transcriptional changes in elav>rCAG and elav>rCUG compared to elav>+ in experiment 1. Log2(ratio) >0.5 or <-0.5, P<0.05.

Gene Title	Gene Symbol	Ensembl	log2 (elav>rCAG to elav>+)	log2 (elav>rCUG to elav>+)	Human orthologue	Function
CG18213	CG18213	CG18213	-0.67 P=0.034	-0.96 P=0.008		
CG15369	CG15369	CG15369	-0.67 P=0.047	-0.80 P=0.019		
hu li tai shao	hts	CG9325	0.60 P=0.002	0.63 P=0.006	ADD1	Cytoskeletal protein, substrate for protein kinase A & C
CG31846	CG31846	CG31846	0.60 P=0.005	0.64 P=0.041		
metabotropic glutamate receptor	mGluRA	CG11144	0.65 P=0.018	0.50 P=0.045	GRM3	Glutamate receptor
CG15216	CG15216	CG15216	0.66 P=0.003	0.55 P=0.026		
Modifier67.2	mod(mdg4)	CG32491	0.68 P=0.008	0.51 P=0.028		Transcriptional regulation
CG13618	CG13618	CG13618	0.84 P=0.004	0.65 P=0.009		
mindmelt	mbl	CG33197	0.88 P=0.022	0.66 P=0.016	MBNL1	Splicing factor

Table 5.11: Common changes for *elav>rCAG* and *elav>rCUG* compared to *elav>+* in experiment 2. Log2(ratio) >0.5 or <-0.5, P<0.05.

Gene Title	Gene Symbol	Ensembl	log2 (elav>rCUG to elav>+)	log2 (elav>rCAG to elav>+)	Human Orthologue	Function
Stellate orphon	Ste12DOR	CG32616	-2.00 P=0.003	-2.88 P=0.002		Spermatogenesis, protein kinase regulator
CG4688	CG4688	CG4688	-0.99 P=0.017	-0.72 P=0.039	MARS	Glutathione transferase
CG32552	CG32552	CG32552	-0.86 P=0.009	-0.55 P=0.013		
CG13077	CG13077	CG13077	-0.79 P=0.027	-0.55 P=0.019	CYB561D2	Electron transport chain
CG13117	CG13117	CG13117	-0.56 P=0.009	-0.53 P=0.015		
CG6752	CG6752	CG6752	-0.54 P=0.027	-0.58 P=0.003	RNF123	Ubiquitin ligase
CG9686	CG9686	CG9686	0.54 P=0.016	0.55 P=0.014		
Hormone receptor-like in 38	hr38	CG1864	0.55 P=0.028	1.05 P=0.040	NUR77/ NGFI-B	Ligand-dependent nuclear receptor activity, transcription activity
CG9186	CG9186	CG9186	0.77 P=0.031	0.58 P=0.017		

5.8 Summary of results from microarray analysis

Microarray analysis was performed on newly eclosed flies expressing rCAG and rCUG repeats specifically in the nervous system in order to investigate transcriptional changes associated with hairpin RNA expression, which may represent some of the earliest pathogenic changes in the expanded repeat diseases. A large degree of variation was observed both between independent microarray experiments – perhaps as a result of the difference in the driver and repeat lines used between the two experiments – and within each experiment, depending on whether *elav>rCAA* or *elav>+* was used as a control. However, components of pathogenic pathways that have been described in other expanded repeat models were identified when either *elav>rCAA* or *elav>+* were used as a control and it is therefore unclear whether one control should be deemed preferable to the other.

Analysis of changes occurring in our model identified a number of pathways which were dysregulated in response to expression of rCAG and rCUG repeats, including transcriptional regulation, redox regulation, axonal transport, RNA processing and lipid metabolism. A number of the genes altered in this microarray study have been previously identified in other *Drosophila* expanded repeat disease models, including *ctp* (244), *mbl* (117)*, hts* (121), *mod(mdg4)* (98) and *mGluRA* (248-249)*,* suggesting that the expression of hairpin RNA alone is sufficient to induce some of the transcriptional dysregulation observed, irrespective of the context of the repeat. It is also indicative that at least a proportion of the cellular changes in these models are not sequence-dependent, since expression of either rCAG or rCUG repeats were both able to induce a number of these changes. Since the presence of repeat-containing RNA is common to both the translated and untranslated expanded repeat diseases, it seems likely that the changes effected by hairpin RNA expression may also play a role in pathogenesis of the polyglutamine diseases.

Chapter 6: Genetic verification of candidates from microarray analysis

A number of transcriptional changes in *Drosophila* expressing expanded untranslated CAG and CUG repeats in neurons were identified using a microarray approach, as described in Chapter 5. From this initial analysis, it is not apparent whether these genes represent early pathogenic changes resulting from hairpin repeat expression, or are more general downstream effects of cellular stress. This is a particularly pertinent distinction since the majority of changes identified in this and other expanded repeat disease microarray studies coincide with changes that have been described in other neurological diseases. For example, axonal transport defects have been described in models of HD (13-14, 119) but also in prion disease (253), ALS (254), Parkinson's disease (255) and Alzheimer's disease (256-257). These sorts of changes are likely to represent common neuronal responses to stress which may play a role in expanded repeat disease, but cannot account for differences in pathology observed between these and other neurological diseases.

In order to determine which of the changes identified in the microarray study can be attributed to expression of hairpin RNA and which are more general responses to cellular stress, we performed a genetic screen of candidates from both microarray experiment 1 and 2. The candidate genes and alleles tested along with a summary of the changes observed in the microarray experiments are shown in Figure 6.1. Since expression of up to four transgene insertions of the rCAG, rCUG or rCAA repeat constructs in the eye with *GMR-GAL4* does not elicit a phenotype, the phenotypes observed when polyglutamine encoded by either a CAG or CAA repeat or polyleucine encoded by a CUG repeat are expressed in the eye (described in section 2.1) were utilised for a primary screen. In this screen, *Drosophila* expressing either polyglutamine or polyleucine specifically in the eye were crossed to candidate gene stocks and scored for modification of the eye phenotype. In cases where modification of the eye phenotype is observed in *Drosophila* expressing CAG, CUG and CAA repeats, this indicates that the alteration to the phenotype is unlikely to be due to a specific interaction with the hairpin repeat RNA – since the CAA RNA is unable to form this structure – and these transcriptional changes are therefore more likely to represent general outcomes of cellular stress. A similar methodology has been successfully used to identify *mbl* as a candidate gene which is able to modify the phenotype caused by a polyglutamine tract encoded by a pure CAG repeat but not a mixed CAG/CAA repeat (117).

We obtained *Drosophila* stocks from the Vienna *Drosophila* RNAi Centre (VDRC) and Bloomington stock centre to reduce expression of candidate genes by either P-element insertion or RNAi (listed in Table 6.1). Flies were generated with a recombinant chromosome containing *GMR-GAL4* and either a CAG, CUG or CAA transgene insertion. These flies were crossed to each of the candidate gene stocks to produce flies with expression of either polyglutamine or polyleucine and reduced expression of the candidate gene (depicted in Figure 6.1). Where possible, independent candidate gene stocks were obtained to verify interactions and rule out the effects of background mutations.

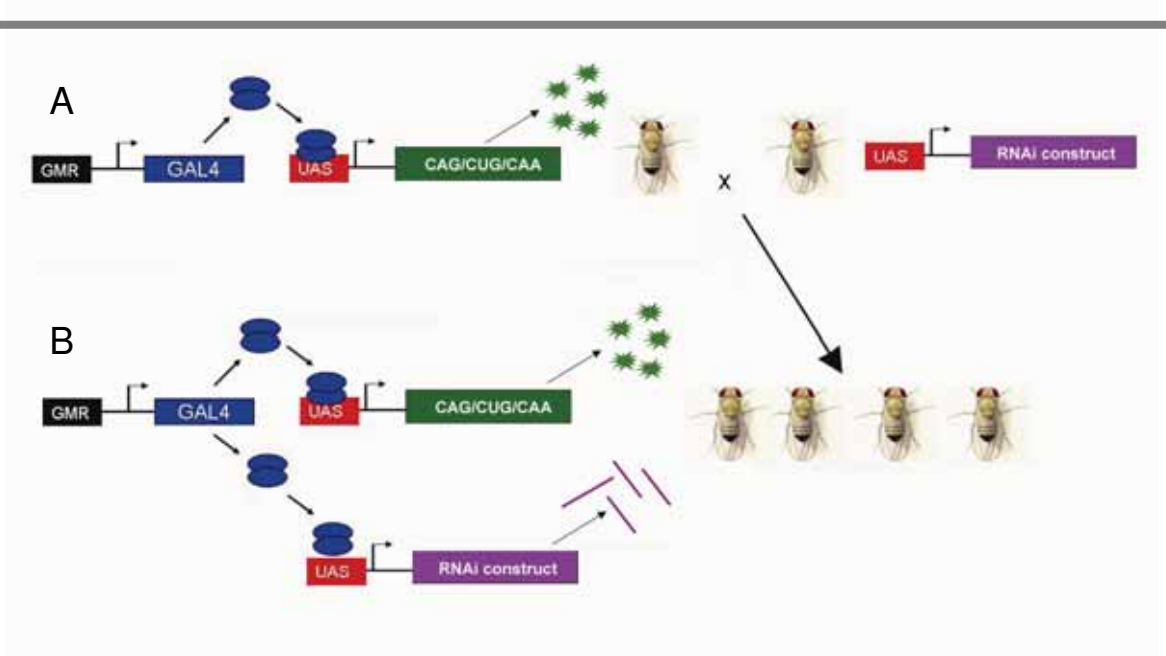


Figure 6.1: Method to generate *Drosophila* expressing polyglutamine or polyleucine in the eye along with an RNAi construct targeting a candidate gene. A) Flies were generated with a chromosome recombinant for *GMR-GAL4* and a CAG, CAA or CUG repeat tract encoding polyglutamine or polyleucine. Each of these stocks was then crossed to the candidate gene stocks (listed in Table 6.1) to produce progeny which express both polyglutamine or polyleucine and the RNAi construct specifically in the eye **(B)**. Where P-element lines were obtained for a candidate gene, the same procedure was followed except in this case the progeny have reduced expression of the candidate gene in all tissues. A control was included in each set of crosses, consisting of the recombinant line carrying *GMR-GAL4* and the repeat tract out-crossed to a wild-type line (w^{1118}). The progeny of this cross were used to indicate the severity of the eye phenotype in flies without knock-down of the candidate gene.

		rCAG				rCUG			
Gene	**Alleles tested**	**Experiment 1**		**Experiment 2**		**Experiment 1**		**Experiment 2**	
		elav> rCAA	*elav>+*	*elav> rCAA*	*elav>+*	*elav> rCAA*	*elav>+*	*elav> rCAA*	*elav>+*
hts (CG9325)	01103, **KG06777**	-	**0.6** P=0.002	-	-	-	**0.63** P=0.006	-	**-0.62** P=0.032
mGluRA (CG11144)	v1793, v**1794**	-	**0.65** P=0.018	-	-	-	**0.5** P=0.045	-	-
CG5669	**v45300**	**-0.54** P=0.011	-	-	-	**-0.62** P=0.007	-	-	-
insc (CG11312)	**v31488**	-	-	**0.89** P=0.013	-	-	-	**0.55** P=0.05	-
ctp (CG6998)	v43116, v**43115**	-	-	**0.7** P=0.001	-	-	-	**0.55** P=0.002	-
mod(mdg4) (CG32491)	**v52268**	-	**0.68** P=0.008	-	**-0.76** P=0.012	-	**0.51** P=0.028	-	-
mef2 (CG1429)	**v15550**	-	-	-	-	-	-	**-0.9** P=0.007	-
mbl (CG33197)	**v28731**	**-0.62** P=0.037	**0.88** P=0.022	-	-	-	**0.66** P=0.016	-	**-0.6** P=0.022

Table 6.1: Overview of genes and alleles tested for genetic interaction with expanded repeats. Candidate genes were selected based on the microarray experiments described in Chapter 5 (Log2(ratio)>0.5 or <-0.5 and P<0.05). Alleles preceded by "v" were obtained from the VDRC stock collection. Alleles of *hts* were sourced from the Bloomington stock centre. All VDRC lines were tested for presence of an eye phenotype when expressed with *GMR-GAL4* and, in all cases, no disruption to the exterior appearance of the eye was seen. Alleles in bold are shown in figures.

While a modification of the polyglutamine or polyleucine phenotypes would be suggestive of a role for the candidate gene in expanded repeat disease pathogenesis, the lack of a change does not necessarily rule out a candidate. In this screen, we are utilising the disruption of the ordered structure of the *Drosophila* eye as a marker of cellular dysfunction and death. Since not all of the cells in the eye are of neural origin, it is possible that they do not respond in the same manner to expanded repeat expression as the cells of the nervous system which were investigated by microarray analysis. Furthermore, while the candidate gene lines have been tested by quantitative real-time PCR for their ability to knock-down expression of their target RNA, there may be cases where the reduction in gene expression is insufficient to modify the phenotype. This screen is intended to be a primary validation of the experimental methods of the microarray study and not an exhaustive investigation of pathogenic pathways in this model.

6.1 Modification of translated repeat phenotypes by cytoskeletal and trafficking components

Defects in axon transport processes have been hypothesised to play a role in polyglutamine disease pathogenesis due to the propensity of proteins containing polyglutamine tracts to cause axonal blockages. The transport of newly synthesised proteins and RNA as well as anterograde signals down the axon to the synapse and transport of retrograde signals and waste material in the opposite direction are essential to the function of neurons. Axonal transport also plays an important role in energy regulation via transport of mitochondria to provide local energy needs (258). Mutations in components of the axonal transport machinery, including motor proteins kinesin and dynein, have been demonstrated to be sufficient to cause neuronal dysfunction and death. Kinesin KIF1A mutations have been shown to cause disruptions to synaptic vesicle transport and cell death in mice (259) and the neuronal kinesin KIF1B is mutated in Charcot-Marie-Tooth disease type 2A, a neuropathy characterised by axonal degeneration resulting in progressive weakness and atrophy of muscles. Mutations in dynein heavy chain 1 have similarly been demonstrated to result in sensory neuropathy associated with “hind-limb clasping” in a mouse model (260). These observations suggest a general link between axonal transport dysfunction and neuronal death.

Components of the cytoskeleton and trafficking machinery were found to display altered expression in *Drosophila* expressing either rCAG or rCUG repeats pan-neuronally in both microarray experiment 1 and 2 (Chapter 5). Alterations to cytoskeletal components have been demonstrated to be amongst the early changes observed in HD brains (261) and may be involved in regulation of toxic protein aggregation rate, suggesting a broader role in the polyglutamine diseases (262). Altered splice-form expression of the microtubule-associated protein Tau, a protein involved in stabilisation of axons, has been detected in the brains of DM1 individuals (263), suggesting that alterations to the structure and function of the cytoskeleton may also play a specific role in CNS pathology in this disease. Expression of the transcript encoding the cytoskeletal protein Hts was upregulated in both rCAG and rCUG expressing flies compared to the *elav>+* control in microarray experiment 1 and downregulated in rCUG expressing flies compared to *elav>+* in experiment 2. While there is not concordance in the direction of the

observed change in expression, the consistent alteration to expression of *hts* between the microarray experiments suggests that cytoskeletal organisation is affected by the expression of expanded repeat RNA.

Trafficking defects have also been demonstrated in HD including perturbation of the retrograde transport of BDNF (264) and impairment of mitochondrial movement (265). It is unclear whether these defects are the result of a loss of normal HTT function, since HTT is involved in transport processes via its interaction with Huntington-associated protein 1 (HAP1), or axonal blockages caused by the polyglutamine aggregates themselves (13-14, 119, 264, 266-267). However the ability of polyglutamine tracts within a broad range of contexts to cause axonal transport defects suggests that these pathways may be commonly perturbed in the polyglutamine diseases. In a mouse model of SBMA, expression of mutant AR protein was found to result in reduced expression of dynactin 1, correlating with accumulation of neurofilaments and synaptophysin at the distal end of motor axons and retrograde transport defects resulting in neuronal dysfunction. The neuronal toxicity in this model could be mitigated by overexpression of dynactin 1, supporting the idea that the transport defects observed did play a major role in pathology (268).

Studies in *Drosophila* show that expression of either polyglutamine tracts alone or in the context of Ataxin-3 are able to induce changes in levels of components of the axonal transport machinery – including kinesin light and heavy chains, dynein light and heavy chains and the dynactin complex component $p150^{glued}$ – as well as accumulation of organelles consistent with axonal transport dysfunction (14). In our *Drosophila* model, no observable modification of a mild CAA polyglutamine phenotype was observed when mutations in *kinesin heavy chain*, *kinesin light chain*, *dynein heavy chain*, *roadblock*, $p150^{glued}$ or *dynamitin* were introduced (133), however this experiment did not investigate a role for CAG repeat RNA in the dysfunction. The results obtained from microarray analysis of *Drosophila* expressing rCAG and rCUG repeat tracts in neurons suggest that expanded repeat RNA alone may be capable of disrupting axonal transport pathways.

RNAi lines or mutants for the candidate genes *hts, ctp* and *insc* (as listed in Table 6.1) were obtained and tested for the ability to modify the CAG and CAA-encoded polyglutamine and CUG-encoded polyleucine eye phenotypes (Figure 6.1).

Using both CAG and CAA-encoded polyglutamine expressing flies will allow the distinction between modifications resulting from an interaction at the RNA level and those resulting from an interaction with the polyglutamine peptide itself, since CAA repeat RNA is not able to form a hairpin secondary structure in the same manner as CAG and CUG repeat RNA. The ability of polyglutamine-containing proteins to perturb axon transport and alter levels of cytoskeletal components is well-documented, however the focus of this experiment is to investigate the contribution of the hairpin RNA to transport defects.

It has been previously demonstrated that a *Drosophila* eye phenotype caused by expression of human *HTT* exon 1 containing 128 glutamines is suppressed by *hts* loss of function alleles 01103 and KG06777 (121). A slight suppression of both the CAG and CAA-encoded polyglutamine phenotypes was also observed in our model when these repeats were expressed with *GMR-GAL4* in flies heterozygous for either of these *hts* alleles (KG06777 shown, Figure 6.2 A&B compared to D&E). This suppression consists of a slight reduction in the area of the eye showing loss of pigment. The eye appeared slightly darker in flies expressing polyleucine and heterozygous for either of the *hts* alleles compared to eyes of flies expressing polyleucine in a wild-type background, which may also indicate a slight suppression (Figure 6.2, C compared to F). While there appears be a modification of both the CUG and CAG repeat phenotypes when expression of Hts is decreased, this effect is unlikely to be mediated through an interaction with hairpin-forming RNA since the phenotype associated with expression of a CAA-encoded polyglutamine tract, which cannot form a hairpin structure at the RNA level, was also suppressed by both *hts* alleles tested.

The ability of either *ctp* or *insc* to interact with polyglutamine has not been previously examined. Expression of RNAi constructs targeting either ctp or insc with *GMR-GAL4* did not cause a visible change to the appearance of the eye (data not shown). Co-expression of polyglutamine encoded by either CAG or CAA repeats with RNAi constructs targeting ctp resulted in a slight increase in the area of the eye showing loss of pigment compared to eyes where only polyglutamine is expressed (allele v43115 shown, Figure 6.2 A&B compared to G&H). No significant change was observed when an RNAi construct targeting insc was co-expressed with polyglutamine (Figure 6.2 A&B compared to J&K). There was also no dramatic

change in the exterior appearance of the eye in flies co-expressing the ctp or insc RNAi constructs with polyleucine compared to those expressing polyleucine alone (Figure 6.2 C compared to I & L). The enhancement observed when *ctp* levels were reduced in the eyes of polyglutamine-expressing flies is unlikely to be mediated through hairpin RNA, since it is observed in both CAG and CAA expressing flies, but not in CUG expressing flies.

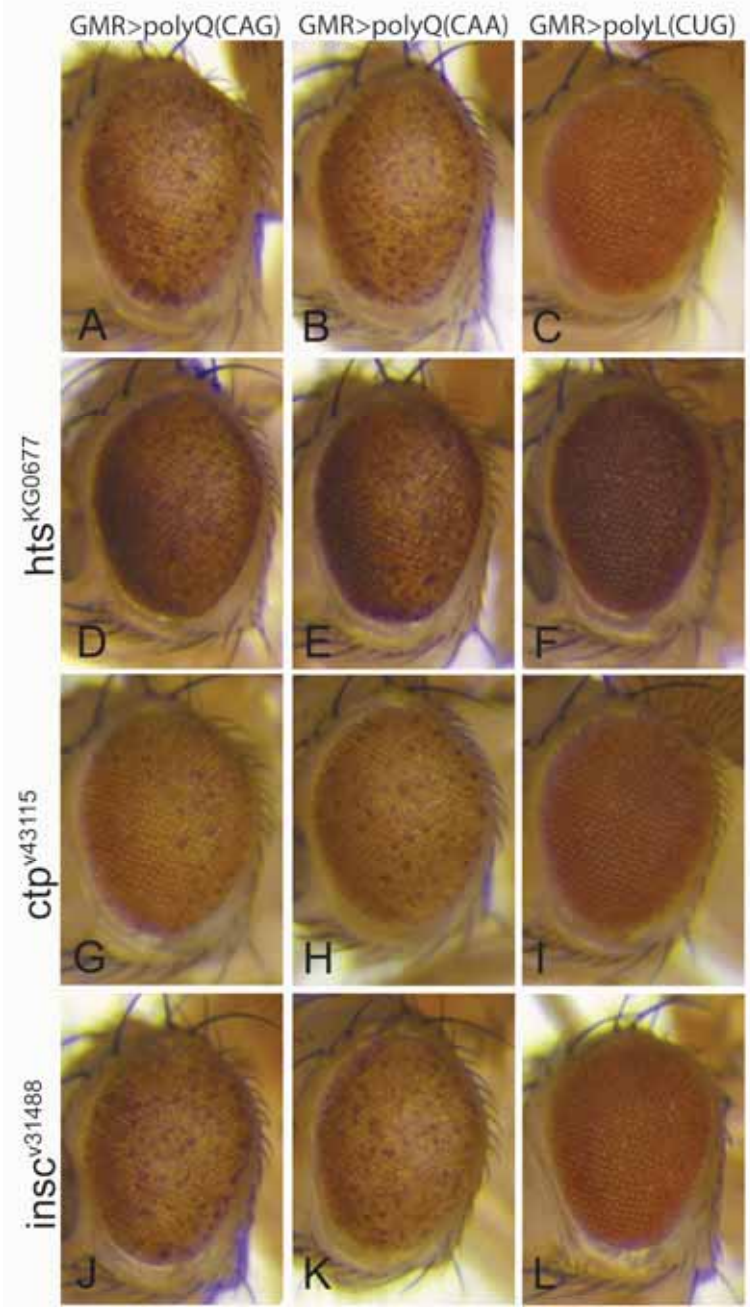


Figure 6.2: Effect of altering levels of cytoskeletal and trafficking components on polyglutamine and polyleucine eye phenotypes in *Drosophila*. In all cases repeat and RNAi construct expression is driven by *GMR-GAL4*. **A&B)** Expression of polyglutamine encoded by either CAG or CAA results in loss of pigment in the eye. **C)** Expression of polyleucine results in a mild rough eye phenotype. **D&E)** Flies heterozygous for a mutation in *hts* and expressing polyglutamine encoded by either a CAG or CAA repeat show a reduction in the area of loss of pigment compared to flies expressing polyglutamine in a wild-type background. **F)** Flies expressing polyleucine in a heterozygous *hts* mutant background have a darker eye colour compared to flies expressing polyleucine in a wild-type background. **G,H)** Co-expression of an RNAi construct targeting ctp with polyglutamine encoded by either CAG or CAA results in an slight increase in the area and severity of the loss of pigment eye phenotype. **J&K)** Co-expression of an RNAi construct targeting insc with polyglutamine encoded by either CAG or CAA does not significantly alter the exterior appearance of the eye. **I&L)** Co-expression of RNAi constructs targeting either ctp or insc with polyleucine does not dramatically modify the appearance of the eye compared to polyleucine expression alone.

The results presented here are not supportive of a role for expanded repeat RNA in the axonal transport defects observed in the polyglutamine diseases, since phenotypes resulting from expression of polyglutamine encoded by either an RNA hairpin-forming CAG repeat or a CAA repeat which is not able to form a secondary structure were both modified by reduced expression of *hts* and *ctp*. However, the ability of reduced expression of cytoskeletal proteins to alter the phenotypes associated with expression of polyglutamine in the eye does support a role for axonal transport proteins in polyglutamine pathogenesis in our model. The inability of reduced expression of kinesin light and heavy chain, dynein light and heavy chain and $p150^{glued}$ to modify a CAA-encoded polyglutamine phenotype (133) may be a result of differences between our model and the model in which the alteration of expression of these genes was originally observed and does not appear to be indicative of a lack of axon transport defects in our model.

6.2 Modification of translated repeat phenotypes by mod(mdg4), mGluRA and CG5669

The ability of co-expression of RNAi constructs targeting several other microarray candidates to modify the polyglutamine and polyleucine eye phenotypes were also examined (listed in Table 6.1). The choice of candidates to test in this initial screen was based on their identification in previous studies of expanded repeat disease. Of the candidates tested, two have been associated with untranslated repeat diseases: *mod(mdg4)* was identified in a modifier screen using a *Drosophila* model of SCA8 (98) and the metabotropic glutamate receptor (*mGluRA*) has been shown to have a mutually antagonistic relationship with the *Drosophila* FMRP orthologue (249). The remaining candidate, *CG5669*, encodes an orthologue of the SP1/SP3 transcription factor family. SP1 has been implicated in transcriptional dysregulation in HD (226-227) and is also downregulated in muscle biopsies from DM1 and DM2 patients (83) and therefore may represent a link between pathogenesis in the untranslated and translated repeat diseases.

Reducing expression of either *mGluRA* or *CG5669* did not cause a dramatic change to the exterior appearance of the eye in flies expressing polyglutamine encoded by either a CAG or CAA repeat (Figure 6.3 D,E & J,K compared to A,B).

This result does not support a role for either *CG5669* or *mGluRA* in polyglutamine pathogenesis in our *Drosophila* model, although it is possible that the level of knock-down elicited by these RNAi constructs is insufficient to cause a modification. A slight darkening of the eye was observed when either the CG5669 or mGluRA RNAi constructs were co-expressed with polyleucine however, since this is a very mild effect in both cases, it could simply be a result of decreased expression of the repeat construct resulting from titration of the available GAL4 protein by the UAS sites of the RNAi construct. It was therefore concluded that altering expression of *mGluRA* and *CG5669* does not have a significant effect on pathogenesis in our *Drosophila* model.

Expression of an RNAi construct targeting mod(mdg4) showed a strong interaction with both polyglutamine encoded by a CAG repeat and polyleucine (Figure 6.3 G&I). In the case of CAG-encoded polyglutamine, this enhancement resulted in nearly complete lethality of flies expressing both the RNAi construct and the CAG construct, while flies expressing either construct alone were viable. The observed lethality in these flies may be the result of some expression of these constructs in tissues other than the eye, an effect which has been previously reported for the *GMR-GAL4* driver (109). The few flies expressing both the mod(mdg4) RNAi construct and polyglutamine encoded by a CAG repeat which did eclose showed significantly greater area of loss of pigment and necrotic patches on the eye. Co-expression of the mod(mdg4) RNAi construct with polyleucine also resulted in a profound change in the colour of the eye and the appearance of necrotic patches. In contrast, expression of the RNAi construct with CAA-encoded polyglutamine resulted in only a relatively mild increase in the loss of pigment phenotype (Figure 6.3 H) despite the similar starting phenotypes in the chosen CAG and CAA polyglutamine lines (Figure 6.3 A&B).

Since *mod(mdg4)* was previously identified in a P-element screen for modifiers of a phenotype caused by expression of the human SCA8 non-coding RNA in the *Drosophila* eye (98), there is already support for an interaction between Mod(mdg4) and expanded repeat RNA. In our *Drosophila* model a much stronger interaction was observed with the CUG and CAG expanded repeat constructs than the CAA construct, suggesting that the sequence of the repeat may be important for this interaction. Mod(mdg4) has numerous biological roles mediated through highly

complex developmental splicing, including regulation of position effect variegation and silencing via chromatin remodelling as well as regulation of apoptotic pathways (269). Since Mod(mdg4) plays a role in chromatin remodelling, it is also possible that the interaction with the expanded repeats may be a sequence-dependent effect occurring at the DNA level. One possible mechanism for such an enhancement may be mediated via an increase in the expression of the repeat construct resulting from removal of Mod(mdg4)-mediated silencing.

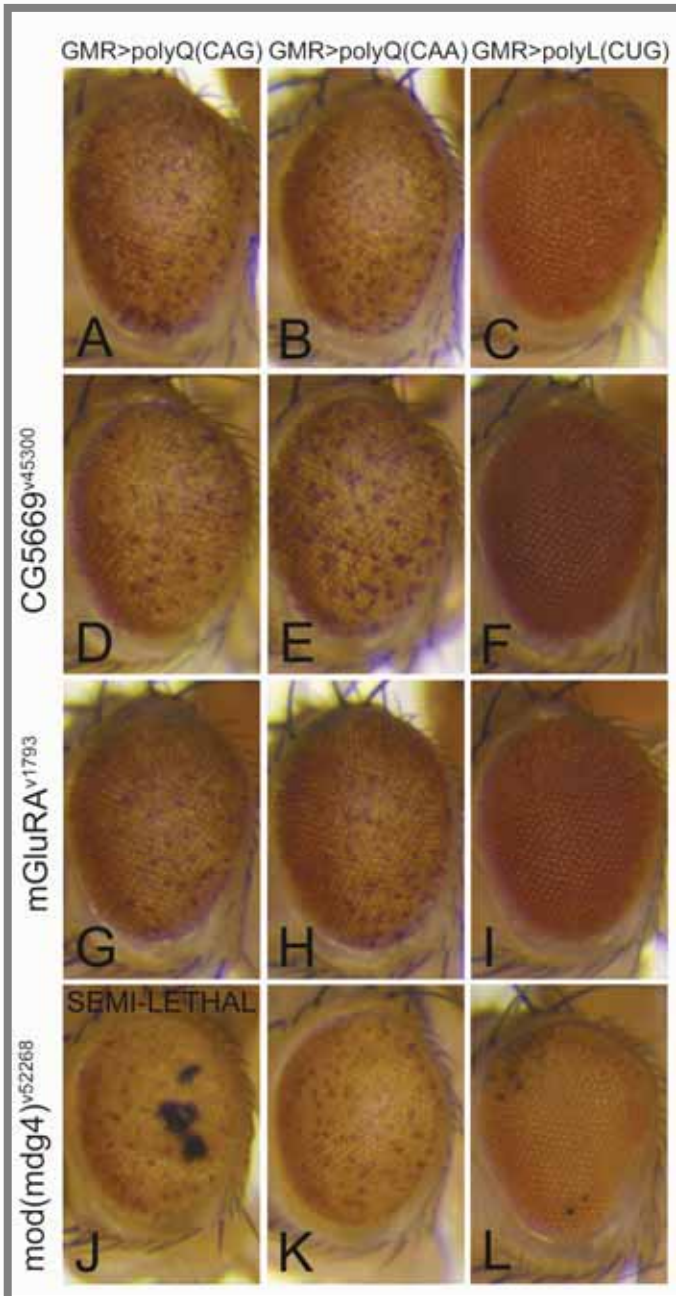


Figure 6.3: Modification of polyglutamine and polyleucine eye phenotypes by altering levels of Mod(mdg4), mGluRA and CG5669. In all repeat and RNAi construct expression was driven by *GMR-GAL4*. **A&B)** Polyglutamine lines with a similar loss of pigment phenotype were selected. **C)** Expression of polyleucine results in a mild rough eye phenotype. **D,E,G,H)** Co-expression of an RNAi construct targeting CG5669 or mGluRA with polyglutamine encoded by either CAG or CAA does not significantly alter the exterior appearance of the eye. **F&I)** Co-expression of an RNAi construct targeting CG5669 or mGluRA with polyleucine results in a slight darkening of the eye compared to expression of polyleucine alone. **J)** Co-expression of an RNAi construct targeting mod(mdg4) with polyglutamine encoded by CAG is semi-lethal. The few flies that do eclose have a severe loss of pigment eye phenotype with necrotic patches. **K)** Co-expression of an RNAi construct targeting mod(mdg4) with polyglutamine encoded by CAA results in a slight decrease in pigmentation compared to flies expressing CAA polyglutamine alone. **L)** Co-expression of an RNAi construct targeting mod(mdg4) with polyleucine causes a dramatic change in eye colour and appearance of necrotic patches on the eye.

6.3 Modification of translated repeat phenotypes by altering levels of *mbl* and *mef2*

Another gene identified by microarray analysis which has a known role in expanded repeat disease pathogenesis is *mbl.* While a role for the human *mbl* orthologue (*MBNL*) has been demonstrated in DM1 and DM2, a more general role for MBNL in other expanded repeat disorders has only recently been suspected. The first indication of this possibility related to the discovery that one of the spinocerebellar ataxias, SCA8, displays components of both polyglutamine and CUG repeat RNA pathology through bi-directional transcription of *ataxin-8* (97). The CUG repeat-containing transcripts in this disorder have also been shown to co-localise with MBNL1 suggesting that there may be overlap in pathogenesis with myotonic dystrophy (270). MBNL1 has also been shown to have similar affinity for both expanded CAG and CUG repeat tracts *in* vitro (116). The ability of *Drosophila mblA* to interact with CAG repeat transcripts *in vivo* has also been demonstrated, supporting a role for MBNL in polyglutamine disease pathology (117).

Despite evidence of mental impairment in many patients, the role of MBNL1 in DM1 has mostly been investigated in relation to muscle phenotypes and there is currently only limited evidence of a role for MBNL in neuronal pathology (271). The finding that *mbl* expression is altered in flies expressing rCAG or rCUG repeats pan-neuronally supports the idea that similar processes to those disrupted in muscle cells in myotonic dystrophy may be perturbed in *Drosophila* neurons in the presence of CAG or CUG repeat containing RNA. In order to validate a role for *mbl* in our model of expanded repeat disease, we tested the ability of an RNAi construct targeting mbl to modify the phenotypes associated with expression of polyglutamine or polyleucine in the eye (Figure 6.4 A-F). Co-expression of this RNAi construct enhanced the polyleucine phenotype, resulting in a decrease in the size of the eye and a glazed appearance, while only a very mild enhancement of the CAG-encoded polyglutamine eye phenotype was seen. No change was observed when *mbl* expression was reduced in flies co-expressing CAA-encoded polyglutamine in the eye, suggesting that this interaction is dependent upon the ability of the RNA to form a secondary structure. While the observation of a sequence-dependent interaction between polyglutamine and Mbl agrees with the findings of a recent study investigating the contribution of RNA toxicity to SCA3 pathogenesis (117), in our model altering levels

of Mbl produced only a very mild effect and therefore it is not clear whether this pathway represents a major component of polyglutamine pathogenesis in this case.

Closer examination of the microarray data also revealed altered regulation of *myocyte enhancing factor 2* (*mef2*) in flies expressing rCUG repeats compared to the *elav>+* control in microarray experiment 2. Mef2 is a developmentally regulated transcription factor involved in neuronal and muscle survival and development (272) which has also been demonstrated to regulate *mbl* expression in *Drosophila* (273). One mammalian MEF2 isoform, MEF2D, has been demonstrated in cultured rat neurons to be induced in response to stimulation of distal axons by neurotrophic signals. This result suggests that MEF2D is a component of the transcriptional response to retrograde signalling and is likely to be important in promoting neuronal survival (274). In a mouse model of Fragile X syndrome, MEF2 has been demonstrated to be involved in activity-dependent pruning of dendritic spines via an interaction with FMRP (275). This sort of role in neuronal plasticity has also been demonstrated in the striatal medium spiny neurons which are amongst the most vulnerable in HD (276). In mouse cerebellar neurons, MEF2 has also been demonstrated to co-localise with the wild-type Ataxin-1 protein. It is predicted that the presence of a CAG expansion in the Ataxin-1 protein may result in repression of MEF2 activity or sequestration of the protein in nuclear inclusions, resulting in a loss of the normal anti-apoptotic function of the protein (277). An interaction between MEF2 and the SP1 transcription factor has also been demonstrated (233), suggesting a link to the transcriptional dysregulation observed in both DM1 and HD.

Since there is a large amount of evidence to support a role for MEF2 in expanded repeat pathogenesis, the effect of reducing Mef2 levels in our *Drosophila* model was tested. Co-expression of an RNAi construct targeting mef2 with polyleucine strongly enhanced the eye phenotype, causing the appearance of a large number of black spots, discolouration and a reduction in the size of the eye (Figure 6.4 I). In contrast, co-expression of this RNAi construct with either CAG or CAA-encoded polyglutamine resulted in only a slight enhancement of the eye phenotype in both cases (Figure 6.4 G&H). This result appears to support a sequence-dependent interaction between *mef2* and CUG repeat RNA.

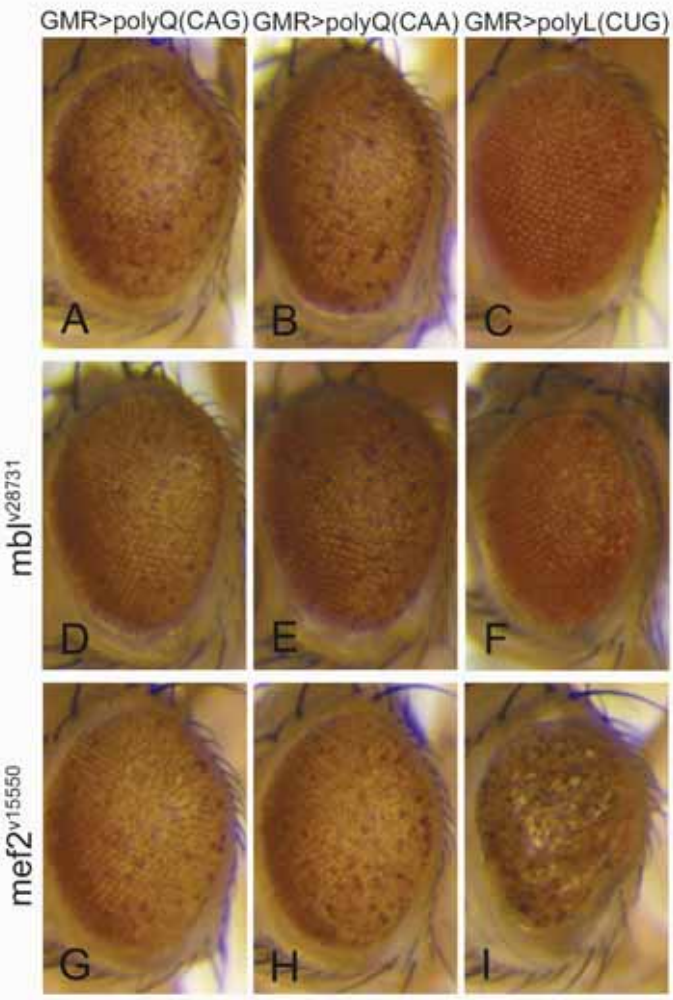


Figure 6.4: Modification of polyglutamine and polyleucine eye phenotypes in *Drosophila* by altering levels of Mbl and Mef2. RNAi and repeat expression are driven by *GMR-GAL4*. **A&B)** Expression of polyglutamine encoded by either CAG or CAA results in a loss of pigment eye phenotype. **C)** Expression of polyleucine results in a very mild rough eye phenotype. **D&G)** Co-expression of an RNAi construct targeting either mbl or mef2 with CAG encoded polyglutamine slightly enhances the loss of pigment eye phenotype. **E&H)** Co-expression of an RNAi construct targeting mef2 with CAA-encoded polyglutamine slightly enhances the loss of pigment eye phenotype while co-expression of an RNAi construct targeting mbl does not significantly alter the appearance of the eye. **F)** Co-expression of an RNAi construct targeting mbl with polyleucine causes an increase in the area of roughness and a decrease in the size of the eye. **I)** Co-expression of an RNAi construct targeting mef2 with polyleucine strongly enhances the eye phenotype resulting in a decrease in size and a glazed appearance with necrotic patches and loss of pigment.

6.4 Investigation of sequence-dependent interactions between expanded repeat RNA and Mef2, Mbl and Mod(mdg4) in *Drosophila*

In order to verify that the interactions observed between expanded repeats and *mbl, mef2* and *mod(mdg4)* are occurring at the RNA level, RNAi constructs targeting mef2 and mod(mdg4) were co-expressed with four transgene insertions of each of the untranslated repeat constructs (rCAG, rCUG and rCAA). Attempts to generate flies recombinant for *GMR-GAL4* and the mbl RNAi construct were not successful, possibly as a result of these insertions being within close proximity on the second chromosome, and therefore a deletion allele derived from a P-element insertion in *mbl* designated as allele E27 was tested with these repeat constructs. Flies heterozygous for this allele only have a 50% decrease in Mbl expression and this allele does not produce as great an enhancement of the polyleucine phenotype as expression of the mbl RNAi construct (data not shown). Flies heterozygous for the mbl^{E27} allele or expressing either the mod(mdg4) or mef2 RNAi constructs or the

rCAG, rCUG or rCAA repeat constructs alone do not show any disruption to the exterior appearance of the eye (Figure 6.5 A-D). Therefore only those gene expression changes which are rate-limiting to pathogenic pathways involved in RNA toxicity will be uncovered in this experiment, since reducing expression of this type of candidate would be expected to result in an increase in toxicity of the RNA species and therefore may result in the uncovering of a phenotype.

Co-expression of the mef2 RNAi construct with rCUG repeat RNA was found to be nearly entirely lethal. The few flies expressing both the mef2 RNAi construct and rCUG repeats that did eclose had a strong eye phenotype involving necrosis and a severe loss of ommatidial structure and died within one day (Figure 6.5 H). A similar effect was also seen when a second independent four transgene insertion line of rCUG was tested. The observed lethality is likely to be the result of some level of expression of the rCUG repeat RNA and mef2 RNAi construct in tissues other than the eye. No change to the appearance of the eye was observed when this RNAi construct was co-expressed with either rCAG or rCAA repeats (Figure 6.5 F&G). This result, along with the interaction of *mef2* with polyleucine, suggests a specific interaction between *mef2* and CUG repeat RNA which is consistent with the reduced levels of *mef2* expression observed only in rCUG expressing flies.

Similarly, co-expression of the mod(mdg4) RNAi construct and the rCUG repeat construct resulted in an eye phenotype consisting of loss of pigment and some roughening of the surface of the eye (Figure 6.5 L). No interaction was seen in flies co-expressing either rCAG or rCAA repeats with this RNAi construct (Figure 6.5 J&K) and therefore there appears to be a specific interaction between *mod(mdg4)* and the CUG repeat. The lack of an interaction between untranslated rCAG repeats and *mod(mdg4)* in this assay is surprising since expression of this RNAi construct with polyglutamine encoded by a CAG repeat resulted in a strong enhancement of the polyglutamine eye phenotype and lethality, an effect which was not seen in flies expressing polyglutamine encoded by a CAA repeat. Analysis of the steady-state RNA level in flies expressing rCAG and rCUG from these particular insertion sites suggests that a difference in the amount of hairpin repeat RNA present is not likely to be responsible for the lack of interaction observed in rCAG repeat expressing flies (S. Samaraweera, unpublished data). It is unclear whether this result suggests a real difference in the mechanism of pathogenesis of the rCAG and rCUG repeat

sequences, or just the degree of toxicity in this particular assay. Nevertheless, since expression of either the mod(mdg4) or mef2 RNAi constructs or rCUG RNA alone do not elicit a phenotype, the ability of co-expression of these constructs to dramatically disrupt the external appearance of the eye is strongly supportive of a genetic interaction between rCUG repeats and *mef2* and *mod(mdg4)*.

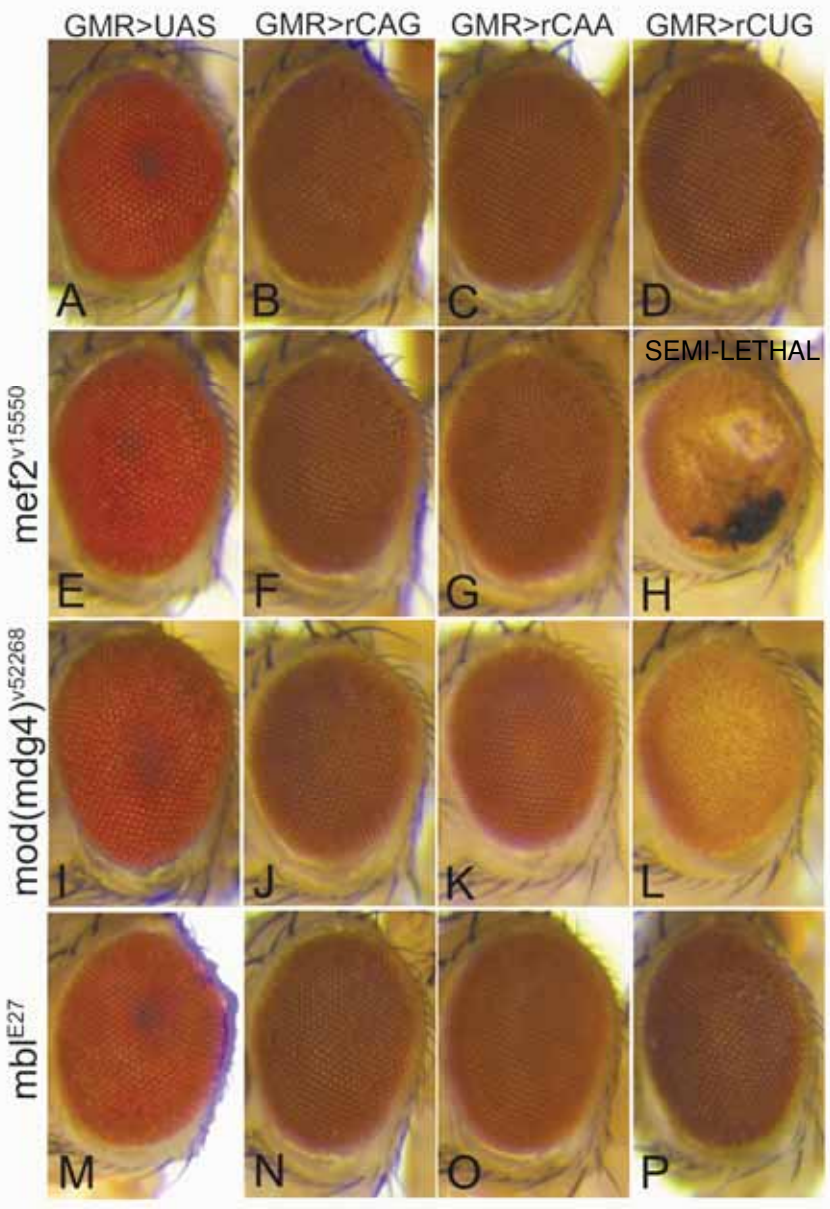


Figure 6.5: Mod(mdg4) and Mef2 show a sequence-dependent interaction with CUG repeat RNA in *Drosophila*. In all cases, expression of the RNAi construct and four transgene insertions of the repeat constructs are all driven by *GMR-GAL4*. **A-D)** Expression of four transgene insertions of rCAG, rCUG, rCAA or the UAS construct in the eye does not disrupt patterning of the eye. There is also no disruption to the normal patterning of the eye when expression of rCAG or rCAA repeats or the UAS construct without an insert are driven in the eye along with either the mef2 RNAi construct **(E-G)** or the mod(mdg4) RNAi construct **(I-K)**. **H)** Co-expression of rCUG repeats with the mef2 RNAi construct is nearly entirely lethal. The few flies that do eclose do not live for more than a day and have a strong eye phenotype consisting of large necrotic areas, a decrease in the size of the eye and nearly complete loss of ommatidial organisation. **L)** Co-expression of rCUG repeats with the mod(mdg4) RNAi construct causes a strong loss of pigment eye phenotype. **M-P)** Introduction of a null *mbl* mutation in a heterozygous manner does not alter the exterior appearance of eyes of flies co-expressing rCAG, rCUG or rCAG repeats or the UAS construct with no insert.

Introduction of mbl^{E27} into flies expressing rCAG, rCUG or rCAA repeats in the eye did not result in any alteration to the exterior appearance of the eye and therefore it was not possible to explore the role of expanded repeat RNA in the sequence-

dependent interaction which was observed between Mbl and the translated repeats. Since this allele also showed a weaker interaction with both polyleucine and CAG-encoded polyglutamine than the RNAi construct (data not shown), the absence of an interaction may be the result of an insufficient reduction in Mbl expression in this case.

6.5 Evidence of a role for *MBNL1* in expanded repeat disease pathogenesis.

A significant amount of data regarding the interaction of the human orthologue of *mbl* (*MBNL*) with expanded repeat RNA in our model had already been accumulated prior to performing the microarray study because of the link between *MBNL* and myotonic dystrophy pathology. There are three isoforms of human MBNL (MBNL1-3) and, while it is unclear how much overlapping function they have in regulation of muscle development and splicing, there is a pool of evidence to suggest that MBNL1 at least is involved in the formation of foci in DM1 and DM2 (278-279). To verify genetically that MBNL1 is involved in expanded repeat pathogenesis, overexpression of a UAS-MBNL1 construct (obtained from (86)) was performed in the *Drosophila* eye using *GMR-GAL4*. Driving expression of *MBNL1* in the eye produces a rough eye phenotype at 23ºC, however this phenotype is nearly completely suppressed by growing the flies at 18 ºC. In order to investigate the ability of *MBNL1* to interact with our expanded repeat constructs, flies carrying the *UAS-MBNL1* construct were recombined with *GMR-GAL4* to produce the *GMR>MBNL1* stock (L. O'Keefe).

Consistent with what is known about the ability of CUG repeat RNA to interact with MBNL, expression of polyleucine encoded by a CUG repeat was able to slightly enhance the phenotype seen in *GMR>MBNL1* flies, causing an increase in roughness and darkening of the eye (Figure 6.6 C&D). This effect does not appear to be additive since there is no phenotype in flies expressing polyleucine alone at this temperature (Figure 6.6 B). If expression of CUG repeat RNA causes sequestration of MBNL1, it would be expected that co-expression of this repeat RNA with MBNL1 would result in a decrease in MBNL1-associated pathology in our model. The observed enhancement may therefore be a result of the effects of both an alteration to the transcriptional profile of the *Drosophila mbl* gene and the ectopic expression of the human MBNL1 protein in these flies. It is unclear how the function of the human MBNL isoforms

corresponds to the *Drosophila* Mbl isoforms and therefore this result is difficult to interpret.

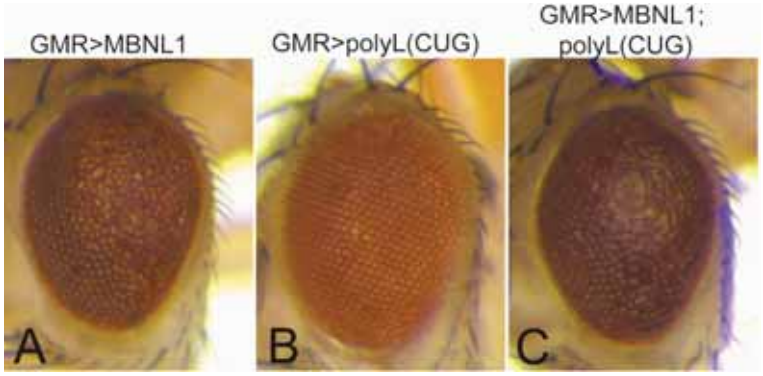


Figure 6.6: Co-expression of CUG-encoded polyleucine enhances the MBNL1 eye phenotype. All flies express UAS constructs specifically in the eye under the control of the *GMR-GAL4* driver and have been grown at 23 ºC. **A)** Expression of MBNL1 alone results in a mild roughening of the surface of the eye. **B)** Expression of polyleucine in flies grown at 23 ºC does not result in a disruption to the structure of the eye. **C)** Co-expression of polyleucine with MBNL1 causes an increase in the area and degree of roughness and a slight darkening of the colour of the eye.

In order to further investigate a role for MBNL1 in the polyglutamine diseases, the effect of co-expressing MBNL1 with a polyglutamine tract was also investigated. Crosses were performed at 18 ºC since no phenotype is observed in flies co-expressing GFP and MBNL1 at this temperature (Figure 6.7 D). Expression of either a CAG-encoded or CAA-encoded polyglutamine tract in the eye at this temperature results in a mild rough eye phenotype consisting mainly of loss of pigment (Figure 6.7 B, C). Co-expression of MBNL1 in either of these cases enhances this phenotype, resulting in a significant increase in the area of the eye affected and the extent of the pigment loss (Figure 6.7 E, F). MBNL1 therefore shows a sequence-independent interaction with polyglutamine in this model, suggesting that this effect is not mediated at the RNA level.

Li *et al*. (117) previously demonstrated that polyglutamine encoded by a pure CAG repeat in the context of a truncated SCA3 transcript is more toxic than polyglutamine encoded by mixed CAG/CAA repeats. Furthermore, they reported that ectopic expression of *mblA* was able to modify this CAG repeat phenotype, but not the phenotype of the mixed CAG/CAA repeat. They concluded that an interaction at the RNA level was at least partly responsible for the enhanced toxicity when *mblA* was overexpressed. This result is consistent with our observation of a sequence-dependent modification of the phenotype associated with expression of polyglutamine encoded by a CAG repeat when Mbl levels are altered, however in our model reducing expression of *Drosophila* Mbl, not overexpression, resulted in an enhancement and

this effect was very mild. Since the overexpression construct used in the study performed by Li *et al*. encodes only one isoform of Mbl and the RNAi construct used in this study targets all known Mbl isoforms, the differences observed may demonstrate perturbation of specific functions of *mbl* depending upon which isoforms are expressed in each case.

In further contrast to what was seen in the SCA3 study, our model shows polyglutamine to be intrinsically highly toxic whether encoded by CAG or CAA repeats, suggesting that the context of the repeat tract may play a vital role in determining toxicity. It was not possible to ascertain whether ectopically expressed human MBNL1 was able to specifically interact with CAG repeat RNA in a sequence-dependent manner using this model of intrinsic polyglutamine toxicity, since a strong enhancement of the polyglutamine eye phenotype was seen when either a pure CAG or CAA repeat tract was expressed (Figure 6.7). There are several explanations for the lack of sequence-specificity observed when human MBNL1 was overexpressed with polyglutamine in our model. Firstly, it is possible that this effect is due to a real difference in binding properties of the human and *Drosophila* proteins at the RNA level. This could also somewhat explain the differences in sequence-specificity between our model and the SCA3 model, since the context of the repeat tract may influence the binding capabilities of the MBNL1 protein. It is also possible that the enhancement of the polyglutamine eye phenotypes is simply a dominant effect resulting from the effects of ectopic expression of MBNL1.

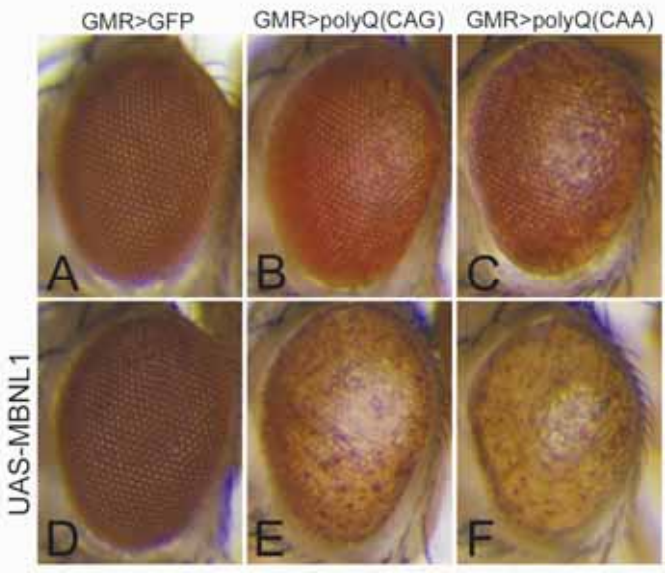


Figure 6.7: Overexpression of MBNL1 enhances both CAA and CAG-encoded polyglutamine eye phenotypes in *Drosophila.* All flies express UAS constructs specifically in the eye under the control of the *GMR-GAL4* driver and have been grown at 18 ºC. Photographs taken by L. O'Keefe. **A)** Expression of GFP in the eye does not disrupt the structure of the eye. **B&C)** Expression of a poly-glutamine tract encoded by CAA or CAG perfect repeats results in a mild eye phenotype consisting of loss of pigment and some disorder of ommatidial arrays. **D)** Co-expression of MBNL1 with GFP does not disrupt the structure of the eye at 18 ºC. **E&F)** Co-expression of MBNL1 with poly-glutamine encoded by either CAG or CAA repeats enhances the eye phenotype causing greater loss of pigment and a greater area of disruption.

In order to investigate the contribution of hairpin RNA to the interactions seen between MBNL1 and polyglutamine and polyleucine, *GMR>MBNL1* flies were also crossed to the untranslated repeat stocks (rCAG, rCUG and rCAA respectively) to produce flies that express MBNL1 and untranslated repeat RNA. In this experiment, flies carrying four insertions of the repeat construct were used to ensure that high levels of expression of the repeat were achieved. For each repeat, two independent four transgene insertion lines were tested (listed in Appendix B, Table B2). A fly stock containing four UAS construct insertions without any transgene present (called UAS) was used as a control to ensure that similar levels of free GAL4 and expression of MBNL1 are achieved.

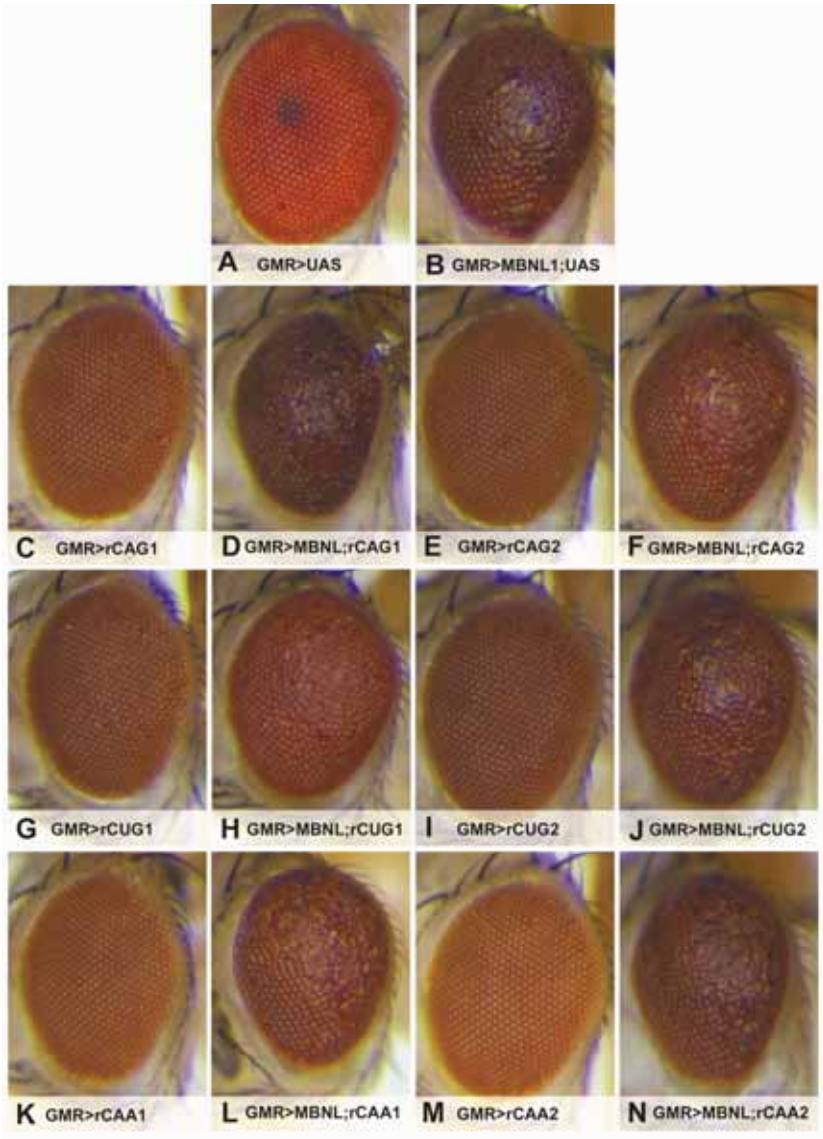


Figure 6.8: Expression of expanded untranslated CAG, CUG and CAA repeats in *Drosophila* overexpressing MBNL1. In all cases flies were grown at 23 ºC and expression of the repeat constructs, the UAS control and *MBNL1* were all driven in the eye by *GMR>GAL4*. All photographs were taken by S. Samaraweera. Repeat line genotypes are listed in Table 3.6. **A,C,E,G,I,K,M)** Expression of either the repeat constructs or the *UAS* control alone do not produce any phenotype. **B**) Co-expression of *MBNL1* with the UAS control produces a rough eye with necrotic patches and reduced size compared to UAS alone. **D,F,H,J,L,N)** Co-expression of *MBNL1* with 2 independent four transgene insertion lines of each of rCAG, rCUG or rCAA repeats did not produce a consistent alteration to the phenotype seen in flies co-expressing MBNL1 and UAS. Eyes of flies expressing rCAA repeats (L&N) are indistinguishable from those of rCAG2(F) and rCUG2(J). Eye phenotypes in flies co-expressing rCAG1 and rCUG1 also look markedly different from the other rCAG and rCUG independent lines tested (D&H).

At 23 ºC, *GMR-GAL4* driven expression of four insertions of each of UAS, rCAG, rCUG or rCAA alone do not produce any disruption to the structure of the eye (Figure 6.8 A,C,E,G,I,K,M). Overexpression of *MBNL1* in flies carrying four transgene insertions of the UAS control results in a reduction in the overall size of the eye and roughening of the surface with some necrotic patches which appear as black spots (Figure 6.5 B). This phenotype appears to be somewhat suppressed in flies expressing one of the rCAG four transgene insertions (rCAG2) or either of the independent rCUG four transgene insertions (rCUG1 & rCUG2), however expression of either rCAA1 or rCAA2 results in a similar suppression, involving an increase in the size of the eye and decrease in the area of roughness and number of necrotic patches (Figure 6.8 F, H and J compared to L and N) and therefore this effect does not appear to be related to the ability of the expressed RNA species to form a hairpin structure. A stronger effect is seen when rCUG1 is expressed than for any other line (Figure 6.8 H), with nearly a complete suppression of the roughness and a return to wild-type size eye. The expression of rCAG1 appears to slightly enhance the MBNL1 phenotype, causing a reduction in the size of the eye and an increase in the area of roughness (Figure 6.8 B compared to D). These results do not support a sequence-dependent interaction between rCAG or rCUG repeats and human MBNL1 in this model.

6.6 Summary of results from genetic screen of microarray candidates

Interactions were observed between a number of the candidates identified in the microarray study and our translated expanded repeat constructs in the *Drosophila* ey*e* (summarised in Table 6.2). In a number of cases, candidate genes showed an interaction with both CAG and CAA-encoded polyglutamine, suggesting that this effect is not mediated through an interaction with hairpin RNA. RNAi constructs for two candidates which did show a sequence-dependent interaction with expanded repeats, *mef2* and *mod(mdg4)*, were also tested with the untranslated rCAG, rCUG and rCAA repeat constructs. A strong interaction was observed with CUG repeat RNA but not with CAG or CAA RNA in these cases (Figure 6.5). This result supports a role for Mef2 and Mod(mdg4) in CUG-repeat RNA toxicity in the expanded repeat diseases. Since altered expression of m*ef2* was only seen in flies expressing rCUG repeats neuronally, this result is consistent with Mef2 playing a

unique role in CUG repeat pathogenesis. However, altered expression of *mod(mdg4)* was observed in both rCAG and rCUG repeat-expressing flies and therefore it is unclear whether the absence of an interaction between rCAG and *mod(mdg4)* in the eye indicates that this is a more important component of CUG repeat pathogenesis, or that different pathways are perturbed in flies expressing CAG repeats in the eye compared to the nervous system. The observation that reducing expression of Mod(mdg4) in flies co-expressing translated CAG repeats encoding polyglutamine resulted in nearly complete lethality suggests that there is likely to be a strong interaction between *mod(mdg4)* and CAG repeats in tissues other than the eye. Since this effect was not seen in flies co-expressing the mod(mdg4) RNAi construct and polyglutamine encoded by a CAA repeat which cannot form a hairpin RNA, this interaction may be mediated via an interaction with the hairpin-forming CAG repeat RNA. Since Mod(mdg4) has been demonstrated to play a role in position effect variegation and transcriptional silencing (269), it is also possible that this interaction is mediated through an interaction with CAG and CUG repeat tracts at the DNA level.

A mild interaction was also observed between *Drosophila mbl* and polyglutamine encoded by a CAG repeat and polyleucine encoded by a CUG repeat, but not polyglutamine encoded by a CAA repeat (Figure 6.4). This result supports a role for Mbl in both CAG and CUG repeat toxicity in our model and suggests that the secondary structure of the RNA is important for this interaction. However, since no interaction was observed when rCAG or rCUG repeats were expressed in a heterozygous *mbl* null background, the biological importance of this interaction in RNA toxicity could not be further investigated. The ability of human MBNL1 to interact with repeat RNAs in a sequence-dependent manner was not supported in our *Drosophila* model. While an enhancement of the phenotype associated with overexpression of human MBNL was seen when CUG-encoded polyleucine was co-expressed (Figure 6.6), a sequence-independent enhancement of both CAG and CAA-encoded polyglutamine phenotypes was also seen (Figure 6.7) which may suggest that human MBNL1 has a dominant toxic effect in our model. Co-expression of MBNL1 with rCAG, rCUG and rCAA repeat constructs did not give a consistent effect between independent transgenic lines (Figure 6.8) and therefore any role of the repeat RNA in the interactions observed between polyglutamine and polyleucine with MBNL1 in this model remains unclear.

The two CUG repeat sequence-dependent interactions uncovered in this analysis are likely to represent real components of RNA toxicity in this *Drosophila* model, since these candidates were identified as showing altered transcription in flies expressing untranslated repeat RNA neuronally and were also genetically verified in flies expressing both translated and untranslated CUG repeats in the eye. The results of this primary screen therefore suggest that further analysis of the microarray data may identify more components of toxicity in both CUG and CAG repeat-expressing flies and that the use of this kind of approach is likely to yield biologically relevant results.

Gene	PolyQ (CAG)	PolyQ (CAA)	PolyL (CUG)	rCAG	rCAA	rCUG
hts	S+	S+	S+	-	-	-
ctp	NS	NS	NS	-	-	-
insc	E+	E+	NS	-	-	-
CG5669	NS	NS	S+	-	-	-
mod(mdg4)	**E+++ (lethal)**	**E+**	**E++**	**NS**	**NS**	**E+**
mGluRA	NS	NS	NS	-	-	-
mef2	**NS**	**NS**	**E++**	**NS**	**NS**	**E+++ (lethal)**
mbl	**E+**	**NS**	**E++**	**NS**	**NS**	**NS**
MBNL1	**E++**	**E++**	**E++**	**NS**	**NS**	**NS**

Table 6.2 Summary of results from genetic screen of microarray candidates. Alleles are all RNAi or P-element insertions which reduce expression of the candidate gene, except for MBNL1 which is an overexpression construct. Alleles tested are listed in Table 7.1. Dashes indicate that interactions were not tested in this study. S=suppression of eye phenotype, E=enhancement of eye phenotype, NS=no significant change to appearance of eye, + indicates a mild interaction, ++indicates a medium strength interaction, +++ indicates a strong interaction.

Chapter 7: Spinocerebellar ataxia 10: a unique untranslated repeat disease?

Spinocerebellar ataxia type 10 (SCA10) is a rare cerebellar ataxia caused by the expansion of a pentameric ATTCT repeat within exon 9 of the *ataxin-10* gene (280). This mutation is believed to have arisen in Latin American populations and SCA10 largely affects individuals of Brazilian and Mexican origin (281). Like other SCAs, SCA10 presents as cerebellar dysfunction resulting in ataxia with other features including cognitive impairment, dementia and seizures in a proportion of patients (282). Genetic anticipation with a bias towards expansion on paternal transmission and a correlation between repeat size and age-at-onset have also been observed in some families (283). Cases of repeat-size mosaicism within tissues and repeat-size variability between tissues are common, suggesting that these repeats are also highly somatically unstable (283).

Despite resulting in a fairly pure cerebellar ataxia, there are several unique features of the SCA10 mutation. Firstly, the disease-causing expansions in *ataxin-10* are extremely large: generally more than 800 repeats and frequently many thousands of repeats are detected in affected individuals (280). The repeat itself is also unusual, firstly because it is very AT rich but also because it is the only SCA-causing repeat tract present within the intron of a gene. These features have a number of implications for disease pathogenesis. Examination of the behaviour of the SCA10 repeat at the DNA level has revealed that it is a DNA unwinding element (284-285) and that the ATTCT strand, but not the anti-sense TAAGA strand, is able to form a secondary structure under physiological conditions (100). This propensity of the repeat tract to unwind is believed to be part of the repeat expansion mechanism responsible for the large size of the SCA10 repeat expansions (285).

The outcomes of the ATTCT repeat expansion in *ataxin-10* are largely unknown, however the discovery that the repeat itself is a DNA unwinding element caused speculation that the expression of *ataxin-10* and surrounding genes might be altered by the expansion. Similar to CAG repeats, the expanded ATTCT repeat has been associated with strong binding of nucleosomes which is further enhanced by the presence of interruptions to the repeat sequence, supporting the idea that the repeat may alter gene regulation (286). The Ataxin-10 protein has been demonstrated to be essential for the survival of primary cerebellar neurons in culture (6, 287), however it has been reported that the mutant *ataxin-10* allele is both

transcribed and processed normally in patient-derived cells, suggesting that loss of function alone is not responsible for the pathogenesis of SCA10 (288). Furthermore, while mice null for *ataxin-10* are embryonic lethal, heterozygotes do not recapitulate any SCA10 symptoms, supporting the idea that a simple gene-dosage effect is unlikely to be responsible for pathogenesis (288).

Since the SCA10 repeat resides within an untranslated region of the *ataxin-10* gene and the disease exhibits dominant inheritance, it has been proposed that this mutation may be pathogenic because it results in the production of a dominant toxic RNA (100). While rCAG/rCUG repeats have been demonstrated to form a simple hairpin structure with a mis-match every third base, the only structure consistent with NMR spectroscopy data for rAUUCU repeats under physiological conditions is an anti-parallel hairpin including a C-C mismatch every 5 bases and an equal ratio of A-U/U-U matches (100). In the context of the SCA10 repeat, the presence of U-U mis-matches is predicted to stabilise the hairpin secondary structure (100). Splicing of the SCA10 transcript may result in the release of the extremely large hairpin-forming AUUCU RNA in the cell which may have the potential to bind RNA binding proteins inappropriately in a similar manner to the CUG/CCUG repeats in DM1 and DM2 respectively.

7.1 Modelling SCA10 in *Drosophila*

The observation that expanded AUUCU repeat RNA is also able to form a hairpin structure supports our hypothesis that RNA secondary structure may be involved in pathogenesis of the expanded repeat diseases. In order to investigate the contribution of rAUUCU repeat-containing RNA to pathology in SCA10, a *Drosophila* model was generated. To do this, the intron 9 region from *ataxin-10* containing the repeat tract and 141 bp of surrounding sequence was amplified from HeLa cell DNA and the ATTCT repeat expanded from a starting size of 13 repeats using a PCR method. Both GFP-tagged and untagged constructs were generated as depicted in Figure 7.1. These constructs contain an expanded ATTCT repeat either in the 3'UTR of a short peptide, as described for rCAG, rCUG and rCAA constructs, or in the 5'UTR of the GFP transcript. Repeat tracts of 65 repeats for the 3'UTR insertion and 67 repeats for the GFP tagged construct were obtained and completely

sequenced. A clone for the GFP-tagged construct was also obtained which gave a PCR product of the expected size for a repeat tract of around 100 repeats. On sequencing at least 100 repeats were detected, however it was not possible to completely sequence across the repeat tract, presumably because of the AT-rich nature of the sequence. In the case of the repeat tracts which were completely sequenced, the injected constructs also contained interruptions to the repeat tract as listed in Table 7.1. The AT-rich nature of the sequence and the instability observed during cloning once the repeat number exceeded this range may have prevented the generation of repeat tracts within the range seen in SCA10 individuals.

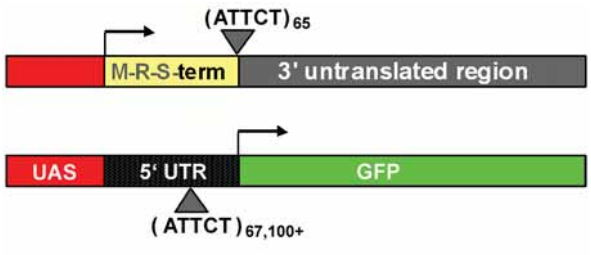


Figure 7.1: Constructs generated to model SCA10 pathogenesis in *Drosophila*. A) A construct was generated with 65 ATTCT repeats inserted into the 3'UTR of a short peptide. **B)** Constructs were also generated with the ATTCT repeat tract inserted upstream of a GFP tag, such that the repeat is effectively in the 5'UTR. Repeat tracts of 67 and at least 100 repeats were generated in this case.

Construct	Repeat sequence	Number of lines generated
$rAUUCU_{65}$	$(ATTCT)_{20}$ ACTCT $(ATTCT)_{23}$ ATTCC $(ATTCT)_{15}$ ATTTT $(ATTCT)_{7}$	7
$rAUUCU_{67}$-GFP	$(ATTCT)_{20}$ ACTCT $(ATTCT)_{25}$ ATTCC $(ATTCT)_{15}$ ATTTT $(ATTCT)_{7}$	9
$rAUUCU_{100+}$-GFP	-	8

Table 7.1: rAUUCU repeat constructs injected into *Drosophila*. Constructs are represented in Figure 8.1A for rAUUCU65 and Figure 8.1B for rAUUCU67-GFP and rAUUCU100+-GFP. Interruptions were introduced into repeat tracts during expansion. In all cases, multiple independent lines were generated by random P-element mediated integration.

Although the repeat copy number of the longest un-tagged rAUUCU repeat construct generated is less than those generated for rCAG/rCUG repeats, which contain 93 and 114 repeats respectively, the actual length of the repeat tract is approximately 300 bp in all cases and therefore the hairpin formed by this RNA would be predicted to be approximately the same size as those formed by the

rCAG/rCUG repeats. Furthermore, while the secondary structures formed by AUUCU repeats are not as stable as those formed by similar lengths of repeats associated with other expanded repeat diseases, rAUUCU RNA containing as few as 9 repeats has been demonstrated to form a hairpin structure under physiological conditions *in vitro* (100) and therefore a repeat tract of this size should form a stable secondary structure capable of eliciting dominant toxic effects. Expression of up to four transgene insertions of either $rAUUCU_{65}$ or the largest GFP-tagged construct, $rAUUCU_{100+}$-GFP, either in the eye using the *GMR-GAL4* driver (Figure 7.2) or in the nervous system using the *elav-GAL4* driver did not result in a phenotype.

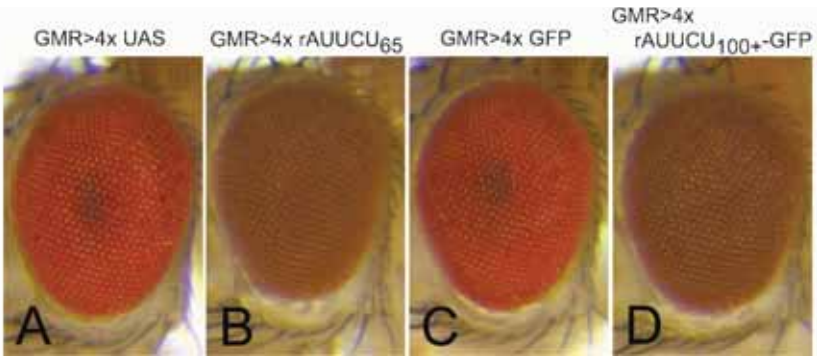


Figure 7.2: Expression of four transgene insertions of an expanded ATTCT repeat does not alter the exterior appearance of the *Drosophila* eye. Expression of constructs is driven by *GMR-GAL4*. **A&C)** Driving expression of four transgene insertions of GFP or the UAS construct without an insert (UAS) results in eyes of wild-type appearance. **B&D)** Expression of four transgene insertions of either $rAUUCU_{65}$ or the largest GFP-tagged construct ($rAUUCU_{100+}$-GFP) does not cause a phenotype in the eye.

7.2 Investigation of cellular localisation of expanded rAUUCU repeats

One of the mechanisms by which RNA is suggested to act as a pathogenic agent in the expanded repeat diseases is through formation of RNA foci which may result in the sequestration of RNA binding proteins. Localisation of expanded CUG repeat-containing RNAs to foci has been demonstrated in a number of studies using either patient tissues or animal models of DM1 (72, 81, 86, 289-290), DM2 (81), HDL-2 (107) and SCA8 (270). The splicing factor MBNL has been demonstrated to be present in foci with CUG repeat-containing RNA both in human tissue and in *Drosophila* models (72, 86, 147, 278), supporting the idea that foci formation is involved in pathogenesis and can explain splicing defects observed in DM1 and DM2. More recently, CAG repeat-containing RNA has also been demonstrated to be able to form RNA foci in *Drosophila* (71) and it has therefore been suggested that the ability to

aggregate in this manner is a more general property of the expanded repeat-containing RNAs. It is not clear whether the formation of foci necessarily correlates with pathogenesis however, since altered splicing was observed in the presence of CUG repeats irrespective of whether foci were evident and was not observed in the presence of CAG repeats even when foci were evident (71).

Since the expanded rAUUCU RNA present in SCA10 has also been demonstrated to form a stable secondary structure (100), we examined the ability of this RNA to localise to foci in *Drosophila* cells. We expressed four transgene insertions of untranslated GFP-tagged CUG, CAA and AUUCU repeats ubiquitously using the *da-GAL4* driver and cut 10 µm sections from 3rd instar *Drosophila* larvae. In each case, we detected the repeat-containing RNA using a Cy3 labelled oligonucleotide probe against the GFP tag – since this allowed us to detect all three repeats under the same hybridisation conditions – and co-stained the samples with DAPI to mark the nuclei of cells. We frequently observed nuclei with four hybridised spots in cells expressing each of the different repeat sequences (Figure 7.3 B, C, E filled arrow heads) which we predict are unlikely to be RNA foci, but rather are sites of transcription correlating to the four transgene insertion sites. This is supported by the observation that larvae expressing GFP without a repeat sequence also show hybridised spots in many cells, with the number of hybridised spots observed corresponding to the number of transgene insertions being expressed (K. Lawlor, unpublished data).

In sections from larvae expressing rCUG-GFP (Figure 7.3 B) and rAUUCU-GFP (Figure 7.3 D-F) but not rCAA-GFP (Figure 7.3 C), we also observed a subset of nuclei which contain a large number of hybridised spots. We predict that these cells are likely to be muscle cells because of their morphology. It therefore appears that rAUUCU repeat-containing RNA is able to aggregate in a similar manner to rCUG repeat RNA in *Drosophila* and, since aggregation was not observed in cells expressing rCAA repeat-containing RNA, that this propensity to aggregate may be related to the ability of the RNA to form a stable secondary structure. While these foci were only observed in a subset of cells, this may be more indicative of the limitations of the detection method and does not necessarily reflect a lack of foci formation in other cells such as neurons. It is unclear how these foci in non-neuronal cells relate to pathogenesis in neuronal cells in SCA10, however this result points towards common

behaviour of expanded repeat-containing RNA in cells and supports the hypothesis that repeat expansions with different sequences may have similar effects on cellular biology.

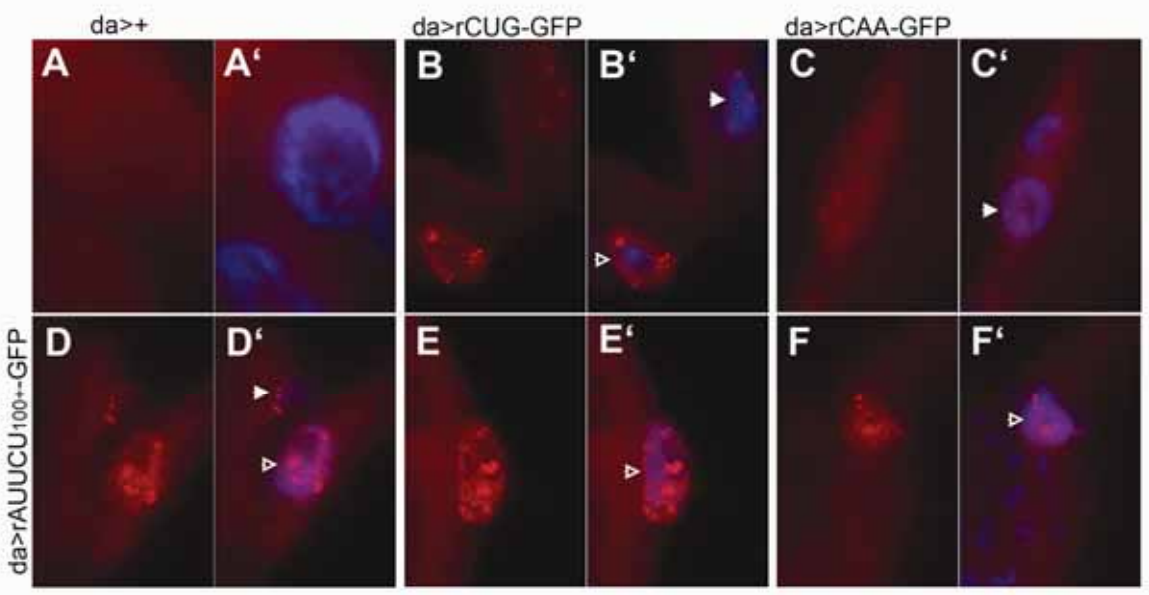


Figure 7.3: rAUUCU repeat-containing transcripts form aggregates in a sub-set of *Drosophila* cells. In each case, 3rd instar larvae expressing four transgene insertions of GFP-tagged repeats driven by *da-GAL4* were frozen and 10 µm cryostat sections were performed. All sections cut by K. Lawlor. In all cases, a Cy3-labelled oligonucleotide probe specific for the GFP-tag was hybridised. **A-F** show Cy3 signal, **A'-F**' are merged images showing DAPI staining overlaid on the Cy3 signal. **A, A')** The *da-GAL4* stock was out-crossed to a wild-type stock (*w^{1118}*) to generate *da-GAL4>+* larvae. No specific probe hybridisation was detected in this case, even when the image was over-exposed. **B, B')** A number of CUG-repeat expressing nuclei were observed containing a large number of hybridised spots (un-filled arrowhead). Many cells were also observed showing only four hybridisation spots (filled arrowhead). **C, C')** Expression of CAA repeat RNA tagged with GFP did not result in the appearance of RNA aggregates. Some cells were observed with up to four hybridisation spots (filled arrowhead), again probably corresponding to the location of transcription from the four transgene insertions. **D, D'-F, F')** Expression of AUUCU repeat RNA resulted in many cells showing four hybridisation spots (filled arrowhead). A subset of cells was observed containing multiple hybridisation spots (un-filled arrows). These spots were not as defined as those observed for the CUG repeat **(B)** but were distinctly different in appearance to those observed when CAA repeat RNA was expressed **(C)**.

7.3 Identification of transcriptional changes in neuronal cells resulting from expression of SCA10 repeats

Whilst rAUUCU RNA has previously been demonstrated to have the ability to form a complex hairpin secondary structure *in vitro* (100) and appears to aggregate in a similar manner to CUG repeat RNA in at least a sub-set of cells in our *Drosophila*

model, the cellular outcomes of expression of the expanded SCA10 repeat have never been investigated. If there is a common pathogenic mechanism involving RNA toxicity for the expanded repeat diseases, expression of expanded rAUUCU repeat RNA would be expected to elicit similar transcriptional changes as those observed for rCAG and rCUG expanded repeat RNAs. Microarray analysis was therefore performed on flies expressing the $rAUUCU_{65}$ construct to investigate cellular changes resulting from the expression of the SCA10 expanded repeat in *Drosophila* neurons. This experiment was of dual purpose: firstly to test whether expression of rAUUCU RNA is sufficient to induce cellular changes which could explain SCA10 disease pathology and, secondly, to investigate the similarities and differences between these changes and those induced by expression of rCAG and rCUG repeat RNAs. As for the analysis of *Drosophila* expressing expanded rCAG and rCUG repeats, RNA for the rAUUCU analysis was extracted from newly eclosed flies in order to identify early events in disease progression which may provide insight into causative changes.

Microarray analysis was performed on *Drosophila* expressing four transgene insertions of the untagged $rAUUCU_{65}$ construct under the control of the *elavII-GAL4* driver. This experiment is therefore the equivalent of microarray experiment 2 for rCUG, rCAG and rCAA repeats (as described in Chapter 5.2) and therefore all comparisons discussed in this section refer to the data presented for experiment 2. A much larger number of genes were identified as changed in flies expressing $rAUUCU_{65}$ compared to the microarray analysis of *elav>rCAG* and *elav>rCUG* in microarray experiment 2, where only genes commonly altered in two independent four insertion lines were further analysed. An independent four transgene insertion line for $rAUUCU_{65}$ was not able to be generated during this study due to an apparent bias for insertion of this transgene on chromosome 3. 391 genes were identified which showed altered expression in flies expressing rAUUCU RNA compared to both *elav>+* and *elav>rCAA* flies (listed in Appendix C, Table C1). This list should not contain transcripts which are altered as a result of either GAL4 toxicity or effects of rCAA RNA expression and therefore should provide a more robust data set for further analysis.

Analysis of functional information for the genes in this list resulted in categorisation of the altered genes into a large number of ontologies, with the most highly represented categories being “redox regulation” (9.5%), “transcriptional

regulation" (7.2%), "immune response" (8.5%) and "protein modification/ metabolism" (5.9%) (Figure 7.4). The majority of genes with known function detected in this analysis fell into the same categories as those identified in the analyses of *elav>rCAG* and *elav>rCUG* flies in microarray experiment 2 (shown in Figure 5.3), suggesting that expression of rAUUCU RNA alone is sufficient to induce similar cellular changes as expression of rCAG and rCUG repeat RNAs. Since both rCAG and rCUG expanded repeat RNAs have been demonstrated to have dominant toxic effects (117, 270), this result supports the hypothesis that RNA pathology may also play a role in cellular dysfunction in SCA10.

7.4 Investigation of common transcriptional changes in flies expressing rAUUCU, rCAG and rCUG repeats

To investigate common effects of hairpin RNA expression, lists of transcripts which were altered in flies expressing rAUUCU RNA compared to either *elav>rCAA* or *elav>+* flies were generated (log2(ratio)>0.5 or <-0.5). From these lists, genes were then selected which were also altered in the same comparison for either *elav>rCAG* or *elav>rCUG* flies (select transcripts are listed in Figure 7.5 and all transcripts are listed in Tables 7.3-7.4 and Appendix C, Table C2-C5). A large proportion of genes short listed for *elav>rCAG* and *elav>rCUG* compared to either *elav>rCAA* or *elav>+* in microarray experiment 2 were also altered in flies expressing rAUUCU RNA (Table 7.2): the highest concordance between identified genes was 71.4% observed between *elav>rCUG* compared to *elav>+* and the same comparison for *elav>rAUUCU* flies and the lowest concordance was 40.7% observed between *elav>rCAG* compared to *elav>rCAA* and the same comparison for *elav>rAUUCU* flies. Furthermore, of the 9 transcripts commonly altered in flies expressing rCAG and rCUG RNA compared to each of *elav>rCAA* and *elav>+*, 6 genes were also altered in flies expressing rAUUCU RNA compared to *elav>rCAA* and 7 genes for the comparison to *elav>+.* These genes are listed in Table 7.3 and Table 7.4 respectively. This result supports our hypothesis that there are common cellular changes induced by the expression of hairpin-forming expanded repeat RNAs, irrespective of the sequence of the repeat tract.

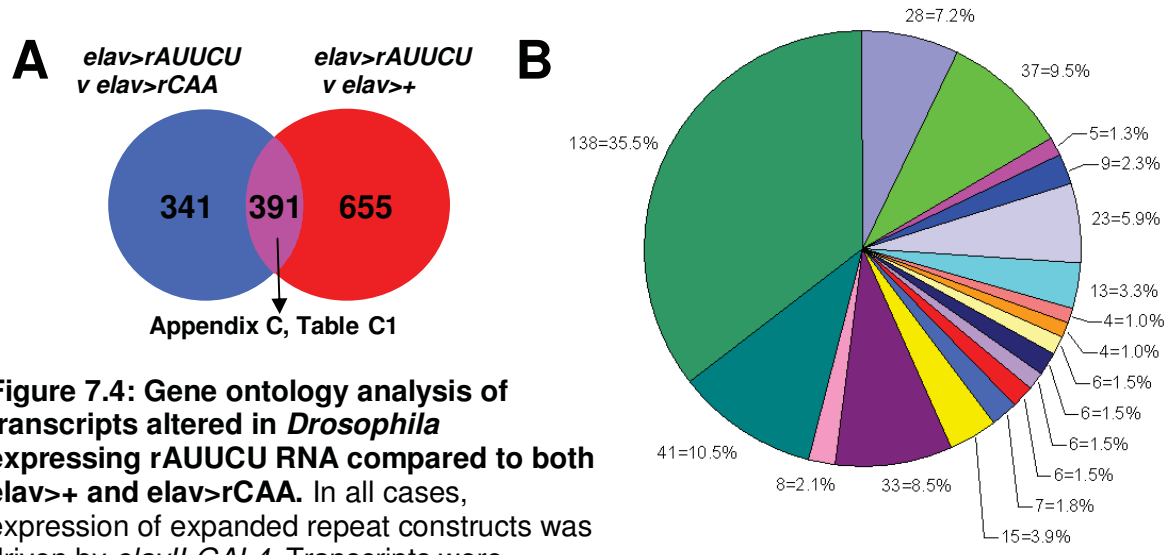


Figure 7.4: Gene ontology analysis of transcripts altered in *Drosophila* expressing rAUUCU RNA compared to both elav>+ and elav>rCAA. In all cases, expression of expanded repeat constructs was driven by *elavII-GAL4*. Transcripts were selected for log2(ratio) >0.5 or <-0.5 when compared to either *elav>+* or *elav>rCAA* flies. **A)** Number of transcripts detected in each comparison, filtered so that a "present" call is achieved for all lines tested. Genes with transcripts detected in both comparisons are listed in Appendix C, Table C1. **B)** Number of genes in each category and the percentage of the total number of genes which they represent. Categories are as listed in figure legend. Genes in the "unknown" category have no known function. Gene ontology was determined either from known phenotypic data or homology with other genes of known function. Genes in the "other" category have a known or suspected function that does not fit into one of the listed categories.

Comparison	Transcripts altered	Percent common to same comparison with elav>rAUUCU
elav>rCAG v elav>rCAA	59	40.7%
elav>rCAG v elav>+	50	52.0%
elav>rCUG v elav>rCAA	172	47.1%
elav>rCUG v elav>+	77	71.4%

Table 7.2: Percent of transcripts commonly altered in *Drosophila* expressing rCAG or rCUG repeats and rAUUCU repeats pan-neuronally. In all cases, expression of expanded repeat constructs was driven by *elavII-GAL4*. Transcripts altered in *elav>rCAG* and *elav>rCUG* flies compared to either *elav>+* or *elav>rCAA* were selected for log2(ratio) >0.5 or <-0.5 (which corresponds to a fold change of approximately ±1.4) with P<0.05. The percent of these transcripts which are also altered when the same comparison is performed for flies expressing rAUUCU repeats are listed.

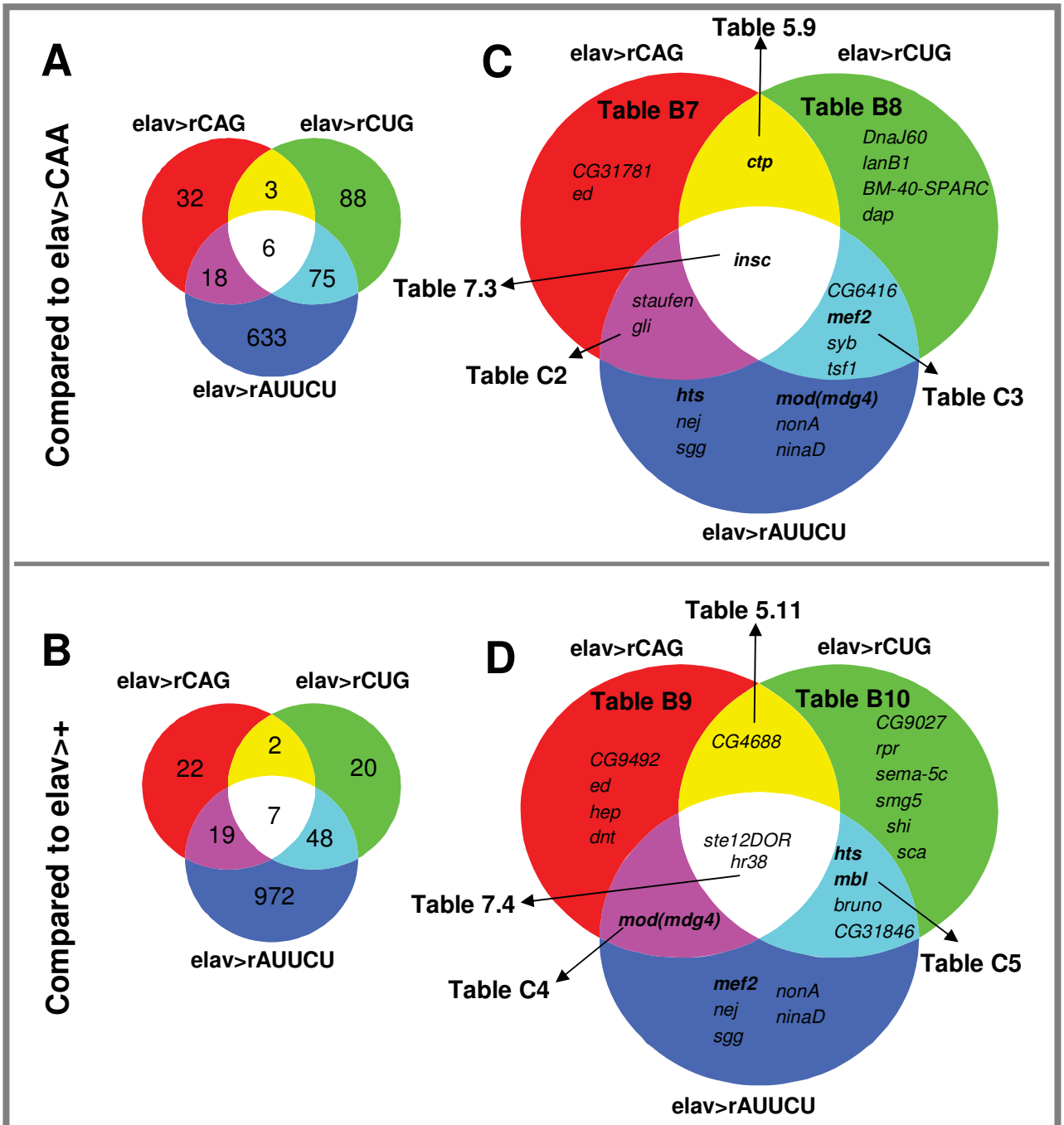


Figure 7.5: Comparison of transcripts altered in *Drosophila* expressing rCAG, rCUG and rAUUCU repeats pan-neuronally. In all cases, expression of expanded repeat constructs was driven by *elavII-GAL4.* Transcripts were selected for Log2(ratio) >0.5 or <-0.5 for each comparison, which represents a fold increase or decrease of ≈1.4. T-tests were performed for genotypes where more than one line was available for analysis and genes filtered for P<0.05. **A)** Number of transcripts altered in *elav>rCAG, elav>rCUG* and *elav>rAUUCU* flies compared to *elav>rCAA.* **B)** Number of transcripts altered in *elav>rCAG, elav>rCUG* and *elav>rAUUCU* flies compared to *elav>+.* **C&D)** Select genes are listed for each comparison. Genes in bold were validated for a role in CAG and CUG repeat pathogenesis by genetic means in Chapter 5. Genes identified in each comparison are listed in tables as indicated.

7.4.1 Common transcriptional changes in *Drosophila* expressing rCAG, rCUG and rAUUCU expanded repeats compared to elav>rCAA

A comparison of transcriptional changes in *Drosophila* expressing rAUUCU repeats and rCAG or rCUG repeats compared to *elav>rCAA* flies resulted in lists of 18 and 75 common transcripts respectively (Listed in Appendix C, Table C2 and C3). Amongst the 18 changes detected which were common to *elav>rCAG* and *elav>rAUUCU* flies compared to *elav>rCAA* (summarised in Figure 7.5), the RNA-binding protein Staufen was found to be downregulated in both genotypes. Staufen has been previously identified as a modifier in a screen of a *Drosophila* SCA8 model (98). This model consists of the non-coding CUG repeat-containing *SCA8* transcript under the control of the UAS-GAL4 system. The observation that expression of Staufen is altered in flies expressing rCAG and rAUUCU hairpin RNAs implicates this RNA binding protein more broadly in expanded repeat pathogenesis.

In *Drosophila*, there is a single Staufen protein which has been associated with RNA transport processes and is important in localisation of transcripts during polarisation of the oocyte (291) and determination of neuroblast asymmetry (292). *Staufen* mutants – along with mutants for several genes encoding RNAs which are normally localised by Staufen – show long-term memory defects, suggesting that localised translation of RNAs is important for memory formation in *Drosophila* (293). There are two Staufen orthologues in mammals, each encoded by a separate gene. While they have similar functions, there is evidence to suggest that they are associated with transport of a unique set of mRNAs. Staufen1 also appears to be fairly ubiquitous, while Staufen2 is specifically expressed in neurons where it is thought to be important in the localisation of RNA to dendrites (193). Staufen2 also plays a role in nuclear RNA export via an interaction with Exportin 5 (294) as well as the nucleoporin Nup62 (193), a protein which was present in the single commonly downregulated spot observed in the proteomics analysis of both *elav>rCAG* and *elav>rCUG* flies compared to *elav>rCAA* flies (described in Chapter 4).

A similar analysis of *elav>rCUG* and *elav>rAUUCU* compared to *elav>rCAA* revealed 75 transcripts commonly altered (summarised in Figure 7.4 and listed in Appendix C, Table C3). These transcripts include the muscle transcription factor *mef2* which was previously demonstrated in this study to be able to modify a CUG-encoded

polyleucine eye phenotype (Chapter 6). Mammalian MEF2 transcription factors have been primarily characterised for a role in transcriptional regulation during muscle development (295) and neuronal survival (296). Recent studies suggest that MEF2 also functions with FMRP in eliminating excitatory synapses in a mouse model (275). The role of FMRP in regulation of synaptic activity is thought to be regulated via alteration of dendritic spine number which is elicited through regulation of transport and translation of a particular sub-set of mRNAs (297). The functional interaction between MEF2 and FMRP implicates MEF2 in RNA localisation and therefore provides a mechanistic relationship via which the expression of hairpin-forming RNAs, which may alter RNA transport dynamics in the cell, might alter MEF2 function.

The cytoskeletal adaptor protein *insc*, was upregulated in *elav>rCAG*, *elav>rCUG* and *elav>rAUUCU* flies compared to *elav>rCAA* (Table 7.3) further supporting the idea that cellular transport is generally disrupted by hairpin repeat expression. Interestingly, a requirement for Insc in Staufen-mediated RNA localisation during neuroblast asymmetrical division has been reported in *Drosophila* (292), supporting a link between alterations to the cytoskeleton and RNA transport processes. The observation that components of RNA transport pathways are altered in flies expressing rCAG, rCUG and rAUUCU expanded repeat RNA suggests that this effect is sequence-independent and may represent a common pathogenic mechanism amongst the expanded repeat diseases.

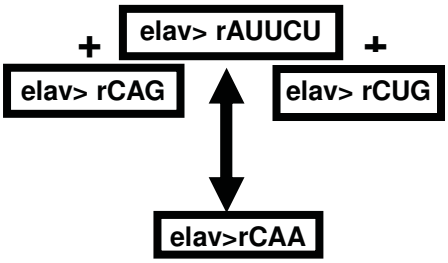


Table 7.3: Changes common to elav>rAUUCU, elav>rCAG and elav>rCUG flies compared to elav>rCAA.

Log2(ratio) >0.5 or <-0.5, selected for P<0.05 for *elav>rCAG* and *elav>rCUG* comparisons.

Gene Title	Gene Symbol	Ensembl	Log2(ratio) elav>rCAG to elav>rCAA	Log2(ratio) elav>rCUG to elav>rCAA	Log2(ratio) elav> rAUUCU to elav>rCAA	Human orthologue	Function
CG9400	CG9400	CG9400	0.84 P=0.021	0.97 P=0.014	0.55		
CG9079	Cpr47Ea	CG9079	0.51 P=0.011	1.34 P=0.003	0.56		
CG34104	CG34104	CG34104	0.58 P=1.43E-5	0.80 P=0.004	0.57		Signal transduction, GTPase activity
Inscuteable	insc	CG11312	0.89 P=0.013	0.55 P=0.050	0.60	INSC	cytoskeletal adaptor, protein and RNA localisation,
CG12998	CG12998	CG12998	0.92 P=0.018	0.71 P=0.023	0.88		
CG14528	CG14528	CG14528	0.72 P=0.024	0.84 P=015	0.91		Metallo-endopeptidase

7.4.2 Common transcriptional changes in *Drosophila* expressing rCAG, rCUG and rAUUCU expanded repeats compared to *elav>+*

A comparison of transcriptional changes in *Drosophila* expressing rAUUCU repeats and rCAG or rCUG repeats compared to *elav>+* flies resulted in lists of 19 and 48 common transcripts respectively (Listed in Appendix C, Table C4 and C5). Several of these transcripts were previously tested in this study for interactions with translated CUG repeats encoding polyleucine or translated CAG and CAA repeats encoding polyglutamine (Chapter 6). One of these, *mod(mdg4)*, was previously identified in a P-element screen for modifiers of a phenotype caused by expression of the human SCA8 non-coding RNA in the *Drosophila* eye and also showed altered expression in both rCAG and rCUG repeat expressing flies in microarray experiment 1. Reducing *mod(mdg4)* expression in flies co-expressing CAG-encoded polyglutamine or CUG-encoded polyleucine resulted in lethality or a strong enhancement in the associated eye phenotypes. Whilst there was also an enhancement to an eye phenotype resulting from co-expression of a CAA-encoded

polyglutamine tract with an RNAi construct targeting mod(mdg4), this interaction did not appear to be as strong (Figure 6.3). The finding that mod(mdg4) transcript levels are also altered in rAUUCU repeat-expressing flies may suggest a broader role for this protein in expanded repeat disease pathogenesis. Since Mod(mdg4) has been shown to play a role in chromatin remodelling and gene silencing, this result may indicate that structural properties of these repeat tracts at the DNA level also play a role in pathogenesis.

Comparing *elav>rCUG* and *elav>rAUUCU* to *elav>+* also revealed a number of interesting transcripts including the cytoskeletal protein and orthologue of ADD1, *hts,* and splicing factor *mbl*. Both Hts and Mbl have been previously implicated in *Drosophila* expanded repeat disease models (98, 117, 121). In our model, reducing expression of *hts* resulted in suppression of phenotypes associated with expression of polyglutamine encoded by either a CAG or CAA repeat or polyleucine encoded by a CUG repeat. This result does not demonstrate a sequence-specific interaction between Hts and expanded repeats at the RNA level, however it is possible that any difference in effect that Hts may have at the RNA level was masked by the interaction with polyglutamine. The observation that *hts* expression is also altered in *Drosophila* expressing rAUUCU repeat RNA supports a role for Hts in RNA toxicity. Altering levels of *mbl* was found to modify eye phenotypes associated with expression of translated CUG or CAG repeats in a sequence-dependent manner, although overexpression of human MBNL1 in *Drosophila* did not show the same sequence dependence. The splicing factor Bruno – an orthologue of CUG-BP1 which is the MBNL1 antagonist implicated in DM1 (86) – was also downregulated in microarray analysis of rAUUCU repeat expressing flies. These results may indicate a broader role for the MBNL/CUG-BP1 pathway in RNA pathogenesis in the expanded repeat diseases.

Seven transcripts were altered in all of *elav>rCAG, elav>rCUG* and *elav>rAUUCU* compared to *elav>+* (Table 7.4), 4 of which have functional information associated with them. Amongst these is *hr38*, a nuclear receptor and orthologue of mammalian *NGFI-B/NUR77* which has been linked to induction of apoptosis via translocation to the nucleus (298) as well as inhibition of dendritic differentiation and synapse formation (250). NUR77 is highly expressed in the striatum and prefrontal cortex (251), regions of the brain which are highly susceptible

to degeneration in Huntington's disease, and has been demonstrated to be regulated by the transcription factors MEF2 and CREB (299), both of which have been implicated in expanded repeat disease (273, 277, 300-302). The CREB binding protein orthologue *nej*, the GSK3-β orthologue *sgg* and *mef2* also show altered expression in flies expressing rAUUCU RNA compared to both *elav>+* and *elav>rCAA* controls (Appendix C, Table C1). GSK3-β has been demonstrated to regulate activity of a large number of transcription factors including CREB (303) and MEF2 (296) and has been linked with numerous human diseases including Fragile X syndrome (304), Alzheimer's disease (305), diabetes (306) and a number of cancers (307). The identification of altered transcription of the *Gsk3-β* orthologue (*sgg*) in flies expressing expanded untranslated repeat tracts may further implicate it in the expanded repeat diseases and therefore a role for GSK3-β signalling in RNA toxicity was further investigated.

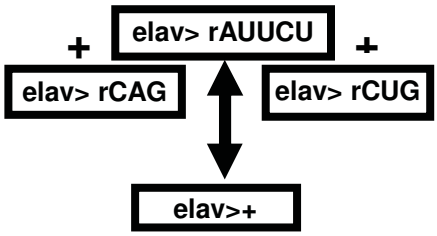


Table 7.4: Changes common to *elav>rAUUCU, elav>rCUG* and *elav>rCAG* flies compared to *elav>+.*

Log2(ratio) >0.5 or <-0.5, selected for P<0.05 for *elav>rCAG* and *elav>rCUG* comparisons.

Gene Title	Gene Symbol	Ensembl	Log2(ratio) elav>rCUG to elav>+	Log2(ratio) elav>rCAG to elav>+	Log2(ratio) elav> rAUUCU to elav>+	Human Orthologue	Function
Stellate orphon	Ste12DOR	CG32616	-2.00 P=0.036	-2.88 P=0.002	-2.33		Spermato-genesis, protein kinase regulator
CG32552	CG32552	CG32552	-0.86 P=0.030	-0.55 P=0.013	-0.87		
CG13077	CG13077	CG13077	-0.79 P=0.014	-0.55 P=0.019	-0.85	CYB561D2	
CG13117	CG13117	CG13117	-0.56 P=0.013	-0.53 P=0.015	-0.84		
CG9686	CG9686	CG9686	0.54 P=0.012	0.55 P=0.014	0.54		
CG9186	CG9186	CG9186	0.95 P=0.005	0.58 P=0.017	0.57	SLC39A6	Metal ion transporter
Hormone receptor-like in 38	Hr38	CG1864	0.55 P=0.028	1.05 P=0.04	0.91	NR4A1/ NGFI-B	Ligand-dependent nuclear receptor activity

7.5 Investigation of a role for the Akt/GSK3- β signalling pathway in expanded repeat disease pathogenesis

Glycogen synthase kinase 3 beta (GSK3-β) was originally identified for its role in regulation of glycogen metabolism, but has more recently been characterised as a central regulator of a number of distinct signalling pathways including the Wnt, insulin and EGF signalling pathways (308). One mechanism by which GSK3-β activity can be regulated is through inhibitory phosphorylation by the protein serine-threonine kinase Akt. This phosphorylation event promotes cell survival and has been demonstrated to be dysregulated in Alzheimer's disease (309). Regulation of signalling of the Akt/GSK3-β pathway is a complex system which is responsive to several different signals including calcium influx and various neurotrophic signals (summarised in Figure 7.6).

7.5.1 Evidence for alterations to Akt/GSK3- β signalling in the expanded repeat diseases

There are several lines of evidence which implicate alterations in Akt activity in expanded repeat disease pathogenesis, including the ability of phosphorylation by Akt to regulate activity of several expanded repeat-containing proteins: Akt is able to phosphorylate Ataxin-1 (126), the androgen receptor (310) and HTT (287) and hence regulate their interactions with other proteins. Upregulation of Akt activity in a mouse model of SBMA has also been demonstrated to alleviate pathology by reducing aggregation of the mutant androgen receptor, an effect which can also be induced through overexpression of IGF-1 (311).

In *Drosophila, akt* has been identified as a common modifier of phenotypes induced by expression of expanded polyglutamine-containing HTT and Ataxin-1 (312). Interestingly, altering Akt levels appears to have opposite outcomes in these two models, with stabilisation of the expanded polyglutamine-containing Ataxin-1 protein increasing toxicity in the SCA1 model and enhanced proteasomal function mitigating toxicity in the HD model. In another study investigating polyglutamine-specific changes in *Drosophila* and cell lines, Nelson *et al.* (2005) identified the target of rapamycin (TOR) pathway, which is regulated by Akt activity, as one component that was consistently deregulated (244). Hernandez-Hernandez et al. (2006) also report

disruption of Akt/GSK3-β signalling in PC12 cells expressing expanded CUG repeats which they suggest is mediated through NGF signalling (313). These observations suggest that Akt/GSK3-β signalling may be disrupted by several mechanisms in expanded repeat disease: firstly, by the aggregation of expanded polyglutamine tracts, but also by alterations to the physical shape of the polyglutamine tract-containing proteins themselves, some of which are targets of Akt. Finally, the expression of hairpin-forming RNA species such as CUG repeats may also perturb signalling. This pathway may therefore represent a key effector of neurodegeneration in the expanded repeat diseases.

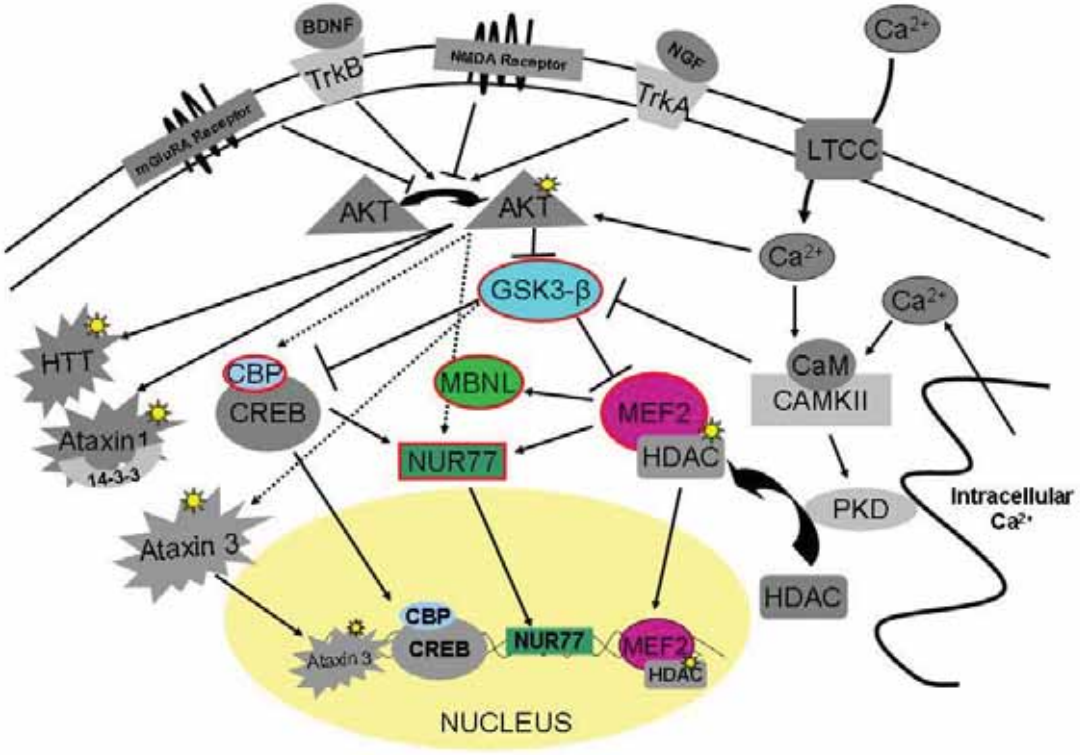


Figure 7.6: Alteration to activity of the Akt/GSK3-β signalling pathway can explain a number of the changes observed in microarray analysis of flies expressing rCAG, rCUG and rAUUCU repeats in the nervous system. Coloured shapes indicate genes which showed altered transcript levels in microarray analysis in flies expressing at least one of the untranslated repeat constructs. A number of links between Akt activity and expanded repeat-containing proteins themselves have also been demonstrated. Akt phosphorylates HTT and Ataxin-1 (represented as a star), altering their interactions with other proteins (126, 287, 314). Phosphorylation of ataxin-3 by GSK3-β (star) has also been recently demonstrated to regulate nuclear entry and therefore may play a role in SCA3 (315). Expression of expanded CUG repeats has also been demonstrated to alter activation of the Akt/Gsk3-β pathway (313). Activation of Akt can be regulated by a number of different signals, including glutamate (316) or neurotrophic (317-318) signals and Ca^{2+} signalling (319). Activated Akt is in turn involved in downregulation of GSK3-β activity which is involved in regulation of a number of transcription factors, including MEF2 (296) and CREB (303). Both CREB and MEF2 have been demonstrated to play a role in regulation of expression of the nuclear receptor NUR77, an orthologue of *Drosophila* Hr38, in a calcium-dependent manner (299). Activation of NUR77 can also be regulated directly by Akt (320). The Akt/GSK3-β signalling pathway is therefore able to have broad downstream transcriptional effects.

Altered transcription of components of the Akt/GSK3-β regulatory pathway was consistently observed in rCAG, rCUG and rAUUCU repeat-expressing flies by microarray analysis, suggesting that this is a key component of cellular dysfunction in our *Drosophila* model of untranslated repeat disease pathogenesis. Transcripts which showed altered regulation in the microarray experiments are shown in colour in Figure 7.6. Importantly, one of the downstream effectors of this pathway, hr38, was consistently upregulated, irrespective of the sequence of the repeat being expressed and in comparisons to both the *elav>rCAA* and *elav>+* control lines. Downregulation of the *Drosophila* GSK3-β orthologue sgg, as well as the downstream targets mef2, mbl and the CREB binding protein orthologue nej was also observed in some cases. While the ability of CUG repeat RNA to disrupt Akt/GSK3-β signalling has been described, this is the first evidence that expression of other hairpin-forming RNA species can also influence activity of this pathway.

The initial stimulus resulting in the disruption of Akt/GSK3-β signalling in our model is unclear, however there is precedent for similar effects in Fragile X syndrome where increased levels of stimulation of the mGluR5 receptor have been demonstrated to increase GSK3-β activity (321). A disruption to mGluR5 signalling has also been described in a pre-symptomatic model of HD (322), and in other HD models alterations to N-Methyl-D-Aspartate Receptor (NMDAR) (323), brain-derived neurotrophic factor (BDNF) (264, 324) and nerve growth factor (NGF) (325) signalling, all of which are associated with activation of the Akt/GSK3-β pathway, have also been observed. Our observations indicate that expression of expanded repeat RNA alone is sufficient to cause transcriptional changes to the Akt/GSK3-β pathway, and therefore that the hairpin RNAs expressed in the disease situation might also interact with components of this pathway to disrupt normal signalling.

The mutation-containing gene in Fragile X syndrome, FMRP, is itself an RNA binding protein involved in translational regulation and transport of numerous RNAs through formation of mRNA/protein complexes. A decrease in FMRP levels perturbs neuronal function via dysregulation or mis-localisation of a subset of mRNAs (297). This kind of dendritic localisation and translation of specific mRNAs is a repeated theme in neuronal signalling: BDNF signalling has also been demonstrated to be important for regulation of GSK3-β activity and neuron survival (326) and BDNF is itself transported and locally translated within neurons (reviewed in (327)).Therefore

expression of hairpin-forming RNAs may be detrimental to neuronal function because it causes mis-localisation of other RNA species, possibly through sequestration of RNA binding proteins involved in transport and processing, and therefore disrupts signalling pathways. Since both the RNA binding proteins themselves and the RNA species which they regulate are likely to be specific to certain neuronal sub-types, this also offers some explanation of the observation that some cells are more vulnerable to degeneration than others.

7.5.2 Effect of altering expression of Akt and GSK3- β in our *Drosophila* model of expanded repeat disease pathogenesis

The *Drosophila* orthologue of GSK3-β, Shaggy (Sgg), has been shown to be concentrated at motor-neuron terminals in larvae, where it is involved in the regulation of dynamics of the cytoskeleton. Mutations in *sgg* are associated with neuromuscular junction over-growth phenotypes (328). A role for Sgg in maintenance of olfactory neurons has also been demonstrated, with loss of activity associated with adult degeneration despite normal development (329). Phosphorylation of Sgg by the *Drosophila* Akt kinase (Akt1) is involved in regulation of Sgg activity, suggesting that these pathways are largely conserved in *Drosophila* (330).

In order to investigate a role for the Akt/GSK3-β pathway in RNA toxicity, the effect of altering expression of *sgg* and *akt1* in flies expressing translated expanded repeats was initially tested. It is predicted that modification of phenotypes associated with expression of polyglutamine encoded by a CAG repeat and polyleucine encoded by a CUG repeat, which are both able to form hairpin secondary structures at the RNA level, but not polyglutamine encoded by a CAA repeat, which cannot form a secondary structure, may be indicative of a specific role for this pathway in RNA-mediated toxicity. RNAi lines targeting akt1 and sgg and an overexpression construct for *sgg* were obtained and expressed in the eye with *GMR-GAL4*. Reducing expression of *sgg* did not result in a disruption to the external appearance of the eye (Figure 7.7 D). Co-expression of an RNAi construct targeting sgg with polyglutamine encoded by either a CAG or CAA repeat tract did not significantly alter the appearance of the eye, although a slight improvement in the ordered arrangement of the ommatidial arrays could be seen (Figure 7.7 E&F compared to

A&B). However, reducing expression of *sgg* in flies expressing polyleucine resulted in a reduction in the severity of the eye phenotype, consisting of a complete suppression of the rough eye phenotype and a darkening of the colour of the eye (Figure 7.7 C compared to G).

Since overexpression of Sgg alone resulted in a severe disruption to the eye consisting of a decrease in size, marked lightening of colour and disorganisation of the patterning of the ommatidial arrays (Figure 7.7 H), it was more difficult to interpret the results of co-expression of the translated repeats in this case. In flies ectopically expressing Sgg, co-expression of polyglutamine encoded by either CAG or CAA repeats resulted in a significant change to the eye phenotype consisting of a significant enlargement of the size of the eye with the appearance of necrotic patches and an enhancement of the loss of pigment phenotype (Figure 7.7 I&J compared to A&B). It is not clear whether this change in phenotype indicates a genetic interaction, or the additive outcome of the polyglutamine and Sgg overexpression phenotypes. Co-expression of polyleucine and ectopic Sgg resulted in complete lethality which supports the conclusion that there is a stronger effect of altering Sgg expression levels in polyleucine-expressing flies than polyglutamine flies.

While expression of an RNAi construct targeting akt1 did not disrupt the *Drosophila* eye (Figure 7.7 L), reducing *akt1* expression resulted in an increase in the loss of pigment in the eyes of both CAG and CAA-encoded polyglutamine expressing flies (Figure 7.7 M&N compared to A&B). Co-expression of polyleucine with the RNAi construct targeting akt1 also resulted in a loss of pigment phenotype, with a complete suppression of the roughness seen when polyleucine is expressed alone in the eye (Figure 7.7 O compared to C). It is not clear whether this constitutes an increase in the severity of the eye phenotype or an alteration to the pathogenic pathway, however it does suggest a role for Akt1 in polyleucine pathogenesis. The effects of altering expression of *akt1* and *sgg* in this assay are consistent with what is known about the mechanism of regulation of signalling through GSK3-β since a decrease in Akt1 expression is expected to result in alleviation of Sgg activity inhibition and therefore should give a similar effect to overexpression of Sgg. However, since both CAA and CAG-encoded polyglutamine expressing flies showed similar alterations to the appearance of the eye when Akt1 expression was

decreased or Sgg was ectopically expressed, it is unclear whether hairpin-forming expanded repeat RNA plays a role in these interactions or if the effect is mediated through an interaction with the polyglutamine peptide.

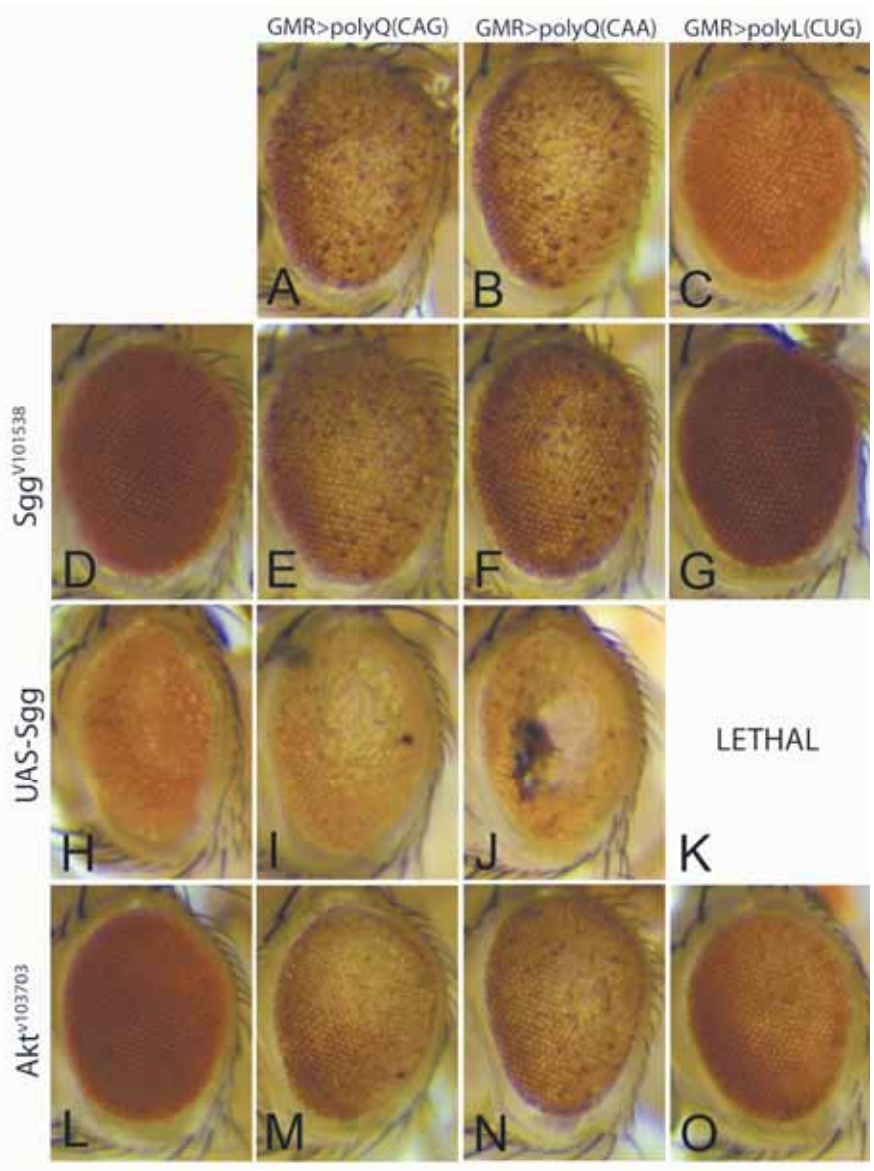


Figure 7.7: Investigation of a role for the Akt/GSK3-β signalling pathway in pathogenesis in polyleucine and polyglutamine-expressing *Drosophila*. In all cases, expression of constructs is driven by *GMR-GAL4*. **A&B)** Expression of polyglutamine encoded by either a CAG or CAA repeat results in an indistinguishable loss of pigment eye phenotype. **C)** Expression of polyleucine encoded by a CUG repeat results in disorganisation of patterning of the ommatidia. **D)** Expression of an RNAi construct targeting transcripts of the *Drosophila* orthologue of *gsk3-β* (*sgg*) does not significantly alter the appearance of the eye. **E&F)** Reducing expression of Sgg in flies expressing polyglutamine encoded by either CAG or CAA repeats does not significantly alter the appearance of the eye. **G)** Reducing expression of Sgg in flies expressing polyleucine completely suppresses the rough eye phenotype and results in darkening of the eye. **H)** Ectopic expression of Sgg in the eye results in a severe rough eye phenotype with a dramatic reduction in the size of the eye and the amount of pigmentation. **I&J)** Co-expression of polyglutamine encoded by either a CAG or CAA repeat with ectopically expressed Sgg results in an increase in the size of the eye compared to ectopic expression of Sgg alone. There appears to be a reduction in the amount of pigment in the eye and in most cases there are necrotic patches and nearly complete loss of the ommatidial array structure. **K)** Ectopic expression of Sgg in flies expressing polyleucine is completely lethal. **L)** Expression of an RNAi construct targeting akt1 does not significantly alter the appearance of the eye. **M&N)** Reducing expression of Akt1 in flies expressing polyglutamine encoded by either a CAG or CAA repeat results in an increase in the loss of pigment phenotype. **O)** Reducing levels of Akt1 in flies expressing polyleucine suppresses the ommatidial disorganisation seen in flies expressing polyleucine alone, but causes a loss of pigment phenotype.

In order to further examine the ability of expanded repeat RNA to disrupt signalling through GSK3-β, the ability of co-expression of rCAG, rCUG, rAUUCU and rCAA repeats to alter the phenotype associated with overexpression of Sgg in the *Drosophila* eye was tested. This phenotype consists of a strong disruption to patterning of the eye, a decrease in overall size and a loss of colour (Figure 7.8 F,F'). Co-expression of four transgene insertions of the UAS construct with no insert in flies overexpressing Sgg resulted in a suppression of all components of this phenotype, probably as a result of titration of GAL4 by the UAS sites and therefore lower expression of ectopic Sgg (Figure 7.8 G,G'). While co-expression of four transgene insertions of the rCAA construct did not reduce the loss of colour phenotype, a slight increase in the size of the eye and a significant suppression of the rough appearance of the surface of the eye were seen (Figure 7.8 I,I'). Again, these are likely to be effects of titration of GAL4 in the cells of the eye.

Flies co-expressing rCAG and rAUUCU repeat RNAs with the Sgg overexpression construct had rougher eyes than flies co-expressing either the UAS construct of rCAA repeat RNA (Figure 7.8 H,H' & K,K'), which may indicate that expression of RNA which is able to form a hairpin structure genetically interacts with *sgg* in *Drosophila*. This effect was stronger in rCAG repeat expressing flies, possibly as a result of the greater stability of the secondary structure formed by this repeat sequence compared to rAUUCU repeats. In support of a hairpin-dependent interaction between *sgg* and expanded repeat RNA, co-expression of rCUG repeat RNA with the Sgg overexpression construct resulted in complete lethality. This result is also consistent with the lethality observed when Sgg was overexpressed in polyleucine expressing flies, which may indicate that the interaction observed with polyleucine is also mediated by expanded repeat RNA and not the polyleucine peptide. Expression of hairpin-forming RNA species therefore seems sufficient to perturb signalling through GSK3-β in our *Drosophila* model, although this effect is significantly stronger in flies expressing CUG repeats. GSK3-β signalling may therefore represent a common pathogenic pathway in the expanded repeat diseases.

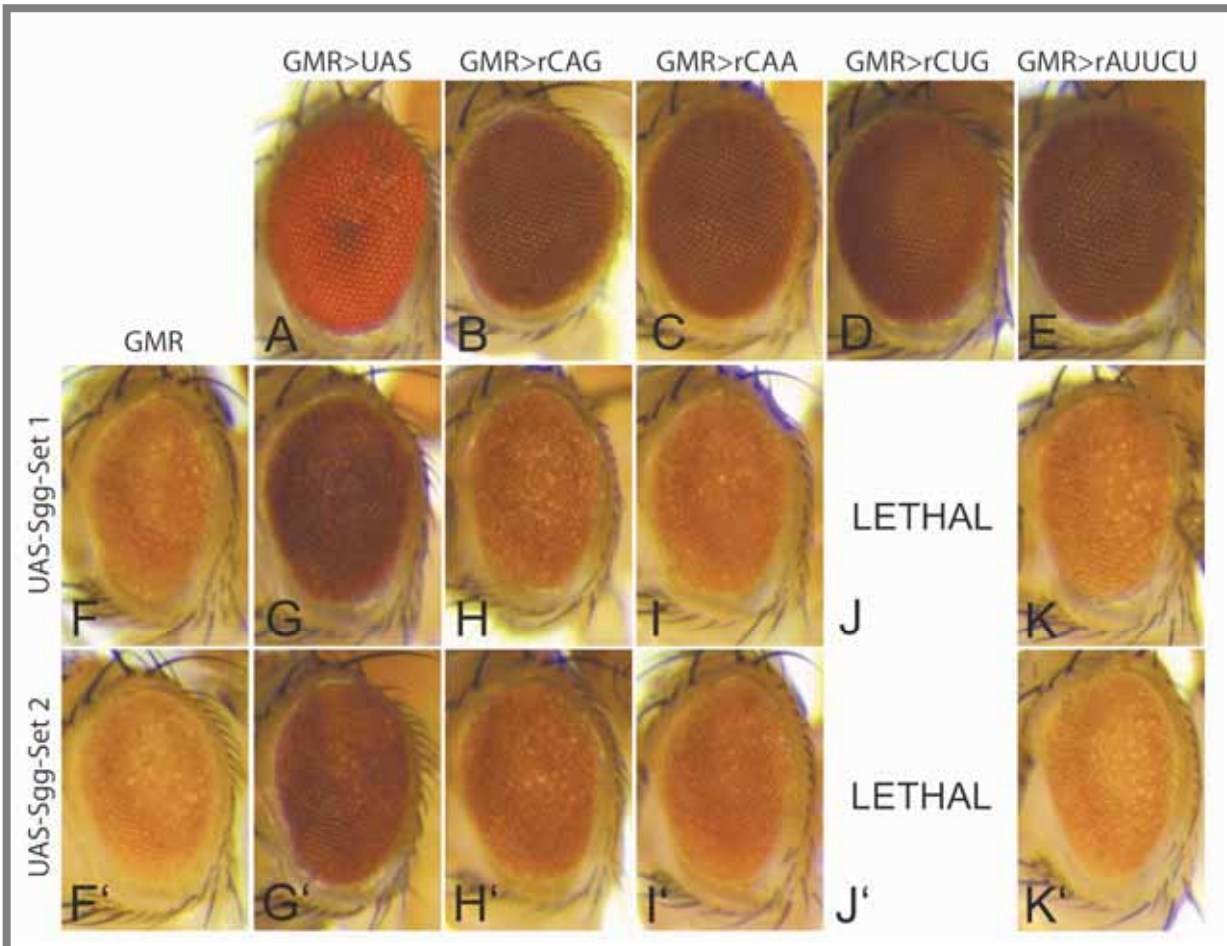


Figure 7.8: Investigation of a role for the Akt/GSK3-β signalling pathway in RNA-mediated pathogenesis in *Drosophila*. In all cases, expression of repeats and the UAS-Sgg construct are driven by *GMR-GAL4.* Set 1 and 2 refer to progeny from independent crosses. **A-E)** Expression of four transgene insertions of the UAS construct with no insert (UAS) or containing the rCAG, rCAA, rCUG or rAUUCU repeats does not alter the external appearance of the eye. **F, F')** Overexpression of Sgg in the eye results in a strong disruption to patterning, a decrease in overall size and a loss of colour. **G, G')** Co-expression of four transgene insertions of the UAS construct with the Sgg overexpression construct dramatically decreases the loss of colour and improves patterning in the eye. **I, I')** Co-expression of four transgene insertions of the rCAA construct with the Sgg overexpression construct improves the patterning of the eye and increases the overall size of the eye but does not alter the colour compared to expression of Sgg alone. **H, H' & K, K')** Co-expression of the rCAG or rAUUCU repeat constructs with the Sgg overexpression construct results in eyes which are consistently rougher than those of flies co-expressing either rCAA or the UAS construct with Sgg. **J, J')** Co-expression of rCUG with the Sgg overexpression construct results in lethality.

7.6 Validation of an interaction between rAUUCU RNA and *mod(mdg4)*, *mbl* and *mef2* in *Drosophila*

A number of other candidate genes which showed altered expression in flies expressing rCAG and rCUG repeats pan-neuronally were also altered in rAUUCU repeat expressing flies. In order to further investigate the role of these

candidate genes in SCA10 pathogenesis, the ability of altering expression of *mbl*, *mef2* and *mod(mdg4)* to induce a phenotype in flies expressing the rAUUCU repeat RNA was also tested. Since no phenotype was elicited by expression of rAUUCU repeat RNA alone, only expression changes which increase the toxicity of this repeat RNA can be identified in this case and therefore candidates were chosen on the basis that they were previously demonstrated to enhance toxicity in translated CUG and CAG repeat-expressing flies. For two of the candidates, *mef2* and *mod(mdg4)*, reducing expression was also demonstrated to enhance toxicity of untranslated CUG repeats, further supporting a role for these genes in RNA toxicity. In rat cerebellar neurons, MEF2D activity has been shown to be regulated by GSK3-β signalling (296) while in *Drosophila*, regulation of *mbl* by Mef2 has been demonstrated (273). These observations suggest a mechanism by which expression of *mef2* and *mbl* might be altered in response to expression of expanded repeat RNA.

Expression of RNAi constructs targeting mef2 and mod(mdg4) has previously been demonstrated to have no effect on the appearance of the *Drosophila* eye (Figure 6.5 E&I). Co-expression of these RNAi constructs with four transgene insertions of the $rAUUCU_{65}$ construct driven by *GMR-GAL4* did not result in a disruption in the external patterning of the eye (Figure 7.9 B&C). Similarly, expression of four transgene insertions of the $rAUUCU_{65}$ construct in the eye of flies heterozygous for the mbl^{E27} loss of function allele did not result in any change in the appearance of the eye. This result, along with the inability of alterations in these candidate genes to elicit a phenotype in flies expressing rCAG repeat RNA, may indicate that there are real differences in the degree or mechanism of toxicity of rCUG repeat RNA in this model. However, since expression of rAUUCU repeats was able to induce similar transcriptional changes to those induced in rCUG repeat-expressing flies, a difference in the degree of toxicity and not the actual mechanism seems to be a more likely explanation. It is unclear whether this difference is a feature of greater tolerance for some repeat sequences in *Drosophila* or a more general property of the repeat sequences themselves, which may therefore be relevant to disease pathology.

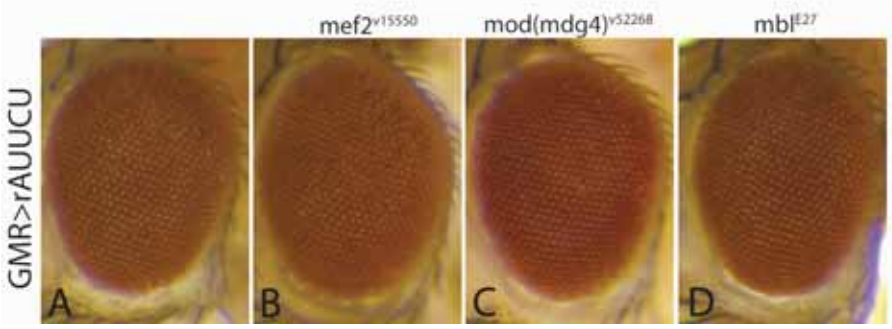


Figure 7.9: Genetic validation of candidates from microarray analysis of *Drosophila* expressing rAUUCU RNA. In all cases, rAUUCU repeat RNA and RNAi construct expression was driven by *GMR-GAL4*. **A)** Expression of four transgene insertions of the $rAUUCU_{65}$ construct does not cause any disruption to the exterior appearance of the eye. **B-C)** Co-expression of RNAi constructs targeting *mef2* or *mod(mdg4)* with four transgene insertions of the $rAUUCU_{65}$ construct does not cause any disruption to the exterior patterning of the eye. **D)** Expression of four transgene insertions of the $rAUUCU_{65}$ construct in flies heterozygous for mbl^{E27} also does not alter the appearance of the eye.

7.7 Further investigation of a role for MBNL1 in expanded repeat pathogenesis

Since MBNL has been demonstrated to co-localise with both expanded CAG and CUG repeat-containing transcripts in a number of disease models, a central role for this splicing factor in RNA pathogenesis has been suggested. While altering expression of the *Drosophila* orthologue of MBNL, Mbl, was not able to elicit a phenotype in flies expressing the expanded rAUUCU repeat construct, it is possible that there are differences in the binding capacity of the *Drosophila* and human proteins and therefore the ability of rAUUCU repeat RNA to interact with MBNL1 in this model was also investigated. The effect of expression of both the $rAUUCU_{65}$ and $rAUUCU_{100+}$-GFP constructs in the eye on the phenotype observed when human MBNL1 is expressed alone was tested.

Flies co-expressing MBNL1 with four transgene insertions of GFP alone (GFP) or four transgene insertions of the UAS construct alone (UAS) were used as controls for these crosses. Co-expression of the UAS control takes into account the contribution of GAL4 toxicity to the MBNL1 eye phenotype, since the presence of four transgene insertions of the UAS construct should reduce the amount of free GAL4 in the cells of the eye in the same manner as the four transgene insertions of the repeat constructs (Figure 7.10 B). Similarly, co-expression of four transgene insertions of the GFP construct takes into account the contribution of GAL4 toxicity but also the effect of GFP expression on the eye phenotype and therefore is the appropriate control for flies expressing the GFP tagged $rAUUCU_{100+}$-GFP construct.

Flies co-expressing either UAS or GFP with MBNL1 showed an increase in the size of the eye compared to those expressing MBNL1 alone (Figure 7.10 A compared to B and D), however eyes of GFP-expressing flies appeared less rough and had lighter colour than eyes of UAS-expressing flies (Figure 7.10 C) which suggests that there is some difference in the effect of expression of these constructs.

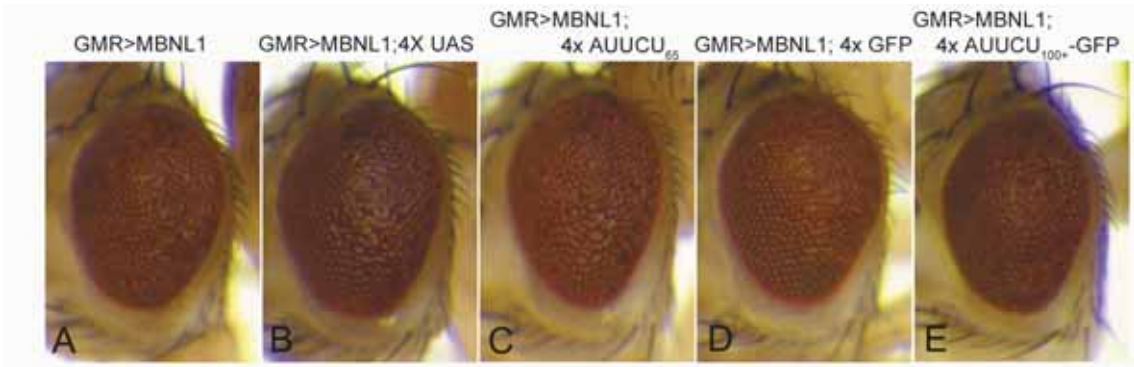


Figure 7.10: Interaction of rAUUCU repeats with MBNL1 in the *Drosophila* eye. All constructs are driven in the eye by *GMR-GAL4* and flies were grown at 23 ºC. **A)** Expression of MBNL1 in the eye results in a rough eye phenotype with necrotic patches. **B)** Expression of MBNL1 and the UAS control results in an increase in the size of the eye compared to expression of MBNL1 alone. **C)** Co-expression of rAUUCU$_{65}$ with MBNL1 results in a slight decrease in the size of the eye compared to co-expression with the UAS control. However, since the eye appears lighter in colour and has fewer necrotic patches, it is unclear whether this demonstrates an enhancement. **D)** Expression of MBNL1 and GFP alone results in a mild rough appearance of the eye with no necrotic patches evident. **E)** Co-expression of rAUUCU$_{100+}$-GFP with MBNL1 results in a slight enhancement compared to the eye phenotypes seen when MBNL1 is co-expressed with GFP alone. Eyes appear smaller and slightly darker in colour, with more necrotic patches evident.

Expression of the un-tagged rAUUCU$_{65}$ construct resulted in a slight decrease in the size of the eye compared to flies co-expressing MBNL1 and UAS, however the eye also appeared to have a smaller area of roughness and necrosis and therefore it is unclear whether this indicates an interaction between MBNL1 and rAUUCU RNA (Figure 7.10 B compared to D). A mild enhancement, consisting of an increase in the appearance of necrotic patches and a decrease in the size of the eye, was observed when the effect of co-expression of rAUUCU$_{100+}$-GFP RNA with MBNL1 was compared to co-expression of MBNL1 with GFP alone (Figure 7.10 C compared to E) suggesting that rAUUCU-GFP RNA may interact with MBNL1, however this was not a strong effect. These results demonstrate inconsistency in the ability of rAUUCU repeat-containing RNA to interact with MBNL1 which appears to depend upon the context of the repeat tract. This effect could be the result of differences in the stability or localisation of each of the repeat-containing transcripts.

7.8 Summary of *Drosophila* model for SCA10

This is the first study to model toxicity of the rAUUCU repeat RNA associated with SCA10. Using expression of this repeat tract in *Drosophila*, the ability of rAUUCU repeat-containing RNA to form foci reminiscent of those observed in models of DM1 *in vivo* was demonstrated in a subset of *Drosophila* cells. This result supports the idea that the rAUUCU repeat RNA may have a dominant toxic effect in the cell and may induce pathology through a similar mechanism to other expanded repeat RNAs.

Analysis of the transcriptional changes resulting from expression of rAUUCU repeats in *Drosophila* neurons further supported the hypothesis that this expanded repeat RNA alone may be sufficient to induce cellular changes which could result in neurodegeneration over an extended period of time. The degree of concordance in genes altered by expression of rAUUCU RNA with those previously detected as altered in rCAG and rCUG repeat-expressing flies, as well as the identification of a number of genes already associated with repeat pathology in different models, strongly supports our hypothesis that expression of expanded repeat RNA alone is sufficient to induce cellular changes consistent with pathology in a sequence-independent manner. This finding is consistent with a model where expression of hairpin-forming RNA may sequester RNA binding proteins and thus cause mis-localisation and subsequent dysregulation of RNA species, although this is likely to be only one component of disease pathology.

Analysis of the common changes identified in flies expressing rCUG, rCAG and rAUUCU repeat RNA identified *hr38* which is a down-stream effector of the Akt/GSK3-β signalling pathway. This observation suggests that this pathway is likely to be important in disease pathology and therefore that components of this pathway may be useful therapeutic targets. An interaction was observed between the *Drosophila GSK3-β* orthologue, *sgg*, and rCAG, rCUG and rAUUCU RNAs suggesting that the ability of the RNA to form a secondary structure, but not the sequence of the repeat itself, is important for an interaction. This result supports a role for signalling through the Akt/GSK3-β pathway in pathogenesis of both translated and untranslated repeat diseases.

Chapter 8: Discussion

8.1 Summary of results

The principle aim of this study was to investigate the cellular pathways involved in toxicity of expanded repeat RNA. Recent evidence (117) suggests a common role for RNA-mediated toxicity in both the translated and untranslated repeat diseases and therefore understanding the mechanism by which the expansion of repeat tracts can be pathogenic at the RNA level is vital to the development of effective therapies for these diseases. Initial experiments in this study tested the intrinsic pathogenicity of a CUG repeat tract, which is the repeat involved in DM1, SCA8 and HDL-2, and a CAG repeat tract, which is the repeat associated with all of the polyglutamine diseases as well as the untranslated repeat disease, SCA12.

The focus of this initial investigation was the elucidation of common pathways of pathogenesis in untranslated rCUG and rCAG repeat expressing flies, since these repeat RNAs form strikingly similar secondary structures and therefore may perturb cellular function in a similar manner. Expression of an expanded CAA repeat RNA was used as a control, as it is unable to form this type of structure. Early alterations to cellular homeostasis were identified in newly eclosed *Drosophila* expressing each of these repeat sequences pan-neuronally by both microarray and proteomic analysis. Applying this method should allow the identification of the primary outcomes of neuronal hairpin RNA expression, which may be the cause of neural dysfunction and death in the disease situation. A role for candidates identified in this manner was then genetically verified by altering expression in the eye of flies expressing hairpin-forming expanded repeat RNA.

Since expression of untranslated CAG and CUG RNA repeats does not cause a neuronal phenotype or a disruption to the organisation of the eye in this *Drosophila* model, initial screening of these candidates tested their ability to modify phenotypes caused by expression of translated repeat sequences. This type of approach has been previously used to uncover components of RNA pathogenesis in a *Drosophila* model of SCA3 pathogenesis (117). Both translated CAG and CAA repeats encode polyglutamine and expression of either repeat in the *Drosophila* eye as part of an open reading frame causes a severe disruption to patterning of the eye. In the case

of the lines used in this study, this disruption mainly consists of a loss of pigment eye phenotype. Expression of translated CUG repeat RNA, which encodes a polyleucine tract, causes a distinct phenotype mainly consisting of a roughening of the surface of the eye. Several candidates identified by microarray or proteomic analyses of flies expressing untranslated hairpin (rCAG and rCUG) repeats also showed some genetic interaction in translated CAG and CUG repeat-expressing flies, supporting a functional role in expanded repeat pathogenesis in this model. Amongst these candidates were genes and proteins involved in functions including nuclear transport, chromatin modification, splicing and transcriptional regulation. A select group of these candidates was then tested for the ability to induce a phenotype in flies expressing expanded untranslated repeats in the *Drosophila* eye. Expression of RNAi constructs targeting two of these candidates, *mod(mdg4)* and *mef2*, or expression of untranslated rCUG RNA alone does not elicit a phenotype in the *Drosophila* eye. However, co-expression of untranslated rCUG RNA and either of these RNAi constructs in the eye induces a significant disruption to the eye, suggesting that *mod(mdg4)* and *mef2* represent rate-limiting steps in RNA pathogenesis in this *Drosophila* model.

Subsequent generation of a *Drosophila* model of SCA10 pathogenesis allowed investigation of the cellular outcomes of expression of an expanded repeat tract with distinctly different sequence composition; in this case a pentanucleotide AUUCU repeat. Microarray analysis was also performed on the *Drosophila* SCA10 model to determine whether similar cellular perturbation is induced by expression of other hairpin-forming disease-associated sequences. A large amount of concordance with the transcriptional changes observed when untranslated rCAG and rCUG RNAs were expressed was seen in flies expressing the untranslated rAUUCU RNA, consistent with a common RNA hairpin-mediated pathogenic mechanism in this *Drosophila* model. Performing comparisons with the different repeat sequences also highlighted common transcriptional changes in a number of down-stream effectors of the Akt/GSK-3β signalling pathway in flies expressing rCAG, rCUG or rAUUCU repeat RNAs. An interaction between the *Drosophila* orthologue of GSK-3β (Sgg) and untranslated rCAG, rCUG and rAUUCU repeats was observed in the *Drosophila* eye, further supporting a role for this pathway in pathogenesis. Using this *Drosophila* SCA10 model, the ability of this pentanucleotide repeat sequence to form RNA foci in a similar manner to those reported for CAG and CUG repeat RNAs was also

demonstrated. It was therefore concluded that formation of RNA foci, which may be regions of high concentration of RNA binding proteins, may also be one outcome of repeat expansion in the SCA10 transcript. The pathogenic potential of RNA foci in expanded repeat disease requires further investigation.

One mechanism by which expanded repeat RNA has been proposed to be toxic is through the sequestration of RNA binding proteins which can result in a loss or reduction of the normal activity of the protein. In DM1, the splicing factor MBNL has been implicated as playing a major role in pathogenesis through a reduction in splicing activity resulting from sequestration by CUG repeat RNA. A role for altered activity of the RNA editing protein Adar was investigated in *Drosophila* expressing CAG repeat transcripts, since a reduction in editing of the normal targets of this enzyme is known to have dramatic neurological effects. No role for the *Drosophila* editing enzyme Adar in either polyglutamine pathogenesis or pathogenesis of untranslated CAG repeat RNA was observed in this model, however this does not rule out a role for the human enzyme in disease pathogenesis.

8.2 Implications for expanded repeat disease pathogenesis

The results presented in this study have demonstrated the ability of expression of hairpin RNA alone to perturb cellular homeostasis and, in conjunction with other components, to act as a cellular toxin. Given the degree of concordance in the pathways disrupted by expression of different repeat sequences, this study also suggests a sequence-independent element to pathogenesis in this *Drosophila* model of expanded repeat disease. Nevertheless, since expression of untranslated expanded repeat tracts in *Drosophila* is not sufficient to induce degeneration within the life-time of the fly – while expression of polyglutamine results in severe early degeneration irrespective of whether it is encoded by a hairpin-forming CAG repeat or a non-hairpin-forming CAA repeat – the extent to which RNA-mediated toxicity contributes to phenotypes in polyglutamine-expressing *Drosophila* remains unclear. Evidence from other *Drosophila* models supports a requirement for high levels of hairpin repeat RNA expression to induce degeneration (117), while expression of polyglutamine peptide is consistently highly toxic from early in development (61,

109). It is unclear whether this high tolerance for expression of hairpin-forming RNAs is a unique feature of *Drosophila*.

Since the expanded repeat diseases generally involve late onset degeneration, the high level of toxicity induced by expression of polyglutamine in this model does not appear consistent with the slow progression seen in the polyglutamine diseases. It has been demonstrated that introduction of amino acids outside of the polyglutamine tract can mitigate toxicity in a *Drosophila* model (109), suggesting that context plays a major role in determining toxicity in the polyglutamine diseases. Furthermore, evidence from other *Drosophila* models of untranslated repeat disease pathogenesis supports a similar role for regions outside of the expanded repeat tract in determining toxicity of expanded repeat RNA (86, 98). It is therefore possible that the polyglutamine diseases demonstrate a greater contribution of RNA-mediated pathogenesis and a lesser contribution of polyglutamine pathogenesis than what is seen in this *Drosophila* model. In the case of the untranslated expanded repeat diseases where there is no toxic peptide expressed, RNA-mediated pathogenesis is presumably sufficient to induce all of the cellular changes leading to neurodegeneration. It is likely that the sorts of changes observed in this model represent components of pathogenesis in these diseases, but that there are also specific effects of expression of the repeat-containing transcript in each disease which are dependent on the context of the repeat tract. Nevertheless, several candidates which showed strong interactions with the context-independent repeats used in this study, including Mod(mdg4) and Mbl, have been previously identified in other *Drosophila* models which used repeats within the disease context (98, 117), suggesting that sequence-independent toxicity does play a role.

8.3 Limitations of the *Drosophila* model

In this study, microarray and proteomic analyses were performed to identify early transcriptional and protein changes which should represent hallmarks of RNA-mediated pathogenesis in *Drosophila*. While each of these techniques is useful in identifying global cellular changes resulting from a particular treatment, they are not comprehensive identification methods. Microarray analysis is limited not only by the number of transcripts represented on the chip, but also by the detection threshold set for analysis which, for the purposes of this study, was deliberately set at a very

stringent level in order to produce a robust data set. Proteomic analyses are largely limited by the abundance of proteins and their ability to be properly resolved on the gel and therefore generally only a small proportion of the total number of proteins from any organism are able to be detected and identified. The candidates investigated in this study are therefore likely to represent only a small number of all of the genes and proteins altered by expression of expanded repeat RNA. Furthermore, in this preliminary investigation of pathogenic pathways, candidates were chosen preferentially on the basis that they were commonly altered in flies expressing more than one of the repeat sequences and therefore this study does not attempt to investigate sequence-dependent effects of expression of each of the repeat sequences.

In using the *Drosophila* eye to model expanded repeat pathogenesis, it is also important to remember that expression of the toxic species is being induced in cells of both neuronal and non-neuronal origin. Therefore, the results obtained by screening candidates in this manner should be considered only as preliminary evidence for an interaction (or lack of interaction) with expanded repeats. Nevertheless, this sort of strategy is routinely used in *Drosophila* studies and has previously been successfully applied to identification of modifiers of neurodegenerative phenotypes. It should be noted, however, that since the candidates tested using the *Drosophila* eye in this study were identified in flies expressing expanded repeat RNA specifically in the neurons, it is quite possible that some of them are either more or less toxic in non-neuronal cells in the *Drosophila* eye than they may be in neurons. A *Drosophila* study examining the ability of modifiers of polyglutamine-induced eye phenotypes to similarly alter polyglutamine toxicity in post-mitotic neurons demonstrated that in a large number of cases examination of interactions in the eye and the brain does give consistent results. Nevertheless, there were three cases identified in this study where candidates which were able to enhance the polyglutamine eye phenotype had no effect on polyglutamine toxicity in the brain (331). For this reason, candidates identified in this study should be further investigated for their specific effect in neurons. Given that expression of these rCAG, rCUG or rAUUCU repeats does not appear to induce degeneration within the life-time of the fly when expressed pan-neuronally, one way in which this could be done is to investigate the ability of expression of these repeat

sequences to modify phenotypes associated with altered expression of the candidate genes in the nervous system.

8.4 Further Experiments

The results presented in this study are a preliminary examination of cellular processes which are disrupted in *Drosophila* expressing different expanded repeat sequences. Analyses performed in this study have successfully identified a number of candidates which show a genetic interaction with expanded repeats and therefore there are likely to be more functional interactors amongst the remaining data obtained by microarray and proteomic analysis. While there are indications that some common pathways are disrupted by expression of different hairpin-forming repeat sequences, candidates which should be further investigated are those which are altered uniquely in response to expression of particular repeat sequences, since these are likely to represent specific pathways involved in different expanded repeat diseases. Given that the candidates tested so far consistently showed stronger genetic interactions with untranslated CUG repeat RNA, it seems likely that there are also other pathogenic mechanisms at play in CAG and AUUCU RNA pathogenesis. It will also be important to confirm that these candidates, or the pathways in which they are involved, are altered during pathogenesis in the human diseases.

Following from the results described in this thesis, a model investigating the context-dependence of CAG repeat toxicity is being generated. This model will test the ability of an expanded CAG repeat tract within the context of the human *ataxin-12* 5'UTR, the only known example of an untranslated expanded CAG repeat involved in human disease, to induce neurodegeneration. The ability of candidates identified in the model of intrinsic rCAG repeat RNA toxicity described in this study to interact with this repeat tract will provide information regarding the degree of context-independence involved in pathogenesis in SCA12. Investigation of unique perturbations resulting from expression of the SCA12 non-coding RNA will also provide insight into disease-specific pathogenic pathways.

Since the commencement of this study, there have also been reports of situations where bi-directional transcription in the region containing the expanded

repeat tract can result in the expression of perfectly double-stranded RNA species in SCA8 and FXTAS (97, 332). This finding has led to suggestions that hairpin RNA may only be one component of pathogenesis and that the formation of perfectly double-stranded RNAs may also cause cellular dysfunction. The ability of this sort of RNA species to induce neuronal dysfunction is also being tested using this *Drosophila* system and will provide information on the contribution of this pathogenic agent in the human disease. The results obtained in this and other studies suggest that hairpin-forming repeat RNAs have the potential to play a role in pathogenesis of the expanded repeat diseases. Further use of *Drosophila* to investigate the contribution of different sorts of RNA species in both the polyglutamine and untranslated repeat diseases – that is hairpin RNAs versus perfectly double-stranded RNAs – as well as the role of the expanded polyglutamine peptide in the polyglutamine diseases will provide information on the degree to which each one of these agents is responsible for the degeneration seen in the disease situation. An understanding of the molecular components and pathways of pathogenesis will hopefully enable the design of rational treatments for these debilitating diseases.

Appendices

Appendix A

Genotype	Line
elav>+	c155
elav>rCAG	"C+D"
elav>rCUG	"A+B"
elav>rCAA	"A+B"

Table A1: Genotypes of flies analysed by 2D DIGE analysis. All repeat-expressing flies carry 2 insertions of the respective expanded repeat construct driven by the *elav^{c155}–GAL4* driver. The control *elav>+* line is the *elav^{c155}–GAL4* driver line out-crossed to the wild-type *w^{1118}* line. Letters are arbitrarily used to denote independent insertions of each repeat transgene and in each case, insertions on the same chromosome have been recombined to make the 2 insertion lines.

Spot Number	***elav>rCAA***		***elav>rCAG***	
	T-test	Average ratio	T-test	Average ratio
215	0.041	-1.25	0.036	-1.29
363	0.009	-1.38	0.032	-1.3
368	0.034	1.28	0.004	1.27
373	0.034	1.22	0.001	1.24
449	0.047	1.47	0.022	1.53
550	0.036	1.39	0.046	1.31
712	0.035	1.34	0.037	1.34
713	0.026	1.46	0.028	1.42
940	0.015	1.27	0.013	1.32
1183	0.003	1.44	0.004	1.48
1439	0.016	1.21	0.041	1.22
1444	0.016	1.63	0.007	1.65
1467	0.002	1.4	0.001	1.44
1520	0.029	1.57	0.010	1.47
1526	0.025	1.40	0.039	1.35
1593	0.039	-1.40	0.006	-1.46
1622	0.015	-1.57	2.2E-5	-1.54
1778	0.007	-1.51	0.006	-1.51
1911	0.049	1.47	0.023	1.46
1974	0.024	-1.77	0.029	-1.86
1983	0.013	-1.38	0.015	-1.37
1989	0.011	1.45	0.013	1.32
2040	0.022	-1.55	0.008	-1.45
2049	0.022	-1.63	0.004	-1.58
2300	0.025	1.62	0.043	1.46

Table A2: Spots altered in *elav>rCAA* and *elav>rCAG* flies compared to *elav>+*. The average ratio is the average change in spot intensity compared to *elav>+* across the four gels for each genotype. Spots were selected for $P<0.05$.

Spot Number	*elav>rCAA*		*elav>rCUG*	
	T-test	Average ratio	T-test	Average ratio
914	0.042	-1.24	0.020	-1.32
1497	0.006	-1.71	0.003	-1.43
1962	1.5E-5	1.27	0.003	1.35
1966	0.018	-1.35	0.017	-1.27

Table A3: Spots altered in *elav>rCAA* and *elav>rCUG* flies compared to *elav>+*. The average ratio is the average change in spot intensity compared to elav>+ across the four gels for each genotype. Spots were selected for P<0.05.

Spot Number	*elav>rCAA*		*elav>rCAG*		*elav>rCUG*	
	T-test	Average ratio	T-test	Average ratio	T-test	Average ratio
595	9.6E-4	1.51	4.6E-4	1.53	0.012	1.35
704	0.012	-1.54	0.002	-1.59	0.018	-1.36
1472	0.015	-1.52	0.002	-1.49	0.008	-1.37
2361	0.008	-4.09	0.013	-3.55	0.016	2.43

Table A4: Spots altered in *elav>rCAA*, *elav>rCAG* and *elav>rCUG* compared to *elav>+*. The average ratio is the average change in spot intensity compared to *elav>+* across the four gels for each genotype. Spots were selected for P<0.05.

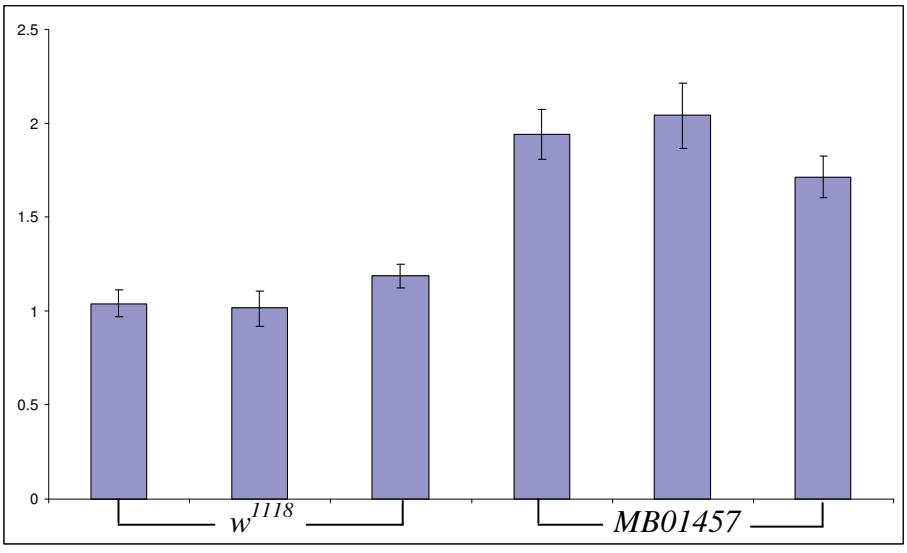


Figure A5: Transcript levels of DPx-2540-1 normalised to rp49. Quantitative real time PCR was performed on biological triplicates of w^{1118} and the *MB01457* insertion line. Transcript levels of DPx-2540-1 were normalised to rp49 levels in all cases. Error bars are ±SEM.

Spot Number	*elav>rCAG*	
	T-test	Average ratio
1101	**0.001**	**-1.51**
1136	0.010	-1.79
1738	0.047	-1.53

Table A6: Spots altered in *elav>rCAG* flies compared to *elav>rCAA*. Average ratio is the average change in spot intensity compared to *elav>rCAA* across the four gels for each genotype. Spots were selected for P<0.05, average ratio >1.5 or <-1.5. Bold text indicates the only spot changed common to *elav>rCUG* and *elav>rCAG* compared to *elav>rCAA*.

Spot Number	*elav>rCUG*	
	T-test	Average ratio
185	0.021	-1.54
366	0.037	-1.68
393	0.038	-2.85
402	0.035	-1.72
668	0.009	-1.6
970	0.044	-1.53
1101	**3.50E-04**	**-1.51**
1403	0.002	-1.84
1520	0.031	1.78
1531	0.002	1.72
1566	0.016	-1.53
1573	0.036	-2.12
1895	5.40E-05	-1.67
1896	0.003	1.52
1983	8.00E-05	-1.74
1998	0.001	-1.55
2026	4.40E-06	1.72
2027	4.40E-06	-1.66
2169	1.40E-05	2.21
2361	3.20E-04	9.93
2455	0.023	-1.79
2495	0.025	-1.59
2542	0.012	-1.66
2552	0.035	1.83
2621	0.035	1.95
2665	0.014	1.68
2681	0.038	-1.58

Table A7: Spots altered in *elav>rCUG* flies compared to *elav>rCAA*. Average ratio is the average change in spot intensity compared to elav>rCAA across the four gels for each genotype. Spots were selected for P<0.05, average ratio >1.5 or <-1.5. Bold text indicates the only spot changed common to *elav>rCUG* and *elav>rCAG* compared to *elav>rCAA*.

Appendix B

Genotype	Line	Insertions	Insertion Chromosome
elav>rCAG	Line 1	"A+E"	2
	Line 2	"C+D"	2
	Line 3	"G+I"	3
elav>rCUG	Line 1	"A+B"	2
	Line 2	"C+D"	2
	Line 3	"E+F"	3
elav>rCAA	Line 1	"A+B"	2
	Line 2	"C+I"	2
	Line 3	"E+F"	3
elav>+	-	-	-

Table B1: Genotypes of two insertion repeat lines analysed in microarray experiment 1. All repeat-expressing flies contain 2 insertions of the respective expanded repeat construct driven by the *elav–GAL4* driver. The control "*elav>+*" line is the *elav–GAL4* driver line out-crossed to the wild-type w^{1118} line. Letters are arbitrarily used to denote independent insertions of each repeat transgene and in each case 2 insertions on the same chromosome have been recombined to make the 2 insertion lines.

Construct	Line	Insertions	Insertion Chromosomes
rCAG	Line 1	"A+E+G+I"	2,3
	Line 2	"J+K+D+H"	1,2
rCUG	Line 1	"C+D+E+F"	2,3
	Line 2	"H+I+J+G"	2,3
rCAA	Line 1	"C+I+E+F"	2,3
	Line 2	"A+B+G+H"	2,3

Table B2: Genotypes of four insertion repeat lines analysed in microarray experiment 2 and *MBNL* interactions. Letters are arbitrarily used to denote independent insertions of each repeat transgene and in each case 2 insertions on the same chromosome have been recombined to make 2 insertion lines. Four insertion lines consist of two insertions on each chromosome listed. In the microarray experiment, expression of these constructs was driven by the *elav-GAL4* driver inserted on the second chromosome.

Table B3: Genes altered in *elav >rCAG* flies compared *to elav> rCAA* in microarray experiment 1. Selected for Log2(ratio) >0.5 or <-0.5, P<0.05.

Gene Title	Gene Symbol	Ensembl	Log2(ratio) *elav>rCAG* to *elav>rCAA*	Log2(ratio) *elav>rCUG* to *elav>rCAA*	Human orthologue	Function
stunted	sun	CG9032	**-1.50** **P=0.009**	-1.13 P=0.033	ATP5E	mitochondrial ATP synthase epsilon chain
defensin	Def	CG1385	**-1.37** **P=0.001**	-1.25 P=0.002		Immune response
CG9377	CG9377	CG9377	**-1.07** **P=0.016**	0.27 P=0.491		Serine-type endopeptidase
CG8297	CG8297	CG8297	**-0.97** **P=0.021**	-1.04 P=0.025	TXNDC15	Thioredoxin
beat-IIIc	beat-IIIc	CG15138	**-0.95** **P=0.044**	-1.01 P=0.035		
CG1531	CG1531	CG1531	**-0.89** **P=0.035**	-0.57 P=0.0127	ZNF650	Ubiquitin protein ligase
act up	capt	CG33979	**-0.88** **P=0.011**	-0.51 P=0.243	CAP1	adenylate cyclase associated protein, actin filament organisation
His3: CG31613	His3: CG31613	CG31613	**-0.88**	-0.65	HIST2H3	H3 histone
CG15313	CG15313	CG15313	**-0.83** **P=0.005**	-0.40 P=0.134		
His1: CG31617	His1: CG31617	CG31617	**-0.83** **P=0.006**	-0.71 P=0.045	HIST1H1	H1 histone
molting defective	mld	CG34100	**-0.82** **P=0.019**	-0.49 P=0.048		Ecdysone biosynthesis
His3: CG31613	His3: CG31613	CG31613	**-0.81** **P=0.004**	-0.74 P=0.014	HIST2H3	H3 histone
fat facets	faf	CG1945	**-0.77** **P=0.005**	-0.32 P=0.068	USP9X	Ubiquitin-specific protease
CG30044	s-cup	CG30044	**-0.75** **P=0.029**	-0.71 P=0.032		
Cad99C	Cad99C	CG31009	**-0.74** **P=0.021**	-0.89 P=0.012	PCDH15	protocadherin 15 precursor, smoothened signaling pathway, cell-cell adhesion
CG15642	CG15642	CG15642	**-0.74** **P=0.023**	-0.17 P=0.452		
Mediator complex subunit 17	MED17	CG7957	**-0.71** **P=0.036**	-0.59 P=0.077	CRSP6	co-factor for Sp1, transcription co-activator
CG9098	CG9098	CG9098	**-0.70** **P=0.002**	-0.44 P=0.092	BCAR3	Tyrosine kinase, estrogen independent cell division, cell cycle regulation
CG31638	CG31638	CG31638	**-0.70** **P=0.002**	-0.58 P=0.021	CCDC102A	Tropomyosin, component of myosin complex
CG9395	CG9395	CG9395	**-0.70** **P=0.020**	-0.23 P=0.352		
CG18437	CG18437	CG18437	**-0.70** **P=0.039**	-0.98 P=0.020		
Glycogenin	Glycogenin	CG9480	**-0.69** **P=0.023**	-0.67 P=0.055	GYG1	Glycogenin, mesoderm development
CG14223	CG14223	CG14223	**-0.69** **P=0.030**	-0.52 P=0.059		

Gene Title	Gene Symbol	Ensembl	Log2(ratio) *elav>rCAG* to *elav>rCAA*	Log2(ratio) *elav>rCUG* to *elav>rCAA*	Human orthologue	Function
CG15011	CG15011	CG15011	**-0.69** **P=0.020**	-0.30 P=0.118	NFXL1	Nuclear transcription factor
CG13001	CG13001	CG13001	**-0.69** **P=0.045**	-0.63 P=0.051	ZC4H2	
CG12814	CG12814	---	**-0.69** **P=0.015**	-0.23 P=0.086		
CG2010	CG2010	CG2010	**-0.67** **P=0.015**	-0.80 P=0.012		
Translocase of outer membrane 40	Tom40	CG12157	**-0.66** **P=0.038**	-0.62 P=0.060	TOMM40	mitochondrial import receptor
CG32056	CG32056	CG32056	**-0.65** **P=0.041**	-0.84 P=0.022	PLSCR1	Phospholipid scramblase, synaptic transmission
CG8833	CG8833	CG8833	**-0.63** **P=0.011**	-0.62 P=0.027	GPATC1	RNA processing
CG8505	Cpr49Ae	CG8505	**-0.63** **P=0.043**	-0.70 P=0.044		Cuticle protein
CG12641	CG12641	CG12641	**-0.63** **P=0.031**	-0.73 P=3.99E-5		
mindmelt	mbl	CG33197	**-0.62** **P=0.037**	-0.52 P=0.168	MBNL1	RNA binding
prickle-spiny legs	pk	CG11084	**-0.62** **P=0.041**	-0.80 P=0.056	PRICKLE2	neurite outgrowth
CG14662	CG14662	CG14662	**-0.62** **P=0.034**	-0.20 P=0.214		
CG31295	CG31295	CG31295	**-0.62** **P=0.006**	0.04 P=0.670		
CG1698	CG1698	CG1698	**-0.61** **P=0.045**	-0.55 P=0.150		
CG34104	CG34104	CG12102	**-0.61** **P=0.033**	-0.65 P=0.085		
CG9264	CG9264	CG9264	**-0.61** **P=0.038**	-0.87 P=0.008		
Zinc/iron regulated transporter-related protein 3	Zip3	CG6898	**-0.61** **P=0.013**	-0.40 P=0.179	SLC39A2	zinc transporter
CG8740	CG8740	CG8740	**-0.60** **P=0.014**	-0.43 P=0.033		
CG33980	CG33980	CG33980	**-0.60** **P=0.019**	-0.50 P=0.064		
elbow B	elB	CG4220	**-0.60** **P=0.010**	-0.53 P=0.173	DDI1	DNA damage repair
CG13214	Cpr47Ef	CG13214	**-0.60** **P=0.010**	-0.40 P=0.073		Cuticle protein
CG10321	CG10321	CG10321	**-0.60** **P=0.002**	-0.60 P=0.027		
Odorant-binding protein 19a	Obp19a	CG11748	**-0.59** **P=0.038**	-0.54 P=0.042		Sensory perception of chemical stimulus
CG8617	CG8617	CG8617	**-0.59** **P=0.016**	-0.67 P=0.085		

Gene Title	Gene Symbol	Ensembl	Log2(ratio) *elav>rCAG* to *elav>rCAA*	Log2(ratio) *elav>rCUG* to *elav>rCAA*	Human orthologue	Function
CG6695	CG6695	CG6695	**-0.59** **P=0.039**	-0.80 P=0.023	SFRS16	Splicing factor, arginine/serine-rich 16
CG9215	CG9215	CG9215	**-0.59** **P=0.029**	-0.47 P=0.030		
CG8925	CG8925	CG8925	**-0.58** **P=0.015**	-0.42 P=0.138		
CG32452	CG32452	CG32452	**-0.58** **P=0.006**	-0.26 P=0.053		
CG17002	CG17002	CG17002	**-0.57** **P=0.036**	-0.38 P=0.278	GPS2	Component of NCoR-HDAC3 complex
Rev1	Rev1	CG12189	**-0.57** **P=0.023**	-0.30 P=0.112	REV1	DNA-template dependent dCMP transferase, DNA lesion bypass
CG16857	CG16857	CG16857	**-0.57** **P=0.047**	-0.62 P=0.024		Cell adhesion
CG8281	CG8281	CG8281	**-0.56** **P=0.002**	-0.32 P=0.200		
Robo2	lea	CG5481	**-0.56** **P=0.037**	-0.58 P=0.037	ROBO1	axon guidance receptor
wolfram syndrome 1	wfs1	CG4917	**-0.56** **P=0.008**	-0.69 P=0.011	WFS1	modulates free calcium in ER, regulated by Sp1
serendipity beta	Sry-beta	CG7938	**-0.56** **P=0.031**	-0.47 P=0.022		
CG12768	CG12768	CG12768	**-0.55** **P=0.050**	-0.38 P=0.140		
tonalli	tna	CG7958	**-0.55** **P=0.043**	-0.60 P=0.032	ZMIZ1	transcriptional coactivator
CG5669	CG5669	CG5669	**-0.54** **P=0.011**	-0.62 P=0.007	SP1	Transcription factor
CG10362	CG10362	CG10362	**-0.54** **P=0.034**	-0.54 P=0.035	PDZK8	Intracellular signaling cascade, DAG binding
CG12910	CG12910	CG12910	**-0.54** **P=0.037**	-0.40 P=0.288		
CG4749	CG4749	CG4749	**-0.53** **P=0.006**	-0.42 P=0.030	NSUN4	
CG1537	CG1537	CG1537	**-0.53** **P=0.048**	0.27 P=0.531		
CG5621	CG5621	CG5621	**-0.53** **P=0.049**	-0.63 P=0.053		
CG17446	CG17446	CG17446	**-0.53** **P=0.046**	-0.36 P=0.103	CXXC1	regulation of histone modification and cytosine methylation
CG14842	CG34388	CG14843	**-0.53** **P=0.004**	-0.27 P=0.530		
Painting of fourth	Pof	CG3691	**-0.53** **P=0.024**	-0.39 P=0.042		
CG3308	CG3308	CG3308	**-0.53** **P=0.009**	-0.32 P=0.193	TATDN1	deoxyribonuclease
CG2713	CG2713	CG2713	**-0.53** **P=0.022**	-0.67 P=0.023	TIMM50	Translocase of inner mitochondrial membrane
CG7277	CG7277	CG7277	**-0.52** P=0.011	-0.41 P=0.214	COQ6	Co-enzyme Q, component of respiratory chain

Gene Title	Gene Symbol	Ensembl	Log2(ratio) *elav>rCAG* to *elav>rCAA*	Log2(ratio) *elav>rCUG* to *elav>rCAA*	Human orthologue	Function
CG9799	CG9799	CG9799	**-0.52** **P=0.034**	-0.47 P=0.122	WDR36	rRNA processing, Human orthologue has association with adult-onset primary open-angle glaucoma (POAG)
CG9394	CG9394	CG9394	**-0.52** **P=0.031**	0.02 P=0.912		Lipid metabolism
CstF-50	CstF-50	CG2261	**-0.52** **P=0.044**	-0.26 P=0.203	CSTF1	polyadenylation and 3' cleavage or pre-mRNA
Suppressor of cytokine signaling at 36E	Socs36E	CG15154	**-0.52** **P=0.047**	-0.61 P=0.033	SOCS1	suppressor of cytokine signaling, JAK-STAT cascade
D19B	D19B	CG10270	**-0.52** **P=0.015**	-0.46 P=0.059		
Rab27	Rab27	CG14791	**-0.51** **P=0.025**	-0.38 P=0.182	RAB27A	vesicle trafficking
CG9213	CG9213	CG9213	**-0.51** **P=0.044**	-0.56 P=0.007	CWF19L2	cell cycle control
CG2950	CG2950	CG2950	**-0.51** **P=0.050**	-0.40 P=0.151		RNA binding
CG6723	CG6723	CG6723	**-0.51** **P=0.005**	0.20 P=0.406		Transmembrane transporter
Bem46	Bem46	CG18642	**-0.51** **P=0.041**	-0.60 P=0.016	ABHD13	hydrolase activity
CG8086	CG8086	CG8086	**-0.50** **P=0.020**	-0.24 P=0.298	ODF3	Outer dense fibre component, cytoskeleton
Cyp4ac3	Cyp4ac3	CG14031	**0.50** **P=0.033**	0.19 P=0.330		Electron carrier
papilin	Ppn	CG33103	**0.51** **P=0.002**	1.02 P=0.063	PAPLN	Glycoprotein, ECM component
Lectin24Db	lectin-24Db	CG2958	**0.52** **P=0.040**	1.52 P=0.100	FCER2	Lymphocyte Ig receptor
CG10026	CG10026	CG10026	**0.52** **P=0.011**	0.94 P=0.026	TTPA	Vitamin E metabolism, deficiency leads to cerebellar degeneration
hormone receptor	GRHR	CG11325	**0.54** **P=0.026**	-0.04 P=0.852	GNRHR	Gonadotropin releasing hormone receptor, lipid metabolism and homeostasis
CG31629	CG31629	CG31629	**0.55** **P=0.045**	0.19 P=0.475		
Mis-expression Suppressor of Ras 1	NFAT	CG11172	**0.58** P=0.013	0.14 P=0.528	NFAT5	Transcription factor, regulation of osmolarity, possible association with SCA4
Cyp28d2	Cyp28d2	CG6081	**0.59** P=0.039	0.30 P=0.247		Electron carrier
Modifier67.2	mod(mdg4)	CG32491	**0.60** **P=0.001**	0.42 P=0.006		Regulation of apoptosis, regulation of chromatin assembly
CG31781	CG31781	CG31781	**0.61** **P=0.037**	0.70 P=9.21E-5	LIP	lipase
CG30151	CG30151	CG30151	**0.62** **P=0.002**	0.50 P=0.110		

Gene Title	Gene Symbol	Ensembl	Log2(ratio) *elav>rCAG* to *elav>rCAA*	Log2(ratio) *elav>rCUG* to *elav>rCAA*	Human orthologue	Function
Immunodeficiency	imd	CG5576	**0.63** **P=0.014**	0.36 P=0.386		Innate immune response
CG14856	CG14856	CG14856	**0.63** **P=0.004**	0.35 P=0.248		
fettucine	cic	CG5067	**0.65** **P=0.048**	0.39 P=0.195	SOX11	neuron survival and outgrowth
CG15293	CG15293	CG15293	**0.71** **P=0.032**	1.26 P=0.018		
CG12116	CG12116	CG12116	**0.73** **P=0.002**	0.56 P=0.015		
Imaginal disc growth factor 5	Idgf5	CG5154	**0.78** **P=0.043**	0.85 P=0.073		Imaginal disc development
CG13783	Pvf3	CG13783	**0.79** **P=0.020**	0.45 P=0.339		Hemocyte migration, VEGF-like activity
CG13784	CG13784	CG13784	**0.81** **P=0.048**	0.90 P=0.075	PQLC1	
CG4306	CG4306	CG4306	**0.82** **P=0.036**	1.23 P=0.020	GGCT	glutathione homeostasis, release of cytochrome c from mitochondria
fragment B	alpha-Est2	CG2505	**0.83** **P=0.035**	0.86 P=0.057	EST2	carboxylesterase 2, lipid metabolism
CG8942	CG8942	CG8942	**0.84** **P=0.026**	1.59 P=0.065	FLJ14712	
CG11072	CG34346	CG11072	**0.84** **P=0.037**	0.68 P=0.024		
CG13086	CG13086	CG13086	**0.89** **P=0.039**	1.40 P=0.024		
CG4716	CG4716	CG4716	**0.90** **P=0.005**	1.33 P=0.007		
CG15068	CG15068	CG15068	**0.97** **P=0.012**	0.91 P=0.016		
CG16836	CG16836	CG16836	**1.39** **P=0.032**	1.45 P=0.071		
Immune induced molecule 1	IM1	CG18108	**1.42** **P=0.018**	1.52 P=0.117		Immune response
Immune induced molecule 2	IM2	CG18106	**2.29** **P=0.005**	1.82 P=0.196		Immune response

Table B4: Genes changed in *elav>rCUG* flies compared to *elav>rCAA* in experiment 1.
Log2 >0.5 or <-0.5, p<0.05.

Gene Title	Gene Symbol	Ensembl	log2(ratio) *elav>rCAG* to *elav>rCAA*	log2(ratio) *elav>rCUG* to *elav>rCAA*	Human orthologue	Function
defensin	Def	CG1385	-1.37 P=0.001	**-1.25** **P=0.002**		Immune response
stunted	sun	CG9032	-1.50 P=0.009	**-1.13** **P=0.033**	ATP5E	mitochondrial ATP synthase epsilon chain
Plum	bw	CG17632	-0.66 P=0.053	**-1.13** **P=0.027**		Eye pigment precursor transport
CG8297	CG8297	CG8297	-0.97 P=0.021	**-1.04** **P=0.025**	TXNDC15	Thioredoxin
beat-IIIc	beat-IIIc	CG15138	-0.95 P=0.044	**-1.01** **P=0.035**		
CG18437	CG18437	CG18437	-0.70 P=0.039	**-0.98** **P=0.020**	C2orf21	
CG33528	CG33528	CG33528	-0.50 P=0.152	**-0.97** **P=0.011**	SLC18A2	vesicular monoamine transporter
CG10914	CG10914	CG10914	-0.60 P=0.059	**-0.91** **P=0.015**	hAtNOS1	mitochondrial GTP binding
Cad99C	Cad99C	CG31009	-0.74 P=0.021	**-0.89** **P=0.012**	PCDH15	protocadherin 15 precursor
Doughnut	dnt	CG17559	-0.45 P=0.024	**-0.89** **P=0.005**	RYK	Receptor-like tyrosine kinase, required for neurite outgrowth
CG9264	CG9264	CG9264	-0.61 P=0.038	**-0.87** **P=0.008**		
sparkling	sv	CG11049	-0.63 P=0.080	**-0.87** **P=0.038**	PAX5	B cell specific transcription factor/ midbrain dopaminergic neuron specification
turtle	tutl	CG15427	-0.73 P=0.095	**-0.86** **P=0.015**	IGSF9	Required for coordinated motor control, Immunoglobulin super family member 9
CG13681	CG13681	CG13681	-0.27 P=0.408	**-0.85** **P=0.013**		
CG8798	CG8798	CG8798	-0.46 P=0.035	**-0.84** **P=0.007**	LONP1	mitochondrial ATP-dependent protease, removal of oxidised aconitase
CG32056	CG32056	CG32056	-0.65 P=0.041	**-0.84** **P=0.022**	PLSCR1	Phospholipid scramblase, synaptic transmission
CG8213	CG8213	CG8213	-0.52 P=0.180	**-0.83** **P=0.030**		
CG2987	alpha-catenin-related	CG2987	-0.54 P=0.073	**-0.82** **P=0.033**	CTNL1	cell-cell recognition
defective pro-ventriculus	dve	CG5799	-0.24 P=0.367	**-0.81** **P=0.012**		Transcription factor activity
CG9948	CG9948	CG9948	-0.42 P=0.237	**-0.81** **P=0.048**		
CG9597	CG9597	CG9597	-0.06 P=0.845	**-0.80** **P=0.007**		
CG6695	CG6695	CG6695	-0.59 P=0.039	**-0.80** **P=0.023**	SFRS16	Splicing factor, arginine/serine-rich 16

Gene Title	Gene Symbol	Ensembl	log2(ratio) *elav>rCAG* to *elav>rCAA*	log2(ratio) *elav>rCUG* to *elav>rCAA*	Human orthologue	Function
CG2010	CG2010	CG2010	-0.67 P=0.015	**-0.80** **P=0.012**		
CG14024	CG14024	CG14024	-0.50 P=0.100	**-0.78** **P=0.045**		
big brain	bib	CG4722	-0.45 P=0.197	**-0.78** **P=0.021**	Aquaporin protein family	
CG15522	CG15522	CG15522	-0.52 P=0.088	**-0.76** **P=0.047**		
Neprilysin 1	Nep1	CG5905	-0.50 P=0.067	**-0.76** **P=0.014**	ECE2	Endothelin converting enzyme, role in clearing Abeta in the brain
CG1764	CG1764	CG1764	-0.49 P=0.070	**-0.75** **P=0.026**	DDAH1	
Open rectifier K[+] channel 1	Ork1	CG1615	-0.56 P=0.130	**-0.75** **P=0.024**	KCNK4	2 pore K+ channel
CG14896	CG14896	CG14896	-0.28 P=0.078	**-0.75** **P=0.010**		
CG17834	CG17834	CG17834	-0.51 P=0.134	**-0.75** **P=0.018**		
CG5022	CG5022	CG5022	-0.23 P=0.152	**-0.74** **P=0.011**	FRMD3	
longitudinals absent	lola	CG12052	-0.44 P=0.056	**-0.74** **P=0.019**	ZBTB3	Transcription factor activity
His3: CG31613	His3: CG31613	CG31613	-0.81 P=0.004	**-0.74** **P=0.014**	HIST2H3	H3 histone
CG5514	CG5514	CG5514	-0.17 P=0.310	**-0.74** **P=0.004**	FAM44A	Structural component of cell wall
CG12641	CG12641	CG12641	-0.63 P=0.031	**-0.73** **P=3.99E-05**		
CG5660	CG5660	CG5660	-0.21 P=0.394	**-0.73** **P=0.038**		
CG13148	CG13148	CG13148	-0.47 P=0.118	**-0.73** **P=0.044**		
His1:CG31617	His1: CG31617	CG31617	-0.83 P=0.006	**-0.71** **P=0.045**	HIST1H1	H1 histone
U26	U26	CG13401	-0.60 P=0.075	**-0.71** **P=0.019**	AASDH	2-aminoadipic 6-semialdehyde dehydrogenase, fatty acid metabolism
CG1688	CG1688	CG1688	-0.44 P=0.042	**-0.71** **P=0.005**	KCNK6	Potassium channel
CG30044	CG30044	CG30044	-0.75 P=0.029	**-0.71** **P=0.032**		
CG14005	CG14005	CG14005	-0.41 P=0.133	**-0.71** **P=0.041**		
ryanodine receptor	Rya-r44F	CG10844	-0.46 P=0.072	**-0.70** **P=0.022**	RYR2	Cardiac muscle calcium regulation
taranis	tara	CG6889	-0.54 P=0.183	**-0.70** **P=0.036**		Chromatin-mediated maintenance of transcription
CG32392	CG32392	CG32392	-0.58 P=0.055	**-0.70** **P=0.031**	RSHL2	ciliary motion, microtubule associated protein

Gene Title	Gene Symbol	Ensembl	log2(ratio) *elav>rCAG* to *elav>rCAA*	log2(ratio) *elav>rCUG* to *elav>rCAA*	Human orthologue	Function
Arginine methyltransferase 4	Art4	CG5358	-0.32 P=0.177	**-0.70** **P=0.028**	CARM1	co-activator interacting arginine methyltransferase, interacts with HuR (elav-L), histone methyltransferase
CG8505	Cpr49Ae	CG8505	-0.63 P=0.043	**-0.70** **P=0.044**		Component of cuticle
CG32105	CG32105	CG32105	-0.55 P=0.147	**-0.70** **P=0.025**	LMX1A	Differentiation of midbrain dopaminergic neurons, transcription factor activity
wolfram syndrome 1	wfs1	CG4917	-0.56 P=0.008	**-0.69** **P=0.011**	WFS1	modulates free calcium in ER, regulated by Sp1
CG9866	CG9866	CG9866	-0.36 P=0.096	**-0.67** **P=0.016**		
CG2713	CG2713	CG2713	-0.53 P=0.022	**-0.67** **P=0.029**	TIMM50	Translocase of inner mitochondrial membrane
CG3160	CG3160	CG3160	-0.27 P=0.123	**-0.67** **P=0.035**	PGAP1	Protein transport and metabolism
male specifc lethal	msl-3	CG8631	-0.36 P=0.162	**-0.66** **P=0.026**	MSL3L1	histone acetylation
CG6512	CG6512	CG6512	-0.48 P=0.027	**-0.65** **P=0.011**	AFG3L2	mitochondrial ATPase, highest expression in skeletal muscle and heart
sequoia	seq	CG32904	-0.40 P=0.102	**-0.65** **P=0.020**		Dendrite morphogenesis
CG12483	CG12483	CG12483	-0.51 P=0.286	**-0.65** **P=0.015**		
SIFamide	IFa	CG33527	-0.06 P=0.812	**-0.64** **P=0.043**		
metabotropic GABA-B receptor subtype 3	GABA-B-R3	CG3022	-0.40 P=0.023	**-0.64** **P=0.007**		
gartenzwerg	garz	CG8487	-0.36 P=0.075	**-0.63** **P=0.021**	GBF1	Guanine nucleotide exchange factor
CG31012	CG31012	CG31012	-0.31 P=0.214	**-0.63** **P=0.034**	SH3D19	Enhances TNF-mediated cell death
CG8187	CG8187	CG8187	-0.43 P=0.014	**-0.63** **P=0.001**		
eIF3-S10	eIF3-S10	CG9805	-0.41 P=0.102	**-0.62** **P=0.022**	EIF3A	Translation initiation factor
male specifc lethal	msl-2	CG3241	-0.36 P=0.043	**-0.62** **P=0.011**	MSL2L1	H4 histone acetylation
CG5669	CG5669	CG5669	-0.54 P=0.011	**-0.62** **P=0.007**	SP1	Transcription factor
---	---	CR33947	-0.46 P=0.092	**-0.62** **P=0.037**		
CG8833	CG8833	CG8833	-0.63 P=0.011	**-0.62** **P=0.027**	GPATC1	RNA processing
CG16857	CG16857	CG16857	-0.57 P=0.047	**-0.62** **P=0.024**		
hoepel2	hoe2	CG15624	-0.36 P=0.110	**-0.61** **P=0.009**	OCA2	regulation of pH of melanocytes

Gene Title	Gene Symbol	Ensembl	log2(ratio) *elav>rCAG* to *elav>rCAA*	log2(ratio) *elav>rCUG* to *elav>rCAA*	Human orthologue	Function
Suppressor of cytokine signaling at 36E	Socs36E	CG15154	-0.52 P=0.047	**-0.61** **P=0.033**	SOCS1	suppressor of cytokine signaling, JAK-STAT cascade
Rep1	Rep1	CG8357	-0.31 P=0.103	**-0.61** **P=0.022**	CIDEA	cell death activator, caspase activated nuclease
CG6073	CG6073	CG6073	-0.08 P=0.499	**-0.61** **P=0.017**	c8orf30A	
CG10321	CG10321	CG10321	-0.60 P=0.002	**-0.60** **P=0.027**		
tonalli	tna	CG7958	-0.55 P=0.043	**-0.60** **P=0.032**	ZMIZ1	transcriptional coactivator
Bem46	Bem46	CG18642	-0.51 P=0.041	**-0.60** **P=0.016**	ABHD13	hydrolase activity
CG4330	CG4330	CG4330	-0.49 P=0.040	**-0.60** **P=0.031**		
CG3703	CG3703	CG3703	-0.36 P=0.084	**-0.59** **P=0.018**	RUNDC1	modulates p53 activity
deep orange	dor	CG3093	-0.44 P=0.018	**-0.59** **P=0.011**	VPS18	vacuolar protein sorting, eye pigment metabolism
CG33331	CG33331	CG33331	-0.30 P=0.068	**-0.59** **P=0.021**		Mitochondrial import protein
CG8273	CG8273	CG8273	-0.65 P=0.068	**-0.59** **P=0.048**	SON	Double stranded RNA binding, DNA binding
CG8511	Cpr49Ag	CG8511	-0.47 P=0.047	**-0.59** **P=0.049**		Cuticle component
CG13792	CG13792	CG13792	-0.39 P=0.122	**-0.59** **P=0.037**		
CG10105	Sin1	CG10105	-0.37 P=0.060	**-0.59** **P=0.008**	MAPKAP1	Stress-activated protein kinase, component of TOR complexes, regulation of apoptosis
CG33696	CG33696	CG33696	-0.10 P=0.364	**-0.59** **P=0.011**		
CG10251	CG10251	CG10251	-0.13 P=0.580	**-0.58** **P=0.027**		
guanylate cyclase 99B	Gycalpha99B	CG1912	-0.45 P=0.027	**-0.58** **P=0.039**	GUCY1A2	cGMP synthesis, A2 isoform is important in neurotransmission
CG15529	CG15529	CG15529	-0.33 P=0.065	**-0.58** **P=0.017**	BLNK	Protein binding
CG9134	CG9134	CG9134	-0.47 P=0.065	**-0.58** **P=0.046**	BCAN	Ig receptor
CG31638	CG31638	CG31638	-0.70 P=0.002	**-0.58** **P=0.021**	CCDC102A	Component of myosin complex
Robo2	lea	CG5481	-0.56 P=0.037	**-0.58** **P=0.037**	ROBO1	axon guidance receptor
prp8	prp8	CG8877	-0.40 P=0.094	**-0.58** **P=0.042**	PRPF8	Component of spliceosomes, highest expression in skeletal muscle and heart, candidate for retinitis pigmentosa
CG13594	CG13594	CG13594	-0.30 P=0.111	**-0.57** **P=0.033**		

Gene Title	Gene Symbol	Ensembl	log2(ratio) *elav>rCAG* to *elav>rCAA*	log2(ratio) *elav>rCUG* to *elav>rCAA*	Human orthologue	Function
CG3570	CG3570	CG3570	-0.17 P=0.216	**-0.57** **P=0.010**	C7orf60	
CG8211	CG8211	CG8211	-0.33 P=0.111	**-0.57** **P=0.003**	INTS2	Subunit of integrator complex, 3' RNA processing of small nuclear RNAs
CG14521	CG14521	CG14521	-0.27 P=0.059	**-0.57** **P=0.004**		
CG6734	CG6734	CG6734	-0.36 P=0.068	**-0.57** **P=0.022**	WDR81	
CG11617	CG11617	CG11617	-0.42 P=0.013	**-0.56** **P=0.026**	MKX	Homeobox protein, transcription factor activity
CG6854	CG6854	CG6854	-0.24 P=0.067	**-0.56** **P=0.003**	CTPS	Synthesis of CTP from UTP
CG5653	CG5653	CG5653	-0.16 P=0.200	**-0.56** **P=0.012**	SMOX	Polyamine oxidase, eye development
spire	spir	CG10076	-0.26 P=0.147	**-0.56** **P=0.022**	SPIRE1	actin organisation, highest expression in cerebellum
CG11695	CG11695	CG11695	-0.34 P=0.230	**-0.56** **P=0.049**		
CG9213	CG9213	CG9213	-0.51 P=0.044	**-0.56** **P=0.007**	CWF19L2	cell cycle control
neither inactivation nor after-potential A	ninaA	CG3966	-0.23 P=0.249	**-0.55** **P=0.041**	PPI	Peptidyl-prolyl cis-trans isomerase
CG17111	CG17111	CG17111	-0.41 P=0.074	**-0.55** **P=0.025**		
CG10702	CG10702	CG10702	-0.33 P=0.021	**-0.55** **P=0.010**		
CG32085	CG32085	CG32085	-0.24 P=0.251	**-0.55** **P=0.045**	FBXL16	Ubiquitination and protein catabolism
CG10440	CG10440	CG10440	-0.40 P=0.058	**-0.55** **P=0.031**	KCTD15	voltage gated potassium channel
CG33249	CG33249	CG33249	-0.42 P=0.010	**-0.55** **P=0.026**		
CG10362	CG10362	CG10362	-0.54 P=0.034	**-0.54** **P=0.035**	PDZK8	signalling
CG1550	CG1550	CG1550	-0.28 P=0.113	**-0.54** **P=0.024**	TTLL12	Tubulin tyrosine ligase, involved in microtubule stabilisation
Odorant-binding protein 19a	Obp19a	CG11748	-0.59 P=0.038	**-0.54** **P=0.042**		Sensory perception of chemical stimulus
Mediator complex subunit 24	MED24	CG7999	-0.21 P=0.323	**-0.54** **P=0.020**	MED24	component of mediator complex, transcriptional co-activator
Spt5	Spt5	CG7626	-0.48 P=0.022	**-0.54** **P=0.018**	SUPT5H	regulator of transcriptional elongation
Protein Kinase D	PKD	CG7125	-0.35 P=0.084	**-0.54** **P=0.020**	PRKD1	serine threonine kinase
CG33717	CG33717	CG33717	-0.15 P=0.540	**-0.54** **P=0.007**	TMEM11	
Von Hippel Lindau	Vhl	CG13221	-0.36 P=0.060	**-0.53** **P=0.017**	VHL	microtubule binding and stabilising

Gene Title	Gene Symbol	Ensembl	log2(ratio) *elav>rCAG* to *elav>rCAA*	log2(ratio) *elav>rCUG* to *elav>rCAA*	Human orthologue	Function
Gene 1	Hsp67Ba	CG4167	-0.27 P=0.021	**-0.53** **P=0.008**		
xl6	xl6	CG10203	-0.17 P=0.231	**-0.53** **P=0.011**	SFRS7	splicing factor, arginine/serine rich
deborg	debcl	CG33134	-0.42 P=0.073	**-0.53** **P=0.014**	BOK	Apoptosis inhibitor
Patsas	Patsas	CG6618	-0.39 P=0.062	**-0.53** **P=0.005**	HIP14	Huntington interacting protein, intracellular trafficking
CG9121	CG9121	CG9121	-0.22 P=0.414	**-0.53** **P=0.004**		
CG9297	CG9297	CG9297	-0.31 P=0.103	**-0.53** **P=0.027**	SRL	Calcium storage and regulation in skeletal and cardiac muscle, cell adhesion
rhomboid-5	rho-5	CG33304	-0.48 P=0.055	**-0.53** **P=0.047**	RHBDF1	serine protease, cleaves spitz
CG3257	CG3257	CG3257	-0.37 P=0.018	**-0.53** **P=0.009**		
CG31211	CG31211	CG31211	-0.21 P=0.074	**-0.52** **P=0.009**	SFRS18	splicing factor, arginine/serine rich
columbus	Hmgcr	CG10367	-0.47 P=0.027	**-0.52** **P=0.019**	HMGCR	HMG-CoA reductase, cholesterol synthesis
Brain Tumor	brat	CG10719	-0.45 P=0.089	**-0.52** **P=0.011**		Negative regulation of neuroblast differentiation
CG15412	CG15412	CG15412	-0.34 P=0.175	**-0.52** **P=0.007**		
CG9170	CG9170	CG9170	-0.26 P=0.199	**-0.52** **P=0.015**		
CG16952	CG16952	CG16952	-0.32 P=0.058	**-0.52** **P=0.009**	BTBD7	cell proliferation
CG17324	CG17324	CG17324	-0.37 P=0.139	**-0.52** **P=0.010**	UGT1A3	UDP-glucuronosyl and UDP-glucosyl transferase
CG12822	CG12822	CG12822	-0.35 P=0.062	**-0.51** **P=0.042**	C9orf156	
CG4630	CG4630	CG4630	-0.40 P=0.034	**-0.51** **P=0.016**	SLC22A3/ Oct3	Transportation of neurotoxins and neurotransmitters
CG16801	Hr51	CG16801	-0.130 P=0.538	**-0.51** **P=0.016**	NR2E3	Retinal nuclear receptor, ligand dependent transcription factor
CG9279	CG9279	CG9279	-0.37 P=0.039	**-0.51** **P=0.025**		
modifier of rpr and grim, ubiquitously expressed	morgue	CG15437	-0.29 P=0.065	**-0.51** **P=0.014**	UBE2	Ubiquitin protein ligase, regulation of cell death
CG8944	CG8944	CG8944	-0.15 P=0.349	**-0.51** **P=0.032**		
Pigment dispersing factor	Pdf	CG6496	-0.13 P=0.698	**-0.51** **P=0.010**		
CG8422	Dh44-R1	CG8422	-0.23 P=0.219	**-0.51** **P=0.047**	CRHR2	Corticotropin-releasing hormone receptor
CG4293	CG4293	CG4293	-0.02 P=0.871	**-0.50** **P=0.016**	ERGIC2	

Gene Title	Gene Symbol	Ensembl	log2(ratio) *elav>rCAG* to *elav>rCAA*	log2(ratio) *elav>rCUG* to *elav>rCAA*	Human orthologue	Function
CG33528	Vmat	CG33528	-0.34 P=0.128	**-0.50** **P=0.020**	SLC18a2	vesicular monoamine transporter
SNFs Protein Partner	pps	CG6525	-0.28 P=0.334	**-0.50** **P=0.046**	DATF1	Apoptotic pathways
CG11658	CG11658	CG11658	-0.41 P=0.003	**-0.50** **P=0.017**	FBXO32	Ubiquitin protein ligase, upregulated during muscle atrophy
Odorant-binding protein 99c	Obp99c	CG7584	0.30 **P=**0.218	**0.51** **P=0.043**		Sensory perception of chemical stimulus
CG6543	CG6543	CG6543	0.03 P=0.908	**0.51** **P=0.030**	ECHS1	Fatty acid oxidation, mitochondrial protein
CG4042	CG4042	CG4042	0.36 P=0.100	**0.51** **P=0.021**		
Cytochrome P450-4e2	Cyp4e2	CG2060	0.26 P=0.006	**0.51** **P=0.042**	CYP4B1	Electron carrier activity
neuroserpin	Spn4	CG9453	0.15 P=0.360	**0.52** **P=0.023**	SERPINI1	regulation of axonal growth and neural plasticity, serine protease inhibitor
CG4074	CG4074	CG4074	0.36 P=0.150	**0.52** **P=0.005**	DSCR3	Down Syndrome critical region 3, vacuolar transport
CG7997	CG7997	CG7997	0.11 P=0.338	**0.53** **P=0.006**	NAGA	lysosomal glycohydrolase
CG4725	CG4725	CG4725	0.13 P=0.705	**0.53** **P=0.006**		
Cytochrome b5-related	Cyt-b5-r	CG13279	0.05 P=0.885	**0.53** **P=0.041**	FADS1	Lipid transport and metabolism
larval-opioid-receptor	FR	CG2114	0.31 P=0.251	**0.54** **P=0.048**		
CG14526	CG14526	CG14526	0.33 P=0.032	**0.55** **P=0.002**		
Peroxiredoxin 6005	Prx6005	CG3083	0.41 P=0.070	**0.55** **P=0.022**	PRDX6	Redox regulation of cell
CG4302	CG4302	CG4302	0.09 P=0.381	**0.55** **P=0.001**	UGT2B7	Conjugation and elimination of toxic compounds
CG1941	CG1941	CG1941	0.39 P=0.111	**0.55** **P=0.018**	MOGAT2	Triacylglycerol synthesis
CG12116	CG12116	CG12116	0.73 P=0.002	**0.56** **P=0.015**		
CG6364	CG6364	CG6364	0.35 P=0.064	**0.56** **P=0.025**	UCK2	pyrimidine ribonucleoside kinase, production of UMP and CMP
CG30002	CG30002	CG30002	0.26 P=0.302	**0.56** **P=0.040**		
CG2241	CG2241	CG2241	0.59 P=0.220	**0.56** **P=0.044**		
Ance-5	Ance-5	CG10142	0.38 P=0.512	**0.57** **P=0.036**		
diazepam binding inhibitor	Dbi	CG8627	0.24 P=0.148	**0.57** **P=0.021**	DBI	Lipid metabolism and modulation of signalling at GABA-A neuron synapses
CG4019	CG4019	CG4019	0.25 P=0.100	**0.57** **P=0.033**	AQP	Aquaporin protein family

Gene Title	Gene Symbol	Ensembl	log2(ratio) *elav>rCAG* to *elav>rCAA*	log2(ratio) *elav>rCUG* to *elav>rCAA*	Human orthologue	Function
CG3603	CG3603	CG3603	0.25 P=0.265	**0.58** **P=0.035**	HSD17B8	Regulation of concentration of biologically active estrogens and androgens
CG15092	CG15092	CG15092	0.30 P=0.151	**0.59** **P=0.044**		
CG14933	CG14933	CG14933	0.23 P=0.447	**0.59** **P=0.016**	SPINK	Trypsin inhibitor
PGRP-SC2	PGRP-SC2	CG14745	0.12 P=0.508	**0.59** **P=0.018**	PGLYRP1	peptidoglycan recognition protein
CG17244	CG17244	CG17244	0.23 P=0.297	**0.59** **P=0.006**		
Cytochrome P450 related BF6-2	Cyp6a22	CG10240	0.25 P=0.202	**0.59** **P=0.049**	CYP3A5	Electron carrier
Esterase-6	Est-6	CG6917	-0.09 P=0.613	**0.60** **P=0.017**		Pheromone biosynthesis
CG31673	CG31673	CG31673	0.25 P=0.014	**0.61** **P=0.018**		
putative noncoding RNA 007:3R	CG6503	CG6503	0.40 P=0.065	**0.61** **P=0.002**		
Cyp12a5	Cyp12a5	CG11821	0.39 P=0.019	**0.62** **P=0.011**	CYP24A1	Electron carrier
CG31469	CG31469	CG31469	0.27 P=8.73-05	**0.62** **P=0.004**		
Trans-aldolase	Tal	CG2827	0.14 P=0.306	**0.62** **P=0.004**	TALDO1	Reduction of reactive oxygen intermediates
CG3663	CG3663	CG3663	0.22 P=0.119	**0.64** **P=0.021**	ISOC1	Isochorismatase hydrolase
CG7322	CG7322	CG7322	0.01 P=0.960	**0.64** **P=0.012**	DCXR	Dicarbonyl L-xylulose reductase
CG10799	CG10799	CG10799	0.34 P=0.064	**0.65** **P=0.031**		
CG6067	CG6067	CG6067	0.41 P=0.019	**0.65** **P=0.008**		
Cyp12c1	Cyp12c1	CG4120	0.15 P=0.520	**0.66** **P=0.003**	CYP24A1	Mitochondrial enzyme that inactivates metabolites of vitamin D
CG6206	CG6206	CG6206	0.14 P=0.172	**0.66** **P=0.029**	MAN2B1	hydrolyses alpha-D-mannose, defects result in lysosomal mannosidosis
CG4447	CG4447	CG4447	0.30 P=0.192	**0.67** **P=0.008**	MRRF	Mitochondrial, release of ribosomes from mRNA at stop codon
CG5840	CG5840	CG5840	0.42 P=0.159	**0.67** **P=0.019**	PYCRL	pyrroline-5-carboxylate reductase-like
CG11072	CG34346	CG11072	0.84 P=0.037	**0.68** **P=0.024**		
Dbeta3	nAcRbeta-21C	CG11822	0.12 P=0.673	**0.68** **P=0.048**		
anon-fast-evolving-1F7	Scamp	CG9195	0.49 P=0.130	**0.68** **P=0.032**	SCAMP1	vesicular transport to cell surface
CG3534	CG3534	CG3534	0.32 P=0.170	**0.68** **P=0.014**	XYLB	Energy metabolism

Gene Title	Gene Symbol	Ensembl	log2(ratio) *elav>rCAG* to *elav>rCAA*	log2(ratio) *elav>rCUG* to *elav>rCAA*	Human orthologue	Function
regucalcin	regucalcin	CG1803	0.26 P=0.367	**0.68** **P=0.042**	RGN	Calcium homeostasis
CG31780	CG18477	CG18477	0.13 P=0.694	**0.68** **P=0.045**		
CG5819	CG5819	CG5819	0.48 P=0.204	**0.68** **P=0.020**		
CG13845	CG34376	CG13845	0.42 P=0.137	**0.69** **P=0.019**		
CG3246	CG3246	CG3246	0.18 P=0.617	**0.70** **P=0.024**		
CG31781	CG31781	CG31781	0.61 P=0.037	**0.70** **P=9.22-05**	LIP	lipase
Adenosine deaminase-related growth factor D	Adgf-D	CG9621	-0.02 P=0.852	**0.71** **P=0.037**	CECR1	Adenosine deaminase, growth factor
CG2083	CG2083	CG2083	0.18 P=0.464	**0.71** **P=0.019**		
CG6084	DyakCG6084	CG6084	0.22 P=0.398	**0.71** **P=0.036**	AKR1B1	Glucose metabolism and osmoregulation
secretory Phospho-lipase A2	sPLA2	CG11124	0.18 P=0.546	**0.72** **P=0.006**	PLA2G3	
CG32368	CG32368	CG32368	0.04 P=0.899	**0.73** **P=0.026**		
CG9119	CG9119	CG9119	0.02 P=0.950	**0.73** **P=0.031**	c11orf54	
CG5793	CG5793	CG5793	0.22 P=0.094	**0.73** **P=0.004**	FAHD1	fumarylacetoacetate hydrolase, mitochondrial
Glutathione S transferase E9	GstE9	CG17534	0.21 P=0.096	**0.74** **P=0.022**		
Seleno-cysteine methyl-transferase	CG10621	CG10621	0.42 P=0.011	**0.74** **P=0.005**	MTR	Methionine synthase
CG5773	CG5773	CG5773	-0.12 P=0.548	**0.74** **P=0.037**		
CG5493	CG5493	CG5493	0.36 P=0.173	**0.75** **P=0.022**	CDO1	Oxidation of cysteine to sulfate
cellular repressor of E1A-stimulated genes	CREG	CG5413	0.40 P=0.401	**0.76** **P=0.049**	CREG1	Transcriptional repressor activity
CG3513	CG3513	CG3513	0.84 P=0.096	**0.76** **P=0.004**		
PGRP-SD	PGRP-SD	CG7496	0.14 P=0.810	**0.76** **P=0.044**		
CG33120	CG33120	CG33120	0.03 P=0.485	**0.77** **P=0.003**	DCI	beta oxidation of fatty acids in the mitochondria
CG4594	CG4594	CG4594	0.15 P=0.467	**0.78** **P=0.026**		
Iris	Iris	CG4715	0.60 P=0.221	**0.79** **P=0.012**		

Gene Title	Gene Symbol	Ensembl	log2(ratio) *elav>rCAG* to *elav>rCAA*	log2(ratio) *elav>rCUG* to *elav>rCAA*	Human orthologue	Function
Immune induced molecule 4	IM4	CG15231	0.40 P=0.066	**0.80** **P=0.004**		
CG3397	CG3397	CG3397	-0.05 P=0.919	**0.81** **P=0.042**		
Lipid storage droplet-1	Lsd-1	CG10374	0.41 P=0.262	**0.83** **P=0.040**		
CG11919	CG11919	CG11919	0.19 P=0.484	**0.83** **P=0.031**	PEX6	cytoplasmic ATPase, peroxisomal import protein
CG4716	CG4716	CG4716	0.49 P=0.017	**0.83** **P=4.74E-5**		
CG15281	CG15281	CG15281	0.14 P=0.612	**0.84** **P=0.011**		
CG33115	CG33115	CG33115	0.47 P=0.004	**0.84** **P=0.027**		
CG14629	CG14629	CG14629	0.35 P=0.217	**0.87** **P=4.50E-5**		
ornithine de-carboxylase	Odc1	CG8721	0.61 P=0.107	**0.88** **P=0.031**	ODC1	polyamine biosynthesis
Ance-4	Ance-4	CG8196	0.05 P=0.888	**0.88** **P=0.026**		Peptidyl dipeptidase
CG31777	CG31777	CG31777	0.48 P=0.074	**0.90** **P=0.007**		
CG6687	CG6687	CG6687	0.07 P=0.754	**0.91** **P=0.035**		
CG10031	CG10031	CG10031	0.50 P=0.027	**0.91** **P=0.044**		
CG15068	CG15068	CG15068	0.97 P=0.012	**0.91** **P=0.016**		
Cyp28d1	Cyp28d1	CG10833	0.23 P=0.153	**0.91** **P=0.017**	CYP3A43	
Vago	Vago	CG2081	0.30 P=0.084	**0.93** **P=0.035**		Immune response
GIP-like	Gip	CG2227	0.40 P=0.050	**0.93** **P=0.014**	HYI	Hydroxypyruvate isomerase
Imaginal disc growth factor1	Idgf1	CG4472	0.40 P=0.062	**0.93** **P=0.043**		Regulation of imaginal disc development
CG10026	CG10026	CG10026	0.52 P=0.011	**0.94** **P=0.026**	TTPA	Vitamin E metabolism, deficiency leads to cerebellar degeneration
CG18067	CG18067	CG18067	0.43 P=0.002	**0.95** **P=0.014**		
CG5288	CG5288	CG5288	0.45 P=0.246	**0.95** **P=0.033**	GALK2	Galactokinase, galactose metabolism
CG16873	CG16873	CG16873	0.56 P=0.058	**0.98** **P=0.034**	VWDE	
CG31370	CG31370	CG31370	0.46 P=0.027	**0.98** **P=2.51E-05**		
CG11395	CG11395	CG11395	0.53 P=0.051	**1.00** **P=0.021**	QRICH2	

Gene Title	Gene Symbol	Ensembl	log2(ratio) *elav>rCAG* to *elav>rCAA*	log2(ratio) *elav>rCUG* to *elav>rCAA*	Human orthologue	Function
hemolectin	Hml	CG7002	0.41 P=0.220	**1.02** **P=0.017**		
CG1092	CG1092	CG1092	0.44 P=0.046	**1.04** **P=0.049**		
CG34054	CG34054	CG34054	0.68 P=0.150	**1.04** **P=0.009**		
Sorbitol dehydrogenase like	Sodh-1	CG1982	0.33 P=0.547	**1.05** **P=0.031**	SORD	Sorbitol dehydrogenase
CG10131	CG10131	CG10131	0.43 P=0.190	**1.08** **P=0.025**	CRYL1	fatty acid metabolism, oxidoreductase activity
regucalcin	regucalcin	CG1803	0.22 P=0.698	**1.09** **P=0.025**	RGN	Calcium homeostasis
CG3699	CG3699	CG3699	0.41 P=0.346	**1.09** **P=0.009**	DHRS10	Short chain dehydrogenase reductase
CG33307	CG33307	CG33307	0.42 P=0.284	**1.15** **P=0.002**		
CG4306	CG4306	CG4306	0.82 P=0.036	**1.23** **P=0.020**	GGCT	glutathione homeostasis, release of cytochrome c from mitochondria
PGRP-SB1	PGRP-SB1	CG9681	0.18 P=0.671	**1.24** **P=0.008**		
CG13422	CG13422	CG13422	0.55 P=0.153	**1.26** **P=0.012**		
CG15293	CG15293	CG15293	0.71 P=0.032	**1.26** **P=0.018**		
CG4716	CG4716	CG4716	0.90 P=0.005	**1.33** **P=0.007**		
CG13086	CG13086	CG13086	0.89 P=0.039	**1.40** **P=0.024**		

Table B5: Genes altered in *elav>rCAG* flies compared to *elav>+* in microarray experiment 1. Selected for Log2(ratio) >0.5 or <-0.5, P<0.05.

Gene Title	Gene Symbol	Ensembl	log2(ratio) *elav>rCAG* to *elav>+*	log2(ratio) *elav>rCUG* to *elav>+*	Human orthologue	Function
CG4945	CG4945	CG4945	**-0.94** **P=0.027**	-0.75 P=0.095	SBK1	Kinase, mesoderm development
CG10625	CG10625	CG10625	**-0.92** **P=0.029**	-0.61 P=0.251		
CG15642	CG15642	CG15642	**-0.91** **P=0.027**	-0.34 P=0.243		
CG8837	CG8837	CG8837	**-0.86** **P=0.025**	-0.22 P=0.253		
CG10017	CG34340	CG10017	**-0.84** **P=0.045**	-0.41 P=0.063		Transcription factor, dendrite morphogenesis and muscle development
CG41136	CG41136	CG41136	**-0.81** **P=0.005**	-0.44 P=0.225		
Zinc/iron regulated transporter-related protein 3	Zip3	CG6898	**-0.81** **P=0.019**	-0.60 P=0.134	SLC39A2	zinc transporter
Glycogenin	Glycogenin	CG9480	**-0.78** **P=0.015**	-0.76 P=0.057	GYG1	glycogen synthesis
CG32850	CG32850	CG32850	**-0.78** **P=0.005**	-0.49 P=0.072	RNF11	induced by mutation to MEN proteins
CG15155	CG15155	CG15155	**-0.75** **P=0.031**	-0.35 P=0.274		
CG32506	CG32506	CG32506	**-0.74** **P=0.029**	-0.13 P=0.774	TRIM24	transcriptional co-repressor
CG31690	CG31690	CG31690	**-0.73** **P=0.045**	-0.77 P=0.088	TMTC2	
PGRP-SC2	PGRP-SC2	CG14745	**-0.71** **P=0.031**	-0.24 P=0.262	PGLYRP1	peptidoglycan recognition
CG9447	CG9447	CG9447	**-0.69** **P=0.019**	-0.30 P=0.265		
CG14569	CG14569	CG14569	**-0.69** **P=0.021**	-0.60 P=0.093		
CG15369	CG15369	CG15369	**-0.67** **P=0.047**	-0.80 P=0.019		
CG18213	CG18213	CG18213	**-0.67** **P=0.034**	-0.96 P=0.008		
CG13504	CG13504	CG13504	**-0.65** **P=0.023**	-0.41 P=0.220		
prickle-spiny legs	pk	CG11084	**-0.62** **P=0.049**	-0.80 P=0.096	PRICKLE2	neurite outgrowth
CG17754	CG17754	CG17754	**-0.62** **P=0.012**	-0.47 P=0.132	KLHL5	actin binding
molting defective	mld	CG34100	**-0.62** **P=0.040**	-0.28 P=0.065		ecdysone synthesis
CG3631	CG3631	CG3631	**-0.61** **P=0.037**	-0.34 P=0.364	FAM20B	Haematopoiesis
CG14258	CG14258	CG14258	**-0.60** **P=0.011**	-0.12 P=0.680		

Gene Title	Gene Symbol	Ensembl	log2(ratio) *elav>rCAG* to *elav>+*	log2(ratio) *elav>rCUG* to *elav>+*	Human orthologue	Function
tramtrack-69	ttk	CG1856	**-0.59** **P=0.045**	-0.43 P=0.115	ZNF499	neural and photoreceptor development, regulated by mef2
Inositol 1,4,5-triphosphate kinase 1	IP3K1	CG4026	**-0.56** **P=0.029**	-0.58 P=0.124		Release of Ca+ from ER, olfaction
CG8925	CG8925	CG8925	**-0.56** **P= CG8925**	-0.41 P=0.256		Carnitine transport
CG18249	CG18249	CG18249	**-0.56** **P=0.001**	-0.05 P=0.723		
Rab27	Rab27	CG14791	**-0.55** **P=0.009**	-0.41 P=0.204	RAB27A	vesicle transport
CG9967	CG9967	CG9967	**-0.54** **P=0.011**	-0.01 P=0.936		
Robo2	lea	CG5481	**-0.53** **P=0.049**	-0.55 P=0.053	ROBO1	axon guidance receptor
CG2083	CG2083	CG2083	**-0.53** **P=0.003**	0.00 P=0.944	KIAA1546	
CG16959	CG16959	CG16959	**-0.53** **P=0.043**	-0.51 P=0.233		Protective against DNA damage induced apoptosis
CG16772	CG16772	CG16772	**-0.53** **P=0.007**	-0.23 P=0.223		
iroquois	mirr	CG10601	**-0.53** **P=0.016**	-0.27 P=0.312	IRX	iroquois homeobox family member, regulation of rhodopsins, MAPK regulated
SP1029	SP1029	CG11956	**-0.52** **P=0.017**	-0.40 P=0.125	ANPEP	
Oregon-R glutamate decarboxylase	b	CG7811	**-0.52** **P=0.026**	-0.24 P=0.334	CSAD	taurine synthesis, excitatory pathways
CG9226	CG9226	CG9226	**-0.5** **P=0.012**	-0.18 P=0.158	WDR79	
Menin 1	Mnn1	CG13778	**0.54** **P=0.001**	0.11 P=0.681	MEN1	Inhibits JunD activity, transcriptional regulation
CG12116	CG12116	CG12116	**0.55** **P=0.039**	0.38 P=0.148		
ion transport peptide	itp	CG13586	**0.55** **P=0.047**	0.37 P=0.012		neuropeptide signalling
lethal (1) G0196	l(1)G0196	CG14616	**0.58** **P=0.009**	0.37 P=0.130	HISPPD2A	production of high energy pyrophosphates
embryonic lethal, abnormal vision	elav	CG4262	**0.60** **P=0.015**	0.46 P=0.022	HuR	RNA binding
CG31846	CG31846	CG31846	**0.60** **P=0.005**	0.64 P=0.041		
hu-li tai shao	hts	CG9325	**0.60** **P=0.002**	0.63 P=0.006	ADD1	Cytoskeletal protein, substrate for protein kinase A & C
beat-IIIa	beat-IIIa	CG12621	**0.61** **P=0.009**	0.33 P=0.222		Possible role in axon guidance
Ariadne-1	ari-1	CG5659	**0.61** **P=0.048**	0.40 P=0.019	ARIH1	Ubiquitin conjugating enzyme

Gene Title	Gene Symbol	Ensembl	log2(ratio) *elav>rCAG* to *elav>+*	log2(ratio) *elav>rCUG* to *elav>+*	Human orthologue	Function
CG32544	CG32544	CG32544	**0.64** **P=0.048**	0.52 P=0.190		
CG30151	CG30151	CG30151	**0.65** **P=0.001**	0.53 P=0.175		
metabotropic glutamate receptor	mGluRA	CG11144	**0.65** **P=0.018**	0.50 P=0.045	GRM3	Glutamate receptor, neuromuscular junction development
CG15216	CG15216	CG15216	**0.66** **P=0.003**	0.55 P=0.026		
disconnected	disco	CG9908	**0.68** **P=0.033**	0.40 P=0.188	BNC2	mRNA processing, regulated by SP1, brain development
Modifier67.2	mod(mdg4)	CG32491	**0.68** **P=0.008**	0.51 P=0.030		Regulation of apoptosis, regulation of chromatin assembly
CG2177	CG2177	CG2177	**0.72** **P=0.004**	0.57 P=0.082	SLC39A9	Zinc transporter
CG13229	CG13229	CG13229	**0.76** **P=0.021**	0.40 P=0.006		G-protein coupled receptor
CG11360	CG11360	CG11360	**0.78** **P=0.015**	0.51 P=0.136	MEX3A	RNA binding, colocalises with DCP1A and AGO1 in P bodies
CG13618	CG13618	CG13618	**0.84** **P=0.004**	0.65 P=0.009		
mindmelt	mbl	CG33197	**0.88** **P=0.022**	0.66 P=0.016	MBNL1	RNA splicing
CG13685	CG13685	CG13685	**0.90** **P=0.046**	0.64 P=0.066		
CG13783	Pvf3	CG13783	**1.20** **P=0.009**	0.86 P=0.213		Blood cell migration, CNS midline formation
Immune induced molecule 2	IM2	CG18106	**1.70** **P=0.034**	1.23 P=0.407		

Table B6: Genes altered in *elav>rCUG* flies compared to *elav>+* in microarray experiment 1. Selected for Log2(ratio) >0.5 or <-0.5, P<0.05.

Gene Title	Gene Symbol	Ensembl	log2(ratio) *elav>rCAG* to *elav>+*	log2(ratio) *elav>rCUG* to *elav>+*	Human orthologue	Function
CG18213	CG18213	CG18213	-0.67 P=0.034	**-0.96** **P=0.008**		
beta-glycoprotein-hormone-related-polypeptide	beta-glycoprotein-hormone-related	CG40041	-0.32 P=0.0226	**-0.89** **P=0.026**		
sparkling	sv	CG11049	-0.59 P=0.080	**-0.84** **P=0.038**	PAX5	B cell specific transcription factor/ midbrain dopaminergic neuron specification
CG31712	CG31712	CG31712	-0.71 P=0.051	**-0.81** **P=0.033**		
CG13675	CG13675	CG13675	-0.13 P=0.214	**-0.81** **P=0.034**		
CG15369	CG15369	CG15369	-0.67 P=0.047	**-0.80** **P=0.019**		Cysteine-type endopeptidase
goliath	gol	CG2679	-0.47 P=0.268	**-0.77** **P=0.035**	RNF150	Transcriptional regulation, formation of mesoderm
defective pro-ventriculus	dve	CG5799	-0.16 P=0.605	**-0.73** **P=0.028**		Transcription factor activity
CG15760	CG15760	CG15760	-0.39 P=0.373	**-0.69** **P=0.022**		
no distributive disjunction	nod	CG1763	-0.27 P=0.090	**-0.67** **P=0.017**		Regulation of meiotic cell cycle
CG32169	CG32169	CG32169	-0.73 P=0.161	**-0.67** **P=0.007**	MSI2	RNA binding, post-transcriptional gene regulation
CG12641	CG12641	CG12641	-0.55 P=0.133	**-0.66** **P=0.008**		
CG17265	CG17265	CG17265	-0.63 P=0.084	**-0.60** **P=0.040**	CCDC85	Putative component of RNAi pathways
Ady43A	Ady43A	CG1851	-0.23 P=0.508	**-0.60** **P=0.010**		Adenosine kinase activity
CG8273	CG8273	CG8273	-0.64 P=0.073	**-0.57** **P=0.019**	SON	Double stranded RNA binding, DNA binding
CG5514	CG5514	CG5514	0.00 P=0.990	**-0.57** **P=0.003**	FAM44A	Structural component of cell wall
CG33331	CG33331	CG33331	-0.27 P=0.056	**-0.57** **P=0.036**	C3orf31	
Mediator complex subunit 24	MED24	CG7999	-0.22 P=0.389	**-0.56** **P=0.039**	MED24	component of mediator complex, transcriptional co-activator
male specifc lethal	msl-2	CG3241	-0.29 P=0.001	**-0.55** **P=0.002**	MSL2L1	H4 histone acetylation
will die slowly	wds	CG17437	-0.50 P=0.231	**-0.55** **P=0.012**	WDR5	
CG8026	CG8026	CG8026	-0.45 P=0.131	**-0.55** **P=0.024**	SLC25A32	folate shuttle, cytoplasm to mitochondria

Gene Title	Gene Symbol	Ensembl	log2(ratio) *elav>rCAG* to *elav>+*	log2(ratio) *elav>rCUG* to *elav>+*	Human orthologue	Function
U26	U26	CG13401	-0.43 P=0.229	**-0.54** **P=0.042**	AASDH	2-aminoadipic 6-semialdehyde dehydrogenase, fatty acid metabolism
clift	eya	CG9554	-0.25 P=0.221	**-0.53** **P=0.007**	EYA1	Transcriptional regulation
CG33174	CG33174	CG33174	-0.33 P=0.102	**-0.53** **P=0.027**	DAGLA	Associated with SCA20, contributes to purkinje cell synapse formation
CG9977	CG9977	CG9977	-0.15 P=0.380	**-0.53** **P=0.039**	AHCYL2	Adenosyl-homocysteinase activity
CG4984	CG4984	CG4984	-0.32 P=0.077	**-0.53** **P=0.009**	CACNG7	Voltage-gated calcium channel
longitudinals absent	lola	CG12052	-0.33 P=0.332	**-0.53** **P=0.028**	ZBTB3	Transcription factor activity
Suppressor of cytokine signaling at 36E	Socs36E	CG15154	-0.43 P=0.028	**-0.52** **P=0.024**	SOCS1	suppressor of cytokine signaling, JAK-STAT cascade
CG13594	CG13594	CG13594	-0.25 P=0.100	**-0.52** **P=0.042**		
deep orange	dor	CG3093	-0.37 P=0.054	**-0.52** **P=0.038**	VPS18	vesicle trafficking
CG17793	CG34402	CG4940	-0.33 P=0.119	**-0.52** **P=0.042**		
xl6	xl6	CG10203	-0.15 P=0.319	**-0.52** **P=0.022**	SFRS7	arginine/serine rich splicing factor
metabotropic GABA-B receptor subtype 3	GABA-B-R3	CG3022	-0.27 P=0.131	**-0.51** **P=0.045**		GABA receptor
transcript-near-decapent-aplegic	oaf	CG9884	-0.32 P=0.015	**-0.50** **P=0.002**	OAF	Nervous system development
CG13845	CG34376	CG13845	0.24 P=0.272	**0.50** **P=0.018**		
metabotropic glutamate receptor	mGluRA	CG11144	0.65 P=0.018	**0.50** **P=0.045**	GRM3	Glutamate receptor
Modifier67.2	mod(mdg4)	CG32491	0.68 P=0.008	**0.51** **P=0.028**		Regulation of apoptosis, regulation of chromatin assembly
CG17761	CG17761	CG17761	0.35 P=0.142	**0.51** **P=0.045**		
CG10399	CG10399	CG10399	0.03 P=0.857	**0.51** **P=0.026**	HMGCL	Hydroxymethyl-glutaryl-CoA lyase, mitochondrial precursor
Cyp12c1	Cyp12c1	CG4120	0.00 P=0.990	**0.51** **P=2.67E-5**	CYP24A1	Electron carrier
CG3246	CG3246	CG3246	-0.01 P=0.983	**0.51** **P=0.040**		
CG7233	CG7233	CG7233	0.28 P=0.228	**0.53** **P=0.040**	SKI	TGFB1 signaling pathway
Cyp6a14	Cyp6a14	CG8687	0.03 P=0.900	**0.53** **P=0.010**		Electron carrier

Gene Title	Gene Symbol	Ensembl	log2(ratio) *elav>rCAG* to *elav>+*	log2(ratio) *elav>rCUG* to *elav>+*	Human orthologue	Function
Turandot X	TotX	CG31193	0.42 P=0.446	**0.54** **P=0.043**		Immune response
lethal (3) neo38	l(3)neo38	---	0.77 P=0.082	**0.55** **P=0.018**		Nucleic acid binding
CG4716	CG4716	CG4716	0.21 P=0.249	**0.55** **P=0.009**		Methylenetetra-hydro-folate dehydrogenase
CG15216	CG15216	CG15216	0.66 P=0.003	**0.55** **P=0.026**		
CG31295	CG31295	CG31295	-0.10 P=0.668	**0.56** **P=0.048**		
Glutathione S transferase D9	GstD9	CG10091	0.21 P=0.428	**0.57** **P=0.050**		Glutathione transferase
Meltrin-like	mmd	CG9163	0.46 P=0.117	**0.58** **P=0.009**	ADAM9	binds mitotic arrest deficient 2 beta protein
CG6364	CG6364	CG6364	0.37 P=0.070	**0.58** **P=0.045**	UCK2	pyrimidine nucleoside triphosphate production
CG3534	CG3534	CG3534	0.22 P=0.278	**0.58** **P=0.027**	XYLB	
CG33472	CG33472	CG33472	1.17 P=0.066	**0.60** **P=0.043**		
CG16965	CG16965	CG16965	0.03 P=0.931	**0.61** **P=0.028**	ATHL1	
jafrac2	Jafrac2	CG1274	0.36 P=0.195	**0.61** **P=0.036**	PRDX4	activation of NF-kappaB
CG10467	CG10467	CG10467	0.21 P=0.216	**0.62** **P=0.014**	GALM	galactose metabolism
hu-li tai shao	hts	CG9325	0.60 P=0.002	**0.63** **P=0.006**	ADD1	Cytoskeletal protein, substrate for protein kinase A & C
CG31846	CG31846	CG31846	0.60 P=0.005	**0.64** **P=0.041**		
CG14629	CG14629	CG14629	0.13 P=0.715	**0.65** **P=0.024**		
CG16890	CG16890	CG16890	0.35 P=0.175	**0.65** **P=0.044**	c10orf4	FRA10AC spanning gene
CG13618	CG13618	CG13618	0.84 P=0.004	**0.65** **P=0.009**		
mindmelt	mbl	CG33197	0.88 P=0.022	**0.66** **P=0.016**	MBNL1	RNA splicing
CG5793	CG5793	CG5793	0.18 P=0.207	**0.69** **P=0.019**	FAHD1	
Chrac-16	Chrac-16	CG15736	0.29 P=0.032	**0.70** **P=0.001**	CHRAC1	Chromatin accessibility protein
CG7900	CG7900	CG7900	0.14 P=0.774	**0.73** **P=0.025**	FAAH	hydrolysis of primary and secondary amides
CG32387	CG32387	CG32387	0.75 P=0.166	**0.76** **P=0.024**	DSCAM	Down syndrome cell adhesion molecule, axon guidance
CG31370	CG31370	CG31370	0.25 P=0.313	**0.78** **P=0.008**		

Gene Title	Gene Symbol	Ensembl	log2(ratio) *elav>rCAG* to *elav>+*	log2(ratio) *elav>rCUG* to *elav>+*	Human orthologue	Function
neuropeptide -like precursor 4	Nplp4	CG15361	0.10 P=0.692	**0.82** **P=0.036**		Neuropeptide signaling
hemolectin	Hml	CG7002	0.44 P=0.288	**1.05** **P=0.047**	VWF	antihemophilic factor carrier and a platelet-vessel wall mediator
PGRP-SB1	PGRP-SB1	CG9681	0.10 P=0.830	**1.16** **P=0.023**		Immune response
CG9080	CG9080	CG9080	0.44 P=0.704	**1.20** **P=0.012**		
CG3777	CG3777	CG3777	0.78 P=0.115	**1.23** **P=0.021**		

Table B7: Genes altered in *elav>rCAG* flies compared to *elav>rCAA* in microarray experiment 2. Selected for Log2(ratio) >0.5 or <-0.5, P<0.05.

Gene Title	Gene Symbol	Ensembl	log2(ratio) *elav>rCAG* to *elav>rCAA*	log2(ratio) *elav>rCUG* to *elav>rCAA*	Human Orthologue	Function
CG3823	CG3823	CG3823	**-0.95** **P=0.006**	0.16 P=0.298		
CG15345	CG15345	CG15345	**-0.69** **P=0.004**	-0.18 P=0.295		
fuzzy onions	fzo	CG4568	**-0.67** **P=0.033**	-0.30 P=0.292		mitochondrial fusion during spermatogenesis
oocyte maintenance defects	omd	CG9591	**-0.66** **P=0.013**	-0.40 P=0.163	INTS5	
Staufen	stau	CG5753	**-0.66** **P=0.029**	-0.57 P=0.081	STAU2	double-stranded RNA binding, RNA localisation involved in cell fate determination
CG10185	CG10185	CG10185	**-0.64** **P=0.014**	-0.47 P=0.019		
***Drosophila* insulin-like peptide 5**	Ilp5	CG33273	**-0.59** **P=0.001**	-0.44 P=0.007		Insulin signalling
thoc6	thoc6	CG5632	**-0.58** **P=0.033**	-0.12 P=0.136	THOC6	mRNA export from nucleus
CG32736	CG32736	CG32736	**-0.51** **P=025**	-0.34 P=0.129		
CG9079	Cpr47Ea	CG9079	**0.51** **P=0.022**	1.34 P=0.003		
CG10211	CG10211	CG10211	**0.51** **P=0.002**	0.15 P=0.620		oxidative stress response
katanin 80	kat80	CG13956	**0.51** **P=0.012**	0.36 P=0.293	KATNB1	microtubule binding/severing
vinculin	Vinc	CG3299	**0.51** **P=0.032**	-0.02 P=0.924	VCL	actin cytoskeleton, cell adhesion, phagocytosis
laminin alpha1,2	wb	CG15288	**0.52** **P=0.021**	-0.14 P=0.382	LAMA1	actin cytoskeleton, cell adhesion
drumstick	drm	CG10016	**0.52** **P=0.025**	-0.07 P=0.856		patterning of gut
engrailed	en	CG9015	**0.54** **P=0.034**	0.16 P=0.615	EN1	homeobox gene, segmentation patterning
Micro-cephalin	MCPH1	CG42572	**0.54** **P=0.013**	0.16 P=0.671		mushroom body development
Juvenile hormone-inducible protein 1	Jhl-1	CG3298	**0.54** **P=0.006**	0.23 P=0.272	ELAC2	pre-tRNA processing
CG31886	CG31886	CG31886	**0.55** **P=0.015**	0.26 P=0.024		
CG6847	CG6847	CG6847	**0.56** **P=0.004**	0.08 P=0.262	LIPC	triacylglcerol lipase
CG32626	CG32626	CG32626	**0.56** **P=0.044**	0.39 P=0.129	AMPD2	AMP deaminase
hamlet	ham	CG31753	**0.57** **P=0.030**	0.43 P=0.220	EVI1	regulation of dendrite morphogenesis, promotes single dendrite morphogenesis

Gene Title	Gene Symbol	Ensembl	log2(ratio) *elav>rCAG* to *elav>rCAA*	log2(ratio) *elav>rCUG* to *elav>rCAA*	Human Orthologue	Function
CG34104	CG34104	CG34104	**0.58** **P=2.87E-05**	0.80 P=**0.004**		Signal transduction, GTPase activity
Peptide O-xylosyltransferase	oxt	CG32300	**0.58** **P=0.038**	0.19 P=0.328	EXYLT1	xylosyltransferase
Odorant-binding protein 56f	Obp56f	CG30450	**0.58** **P=0.001**	0.27 P=0.580		Sensory perception of chemical stimulus
CG1397	CG1397	CG1397	**0.59** **P=0.028**	0.53 P=0.051		
CG33528	CG33528	CG33528	**0.60** **P=0.042**	0.02 P=0.911	SLC18A2	Neurotransmitter secretion, synaptic vesicle amine transport
CG13277	CG13277	CG13277	**0.60** **P=0.045**	0.44 P=0.121	LSM7	pre-mRNA processing
gliotactin	Gli	CG3903	**0.60** **P=0.040**	0.22 P=0.295		septate junction formation, role in polarisation of cells
N-like	CG7447	CG7447	**0.61** **P=0.030**	0.17 P=0.559	EGFL8	
Matrix metallo-proteinase 1	Mmp1	CG4859	**0.62** **P=0.041**	0.57 P=0.160	MMP14	Role in ECM regulation, cell adhesion
CG31781	CG31781	CG31781	**0.62** **P=0.011**	-0.08 P=0.777	LIPF-002	triacylglcerol lipase
CG12877	CG12877	CG12877	**0.62** **P=0.030**	0.56 P=0.428	REXO1	Transcriptional elongation, RNA exonuclease activity
CG4525	CG4525	CG4525	**0.63** **P=0.044**	0.47 P=0.122	TTC26	cilium assembly
D-frizzled2	fz2	CG9739	**0.64** P=**0.028**	-0.21 P=0.331	FZD8	Wnt receptor signalling, axon extension, found on postsynaptic motor neurons, requires Grip to be endocytosed and elicit transcriptional changes.
CG13065	CG13065	CG13065	**0.64** **P=0.046**	0.55 P=0.158		
cut	ct	CG11387	**0.65** **P=0.025**	0.22 P=0.523	CUTL1	transcriptional regulation, regulation of dendrite morphogenesis
CG10632	CG10632	CG10632	**0.67** **P=0.041**	0.45 P=0.011		
CG7296	CG7296	CG7296	**0.67** **P=0.028**	0.71 P=0.358		
CG7714	CG7714	CG7714	**0.68** **P=0.034**	0.41 P=0.227		
dynein	ctp	CG6998	**0.70** **P=0.001**	0.55 P=0.002	DYNLL2	microtubule-based movement, required for proper axon guidance of sensory neurons.
CG14528	CG14528	CG14528	**0.72** **P=0.048**	0.84 P=0.015		
CG13062	CG13062	CG13062	**0.72** **P=0.025**	0.22 P=0.050		
CG7516	CG7516	CG7516	**0.73** **P=0.021**	0.03 P=0.827	NOL10	Putative nucleolar protein

Gene Title	Gene Symbol	Ensembl	log2(ratio) *elav>rCAG* to *elav>rCAA*	log2(ratio) *elav>rCUG* to *elav>rCAA*	Human Orthologue	Function
septin	Sep5	CG2916	**0.73** **P=0.031**	0.24 P=0.421	SEPT11	Filament forming cytoskeletal GTPase, role in vesicle transport cytokinesis
CG3099	CG3099	CG3099	**0.79** **P=0.024**	0.20 P=0.481	HECW2	Ubiquitin protein ligase
yolk protein	Yp3	CG11129	**0.79** **P=0.022**	-0.16 P=0.557		Fat body protein, lipase?
CG32850	CG32850	CG32850	**0.80** **P=0.017**	-0.03 P=0.943	RNF11	
CG18641	CG18641	CG18641	**0.81** **P=0.007**	0.67 P=0.031		lipase activity
CG4078	CG4078	CG4078	**0.82** **P=0.004**	0.48 P=0.021	RTEL1	nucleotide excision repair
G protein gamma30A	Ggamma30A	CG3694	**0.83** **P=0.016**	0.20 P=0.093	GNG13	phototransduction
CG9400	CG9400	CG9400	**0.84** **P=0.043**	0.97 P=0.014		
CG7744	CG7744	CG7744	**0.87** **P=0.017**	0.80 P=0.148		
CG16752	CG16752	CG16752	**0.88** **P=0.024**	-0.21 P=0.512		neuropeptide receptor activity
Ret oncogene	Ret	CG14396	**0.89** **P=0.032**	0.05 P=0.928	RET	homophilic cell adhesion
Inscuteable	insc	CG11312	**0.89** **P=0.025**	0.55 P=0.050	INSC	cytoskeletal adaptor, protein and RNA localisation, localisation is dynein dependent
CG12998	CG12998	CG12998	**0.92** **P=0.036**	0.71 P=0.023		
echinoid	ed	CG12676	**1.05** **P=0.034**	0.18 P=0.574		actin cytoskeleton/cell adhesion, may form a signalling complex with Grip, important in myogenesis
CG4270	CG4270	CG4270	**1.32** **P=0.015**	1.00 P=0.309		

Table B8: Genes altered in *elav>rCUG* flies compared to *elav>rCAA* in experiment 2.
Selected for Log2(ratio) >0.5 or <-0.5, P<0.05.

Gene Title	Gene Symbol	Ensembl	log2(ratio) *elav>rCUG* to *elav>rCAA*	log2(ratio) *elav>rCAG* to *elav>rCAA*	Human Orthologue	Function
CG4161	CG4161	CG4161	**-1.34** **P=0.010**	-0.36 P=0.223		
CG10943	CG10943	CG10943	**-1.21** **P=0.004**	-0.38 P=0.194		
defensin	Def	CG1385	**-1.04** **P=0.041**	-0.83 P=0.123		immune response
defective transmitter release	dtr	CG31623	**-1.02** **P=0.030**	-0.22 P=0.152		synaptic transmission
Cyp6a19	Cyp6a19	CG10243	**-1.01** **P=0.010**	-0.20 P=0.414	CYP3A7	electron carrier
CG6416	CG6416	CG6416	**-0.92** **P=0.027**	-0.06 P=0.379		mesoderm development (Zasp66) - cytoskeletal remodelling
CG33798	CG33798	CG33798	**-0.91** **P=0.007**	0.04 P=0.393		
Complement-ation group C	Mef2	CG1429	**-0.90** **P=0.007**	0.00 P=0.495	MEF2C	transcription factor, muscle development
Cyp6a17	Cyp6a17	CG10241	**-0.88** **P=0.040**	-0.35 P=0.275	CYP3A5	electron carrier
CG18549	CG18549	CG18549	**-0.87** **P=0.005**	0.06 P=0.159		
CG9817	CG9817	CG9817	**-0.82** **P=0.023**	-0.36 P=0.116		
Grip71	Grip71	CG10346	**-0.82** **P=0.001**	-0.34 P=0.012		gamma-tubulin binding, cell-cycle regulation
CG14034	CG14034	CG14034	**-0.78** **P=0.024**	0.05 P=0.397		phospholipase activity, lipid metabolism
CG17177	CG17177	CG17177	**-0.78** **P=0.048**	-1.51 P=0.078		
munin	mun	CG42625	**-0.76** **P=0.034**	-0.33 P=0.270		glial-derived neurotrophic factor
CG32238	CG32238	CG32238	**-0.75** **P=0.024**	0.19 P=0.173	TTLL1	Tubulin tyrosine ligase, essential post-translational modification for normal brain and muscle development
CG6425	CG6425	CG6425	**-0.73** **P=0.035**	-0.47 P=0.168		
CG31365	CG31365	CG31365	**-0.72** **P=0.028**	-0.43 P=0.040		
don juan	dj	CG1980	**-0.71** **P=0.037**	-0.45 P=0.032		associated with mitochondria in flagella during spermatogenesis
CG1961	CG1961	CG1961	**-0.68** **P=0.040**	-0.23 P=0.359		nucleotidase
CG13416	CG13416	CG13416	**-0.67** **P=0.031**	-0.25 P=0.285		
CG12425	CG34354	CG34354	**-0.66** **P=0.035**	-0.09 P=0.369		RNA binding protein

Gene Title	Gene Symbol	Ensembl	log2(ratio) *elav>rCUG* to *elav>rCAA*	log2(ratio) *elav>rCAG* to *elav>rCAA*	Human Orthologue	Function
CG17290	CG17290	CG17290	**-0.66** **P=0.031**	-0.22 P=0.128		
Dalpha1	nAcRalpha-96Aa	CG5610	**-0.64** **P=0.010**	-0.19 P=0.277	CHRNA4	Acetylcholine receptor subunit, neurotransmitter signalling, insulin signalling pathway
CG10320	CG10320	CG10320	**-0.63** **P=0.008**	0.02 P=0.421	NDUFB3	NADH dehydrogenase, electron transport chain, RNA import into nucleus
CG9200	Atac1	CG9200	**-0.62** **P=0.007**	0.20 P=0.257	ZZZ3	histone acetylation
CG8157	CG8157	CG8157	**-0.61** **P=0.024**	-0.22 P=0.075		
CG13300	CG13300	CG13300	**-0.59** **P=0.014**	-0.25 P=0.019		
CG8329	CG8329	CG8329	**-0.59** **P=0.016**	-0.43 P=0.041		protease
DnaJ-like-60	DnaJ-60	CG42567	**-0.59** **P=0.035**	-0.06 P=0.382	DNAJC4	molecular chaperone
CG14803	CG14803	CG14803	**-0.59** **P=0.033**	-0.52 P=0.076		
GAGA factor	Trl	CG33261	**-0.58** **P=0.034**	0.21 P=0.280	BTBD14B	positive regulation of transcription, interacts with mod(mdg4)
CG7065	CG7065	CG7065	**-0.57** **P=0.024**	-0.16 P=0.195		
CG33054	CG33054	CG33054	**-0.56** **P=0.009**	-0.51 P=0.049		
CG12912	CG12912	CG12912	**-0.56** **P=0.001**	0.02 P=0.455		
CG10738	CG10738	CG10738	**-0.56** **P=0.017**	0.12 P=0.275		cell signalling pathway
DNA polymerase epsilon	DNApol-epsilon	CG6768	**-0.55** **P=0.007**	0.14 P=0.273	POLE	DNA-dependent DNA polymerase
skpB	skpB	CG8881	**-0.54** **P=0.041**	-0.05 P=0.428	SKP1A	ubiqutin-dependent protein catabolism, cell cycle?
lethal (1) G0020	l(1)G0020	CG1994	**-0.54** **P=0.005**	-0.59 P=0.082	NAT10	N-acetyltransferase
nmdyn-D6	nmdyn-D6	CG5310	**-0.53** **P=0.037**	-0.16 P=0.055	NME6	nucleoside-diphosphate kinase
suppressor (rdgB) 69	su(rdgB)69	-	**-0.52** **P=0.010**	-0.35 P=0.074		phototransduction, interacts with phospholipase C, mutants are slow to terminate phototransduction
sepia	se	CG6781	**-0.52** **P=0.045**	0.00 P=0.492	GSTO1	glutathione dehydrogenase activity, eye pigment biosynthesis
Psf3	Psf3	CG2222	**-0.52** **P=0.027**	-0.45 P=0.023	GINS3	DNA helicase activity
transcription unit	Cyp4ae1	CG10755	**-0.52** **P=0.009**	-0.06 P=0.389		electron carrier
CG32108	CG32108	CG32108	**-0.52** **P=0.047**	-0.23 P=0.229	GTF3C2	putative transcription factor

Gene Title	Gene Symbol	Ensembl	log2(ratio) *elav>rCUG* to *elav>rCAA*	log2(ratio) *elav>rCAG* to *elav>rCAA*	Human Orthologue	Function
Plum	bw	CG17632	**-0.51** **P=0.018**	-0.44 P=0.008		eye pigment precursor transport activity
CG4290	CG4290	CG4290	**-0.50** **P=0.003**	-0.17 P=0.053	SNF1LK2	protein kinase
CG13397	CG13397	CG13397	**0.50** **P=0.016**	0.29 P=0.015	NAGLU	alpha-N-acetylglucosaminidase, orthologue of MPS IIIB gene
CG1969	CG1969	CG1969	**0.50** **P=0.005**	0.20 P=0.142	GNPNAT1	glucosamine 6-phosphate N acetyltransferase activity, energy metabolism
CG10031	CG10031	CG10031	**0.50** **P=0.015**	0.11 P=0.082		peptidase activity
PioPio	pio	CG3541	**0.51** **P=0.029**	0.20 P=0.095		cell adhesion and microtubule organisation
CG13116	CG13116	CG13116	**0.51** **P=0.020**	-0.38 P=0.069		
CG15006	Cpr64Aa	CG15006	**0.52** **P=0.027**	0.49 P=0.108		cuticle protein
CG10208	CG10208	CG10208	**0.52** **P=0.010**	0.20 P=0.200		
CG10175	CG10175	CG10175	**0.52** **P=0.013**	0.42 P=0.226	CES2	carboxyl esterase
Laminin	LanB1	CG7123	**0.52** **P=0.041**	0.19 P=0.079	LAMB1	sub-unit of laminin, promoter of neurite outgrowth
CG30190	CG30190	CG30190	**0.52** **P=0.002**	0.26 P=0.079		
Melanization Protease 1	MP1	CG1102	**0.52** **P=0.013**	0.06 P=0.345		immune response
CG5224	CG5224	CG5224	**0.53** **P=0.013**	0.10 P=0.185		glutathione transferase
CG4576	CG4576	CG4576	**0.53** **P=0.042**	0.42 P=0.113		acyl transferase
CG5167	CG5167	CG5167	**0.54** **P=0.045**	-0.07 P=0.362		
CG5618	CG5618	CG5618	**0.54** **P=0.009**	0.11 P=0.121		role in cardiogenesis?
CG30380	CG30380	CG30380	**0.54** **P=0.001**	0.16 P=0.253		
CG32641 /// **CG32640**	CG32640 /// CG32641	CG32640/// CG32641	**0.55** **P=0.048**	1.00 P=0.171		
Inscuteable	insc	CG11312	**0.55** **P=0.050**	0.89 P=0.013	INSC	cytoskeletal adaptor, protein and RNA localisation, localisation is dynein dependent
CG18249	CG18249	CG18249	**0.55** **P=0.026**	0.08 P=0.380		
dynein	ctp	CG6998	**0.55** **P=0.002**	0.70 P=0.001	DYNLL2	microtubule-based movement, required for proper axon guidance of sensory neurons.
CG6415	CG6415	CG6415	**0.55** **P=0.039**	0.09 P=0.323	AMT	Aminomethyl-transferase

Gene Title	Gene Symbol	Ensembl	log2(ratio) *elav>rCUG* to *elav>rCAA*	log2(ratio) *elav>rCAG* to *elav>rCAA*	Human Orthologue	Function
CG11400	CG11400	CG11400	**0.55** **P=0.017**	0.47 P=0.002		
CG7144	CG7144	CG7144	**0.56** **P=0.004**	0.41 P=0.137	AASS	lysine ketoglutarate reductase, regulation of histone modification
CG13877	CG13877	CG13877	**0.56** **P=0.008**	-0.05 P=0.195		
BM-40-SPARC	BM-40-SPARC	CG6378	**0.57** **P=0.049**	0.02 P=0.448	SPARCL1	mesoderm development
CG33494	CG33494	CG33494	**0.57** **P=0.005**	-0.10 P=0.323		
cytosolic aconitase	Irp-1B	CG6342	**0.58** **P=0.008**	0.12 P=0.118	ACO1	iron homeostasis
CG6340	CG6340	CG6340	**0.58** **P=0.036**	0.50 P=0.242		
licorne	lic	CG12244	**0.58** **P=0.042**	0.07 P=0.325	MAP2K3	MAPK activity, involved in oocyte polarity
CG11550	CG11550	CG11550	**0.59** **P=0.045**	0.08 P=0.265		
CG7154	CG7154	CG7154	**0.59** **P=0.010**	0.81 P=0.043	BRD7	
Glycyl-tRNA synthetase	Aats-gly	CG6778	**0.59** **P=0.043**	0.20 P=0.326	GARS	glycyl-tRNA synthetase
CG12560	CG12560	CG12560	**0.60** **P=0.023**	-0.08 P=0.359		N-acetyltransferase
DNA ligase III	lig3	CG17227	**0.60** **P=0.010**	-0.12 P=0.191	LIG3	DNA repair
CG17189	CG17189	CG17189	**0.60** **P=0.041**	0.23 P=0.157		
CG7830	CG7830	CG7830	**0.60** **P=0.030**	0.18 P=0.024	TUSC3	
CG11378	CG11378	CG11378	**0.60** **P=0.044**	-0.63 P=0.243		
Olfactory-specific 9	Os9	CG18806	**0.60** **P=0.044**	0.14 P=0.084		possible immune function
CG9657	CG9657	CG9657	**0.61** **P=0.035**	0.18 P=0.035	SLC5A12	sodium:solute symporter
decapo	dap	CG1772	**0.61** **P=0.025**	0.42 P=0.023		cell cycle control, interacts with Dicer1 and spen
CG18302	CG18302	CG18302	**0.61** **P=0.007**	0.06 P=0.268	LIPA/LIPF	lipid metabolism
Ciao1	Ciao1	CG12797	**0.61** **P=0.006**	0.32 P=0.181	CIAO1	orthologue of S. cerevisiae cytosolic iron-sulfur protein assembly
CG17928	CG17928	CG17928	**0.62** **P=0.013**	0.29 P=0.145		
CG5853	CG5853	CG5853	**0.62** **P=0.037**	0.20 P=0.183		ATP coupled transporter, phagocytosis
synaptobrevin	Syb	CG12210	**0.63** **P=0.025**	0.37 P=0.087	VAMP1	neurotransmitter secretion
CG30154	CG30154	CG30154	**0.64** **P=0.002**	0.27 P=0.023		

Gene Title	Gene Symbol	Ensembl	log2(ratio) *elav>rCUG* to *elav>rCAA*	log2(ratio) *elav>rCAG* to *elav>rCAA*	Human Orthologue	Function
CG13222	Cpr47Ee	CG13222	**0.64** **P=0.008**	0.25 P=0.044		Component of cuticle
yippee interacting protein 2	yip2	CG4600	**0.64** **P=0.015**	-0.03 P=0.440	ACAA2	lipid metabolism, fatty acid beta oxidation, mitochondrial
CG12262	CG12262	CG12262	**0.64** **P=0.027**	0.06 P=0.292	ACADM	lipid metabolism, fatty acid beta oxidation, mitochondrial
CG15414	CG15414	CG15414	**0.64** **P=0.007**	0.51 P=0.130		
CG32667	CG32667	CG32667	**0.64** **P=0.047**	0.33 P=0.263		
CG32499	CG32499	CG32499	**0.65** **P=0.030**	0.10 P=0.306		chitin metabolism
CG33054	CG33054	CG33054	**0.66** **P=0.040**	0.27 P=0.055		
cln3	cln3	CG5582	**0.66** **P=2.58E-5**	0.50 P=0.062	CLN3	human gene deficiencies show neuronal degeneration
CG12656	CG12656	CG12656	**0.67** **P=0.026**	-0.10 P=0.361		
CG15093	CG15093	CG15093	**0.67** **P=0.008**	0.03 P=0.245	HIBADH	3-hydroxyisobutarate dehydrogenase
CG32032	CG32032	CG32032	**0.67** **P=0.023**	0.11 P=0.032		
CG18641	CG18641	CG18641	**0.67** **P=0.031**	0.81 P=0.003		lipase
crammer	cer	CG10460	**0.68** **P=0.013**	0.37 P=0.072		cysteine-type endopeptidase inhibitor, inhibits cathepsins, involved in long-term memory formation
CG15825	fon	CG15825	**0.68** **P=0.010**	-0.17 P=0.031		hemolymph clotting
capa receptor	capaR	CG14575	**0.68** **P=0.010**	0.39 P=0.112	NMUR2	neuropeptide receptor
CG1537	CG1537	CG1537	**0.69** **P=0.022**	-0.28 P=0.195		
Glutathione S transferase E7	GstE7	CG17531	**0.69** **P=0.003**	0.21 P=0.148		glutathione transferase
***Drosophila* insulin-like peptide 6**	Ilp6	CG14049	**0.69** **P=2.50E-05**	0.11 P=0.199		insulin receptor
fragment K	alpha-Est10	CG1131	**0.69** **P=0.041**	0.39 P=0.028	CES2	carboxyl esterase
CG17333	CG17333	CG17333	**0.69** **P=0.010**	0.27 P=0.029	PGLS	6-phosphogluconolactonase activity
Limpet	Lmpt	CG32171	**0.69** **P=0.019**	0.56 P=0.138	FHL2	transcription factor, heart development
Flavin-containing mono-oxygenase 1	Fmo-1	CG3006	**0.70** **P=0.028**	0.25 P=0.160	FMO6/FMO5	oxidative stress response, possible neurological role
CG33281	CG33281	CG33281	**0.70** **P=0.048**	0.58 P=0.030		monosaccharide transporter

Gene Title	Gene Symbol	Ensembl	log2(ratio) *elav>rCUG* to *elav>rCAA*	log2(ratio) *elav>rCAG* to *elav>rCAA*	Human Orthologue	Function
CG9691	CG9691	CG9691	**0.7** **P=0.018**	0.26 P=0.002		
CG12998	CG12998	CG12998	**0.71** **P=0.023**	0.92 P=0.018		
CG7442	CG7442	CG7442	**0.72** **P=0.023**	0.12 P=0.230		
CG18107	CG18107	CG18107	**0.73** **P=0.016**	-0.04 P=0.475		
serpin 1	Spn1	CG9456	**0.73** **P=0.031**	0.23 P=0.041	SERPINB4	serine protease inhibitor
CG6723	CG6723	CG6723	**0.73** **P=0.041**	0.13 P=0.388		sodium:solute symporter
glutactin	Glt	CG9280	**0.74** **P=0.034**	0.07 P=0.351		glycoprotein
GH06348	CG1516	CG1516	**0.74** **P=0.032**	0.27 P=0.119	PC	Gluconeogenesis, pyruvate carboxylase
Transferrin 3	Tsf3	CG3666	**0.75** **P=0.050**	0.39 P=0.117		iron transport
CG30026	CG30026	CG30026	**0.76** **P=0.006**	0.12 P=0.220		
CG18067	CG18067	CG18067	**0.76** **P=0.006**	0.05 P=0.277		
Cyp4d21	Cyp4d21	CG6730	**0.77** **P=0.042**	-0.03 P=0.365		electron carrier
Tetraspanin 5D	Tsp5D	CG4690	**0.78** **P=0.047**	0.03 P=0.433	TSPAN9	
CG33115	CG33115	CG33115	**0.78** **P=0.020**	0.21 P=0.005		Nimrod B4
yellow-f2	yellow-f2	CG8063	**0.80** **P=0.005**	0.26 P=0.130		dopachrome conversion, melanazation
virus-induced 1	vir-1	CG31764	**0.80** **P=0.023**	0.12 P=0.253		upregulated in response to viral infection, responsive to RNAi pathway components
fragment D	alpha-Est5	CG1089	**0.80** **P=0.006**	0.25 P=0.083	CES2	
Imaginal disc growth factor 5	Idgf5	CG5154	**0.80** **P=0.021**	0.24 P=0.213		Development of imaginal discs
CG34104	CG34104	CG34104	**0.80** **P=0.004**	0.58 P=1.43E-05		microtubule based G-protein coupled signal transduction
CG3397	CG3397	CG3397	**0.81** **P=0.022**	-0.02 P=0.471		oxidative stress response
CG14872	CG14872	CG14872	**0.81** **P=0.031**	0.08 P=0.307		
bangles and beads	bnb	CG7088	**0.82** **P=0.023**	-0.05 P=0.359		gliogenesis
CG1773	CG1773	CG1773	**0.82** **P=0.008**	-0.11 P=0.031		
beta-galactosidase	Ect3	CG3132	**0.82** **P=0.039**	0.70 P=0.142	GLBL1	autophagic cell death

Gene Title	Gene Symbol	Ensembl	log2(ratio) *elav>rCUG* to *elav>rCAA*	log2(ratio) *elav>rCAG* to *elav>rCAA*	Human Orthologue	Function
CG15293	CG15293	CG15293	**0.83** **P=0.007**	0.31 P=0.131		
Juvenile hormone esterase duplication	Jhedup	CG8424	**0.83** **P=0.020**	0.41 P=0.002		degradation of juvenile hormone
CG4725	CG4725	CG4725	**0.83** **P=0.027**	0.04 P=0.405		Metallo-endopeptidase
CG14528	CG14528	CG14528	**0.84** **P=0.015**	0.72 P=0.024		Metallo-endopeptidase
CG5958	CG5958	CG5958	**0.84** **P=0.040**	-0.04 P=0.428	RLBP1	retinaldehyde-binding protein
CG9396	CG9396	CG9396	**0.86** **P=0.039**	0.10 P=0.026	BRP44	
CG3588	CG3588	CG3588	**0.86** **P=0.044**	0.27 P=0.283		
CG17974	CG17974	CG17974	**0.86** **P=0.019**	0.12 P=0.312		
CG4721	CG4721	CG4721	**0.87** **P=0.046**	0.28 P=0.118		Metallo-endopeptidase
CG6687	CG6687	CG6687	**0.89** **P=4.09E-05**	0.09 P=0.284		hydrogen ion transmembrane transporter
lethal (2) 09851	l(2)09851	CG12792	**0.89** **P=0.003**	0.34 P=0.247	GRWD1	glutamine rich, ribosome biogenesis
Cytochrome P450-9b1	Cyp9b1	CG4485	**0.90** **P=0.024**	-0.29 P=0.054	CYP3A4	electron carrier
Transferrin	Tsf1	CG6186	**0.90** **P=0.048**	0.39 P=0.121		cellular iron ion homeostasis
CG12269	CG12269	CG12269	**0.95** **P=0.021**	0.20 P=0.276		sterol transporter
CG7906	CG7906	CG7906	**0.95** **P=0.038**	0.50 P=0.178		
CG9400	CG9400	CG9400	**0.97** **P=0.014**	0.84 P=0.021		peptidase inhibitor
CG13285	CG13285	CG13285	**0.97** **P=0.041**	0.58 P=0.034		
CG3984	CG3984	CG3984	**1.03** **P=0.019**	0.27 P=0.204		
Mth-like 2	mthl2	CG17795	**1.05** **P=0.001**	0.34 P=0.071		G-protein coupled receptor signalling pathway, extended lifespan
CG10191	CG10191	CG10191	**1.08** **P=0.039**	0.90 P=0.192	WDR51A	
CG16836	CG16836	CG16836	**1.09** **P=0.012**	0.26 P=0.163		
CG34020	CG34020	CG34020	**1.14** **P=0.049**	0.20 P=0.275		
CG9689	CG9689	CG9689	**1.18** **P=0.022**	0.50 P=0.081		
CG7299	CG7299	CG7299	**1.22** **P=0.032**	0.36 P=0.193		

Gene Title	Gene Symbol	Ensembl	log2(ratio) *elav>rCUG* to *elav>rCAA*	log2(ratio) *elav>rCAG* to *elav>rCAA*	Human Orthologue	Function
CG16926	CG16926	CG16926	**1.23** **P=0.006**	0.36 P=0.009		
CG14534	CG14534	CG14534	**1.28** **P=0.027**	-0.01 P=0.495		TweedleE, possible cuticle protein
CG5428	CG5428	CG5428	**1.30** **P=0.033**	0.54 P=0.065	SULT1E1	Sulfo-transferase
CG9079	Cpr47Ea	CG9079	**1.34** **P=0.003**	0.51 P=0.011		structural component of cuticle
CG6188	CG6188	CG6188	**1.40** **P=0.001**	0.07 P=0.457	GNMT	methionine metabolism
metchnikowin	Mtk	CG8175	**1.47** **P=0.009**	-0.08 P=0.364		defense response
Cytochrome P450-4e3	Cyp4e3	CG4105	**1.57** **P=0.010**	0.56 P=0.149		electron carrier
CG17777	CG17777	CG17777	**1.61** **P=0.009**	0.81 P=0.132		
CG5966	CG5966	CG5966	**1.92** **P=0.001**	0.51 P=0.031	PNLIP	triacylglycerol lipase
Turandot	TotA	CG31509	**3.14** **P=0.008**	1.04 P=0.074		stress response
Turandot C	TotC	CG31508	**4.16** **P=0.004**	1.01 P=0.075		stress response

Table B9: Genes altered in *elav>rCAG* flies compared to *elav>+* in microarray experiment 2. Selected for Log2(ratio) >0.5 or <-0.5, P<0.05.

Gene Title	Gene Symbol	Ensembl	log2(ratio) *elav>rCAG* to *elav>+*	log2(ratio) *elav>rCUG* to *elav>+*	Human Orthologue	Function
Stellate orphon	Ste12DOR	CG32616	**-2.88** **P=0.002**	-2.0 P=0.036		Spermatogenesis, protein kinase regulator
CG7526	CG7526	CG7526	**-1.88** **P=0.024**	1.46 P=0.053		
CG31606	CG31606	CG31606	**-0.96** **P=0.009**	1.46 P=0.150		
CG9498	CG9498	CG9498	**-0.82** **P=2.99E-5**	0.39 P=0.467		
CG11893	CG11893	CG11893	**-0.81** **P=0.001**	-0.1 P=0.792		
CG3823	CG3823	CG3823	**-0.8** **P=0.005**	0.3 P=0.119		Vitamin E binding
Modifier67.2	mod(mdg4)	CG32491	**-0.76** **P=0.012**	-0.71 P=0.071	AKAP1	Regulation of chromatin assembly/disassembly, regulation of apoptosis
Odorant-binding protein 19c	Obp19c	CG15457	**-0.75** **P=0.029**	-1.02 P=0.007		Sensory perception of chemical stimulus
CG4688	CG4688	CG4688	**-0.72** **P=0.039**	-0.99 P=0.007		Glutathione transferase activity
CG33054	CG33054	CG33054	**-0.71** **P=0.043**	-0.77 P=2.47E-5		
CG9492	CG9492	CG9492	**-0.67** **P=0.040**	-0.44 P=0.094	DNAH5	Microtubule motor activity, component of axonemal dynein complex
hemipterous	hep	CG4353	**-0.64** **P=0.029**	-0.34 P=0.096	MAP2K7	Jun kinase activity, roles in establishment of planar polarity
CG13895	CG13895	CG13895	**-0.62** **P=0.016**	-0.43 P=0.016		
CG14401	CG14401	CG14401	**-0.62** **P=0.022**	0.44 P=0.368		
Doughnut	dnt	CG17559	**-0.6** **P=0.010**	-0.49 P=0.049	RYK	Axon guidance, muscle attachment, Wnt signalling pathway
CG6752	CG6752	CG6752	**-0.58** **P=0.003**	-0.54 P=0.014	RNF123	
CG13077	CG13077	CG13077	**-0.55** **P=0.019**	-0.79 P=0.014	CYB561D2	
CG32552	CG32552	CG32552	**-0.55** **P=0.013**	-0.86 P=0.030		
CG13117	CG13117	CG13117	**-0.53** **P=0.015**	-0.56 P=0.013		
CG7031	CG7031	CG7031	**-0.52** **P=0.022**	0.25 P=0.170		
CG15545	CG15545	CG15545	**-0.52** **P=0.029**	-0.23 P=0.201		
CG6197	CG6197	CG6197	**-0.52** **P=0.047**	-0.41 P=0.108	XAB2	Regulation of alternative splicing, phagocytosis
CG2162	CG2162	CG2162	**-0.51** **P=0.007**	-0.26 P=0.267	R3HCC1	
CG30143	CG30143	CG30143	**0.51** **P=0.046**	0.53 P=0.053		

Gene Title	Gene Symbol	Ensembl	log2(ratio) *elav>rCAG* to *elav>+*	log2(ratio) *elav>rCUG* to *elav>+*	Human Orthologue	Function
CG8445	CG8445	CG8445	**0.52** **P=0.035**	0.25 P=0.173	BAP1	Ubiquitin thiolesterase
CG4484	CG4484	CG4484	**0.54** **P=0.039**	0.48 P=0.035	SLC45A1	Glucose trans-membrane transport
CG7714	CG7714	CG7714	**0.55** **P=0.016**	0.28 P=0.290		
CG9686	CG9686	CG9686	**0.55** **P=0.014**	0.54 P=0.119		
Marionette	mrn	CG7764	**0.56** **P=0.030**	0.3 P=0.075	GTF2H4	Regulation of transcription, DNA repair
CG2540	CG2540	CG2540	**0.56** **P=0.030**	0.39 P=0.184	CHAC2	Cation transporter
Cyp311a1	Cyp311a1	CG1488	**0.57** **P=0.041**	0.59 P=0.255		Electron carrier
CG15107	CG15107	CG15107	**0.57** **P=0.031**	0.4 P=0.270		
CG5945	CG5945	CG5945	**0.58** **P=0.024**	0.49 P=0.178		
CG9186	CG9186	CG9186	**0.58** **P=0.017**	0.77 P=0.005		
CG7777	CG7777	CG7777	**0.59** **P=0.018**	0.45 P=0.011	AQP1	Water transporter
CG3056	CG3056	CG3056	**0.59** **P=0.049**	0.47 P=0.062		mRNA binding
CG4004	CG4004	CG4004	**0.6** **P=0.026**	-0.26 P=0.570		
wunen2	wun2	CG8805	**0.6** **P=0.040**	1.09 P=0.048	PPAP2A	Lipid phosphatase activity, germ cell development
CG15893	CG15893	CG15893	**0.61** **P=0.045**	0.16 P=0.380		
CG2812	CG2812	CG2812	**0.61** **P=0.018**	0.19 P=0.530	WDR47	
CG9657	CG9657	CG9657	**0.61** **P=0.030**	1.04 P=0.022	SLC5A12	Sodium:iodide transporter
CG7634	CG7634	CG7634	**0.61** **P=0.006**	-0.3 P=0.451		
Adenosine deaminase-related growth factor D	Adgf-D	CG9621	**0.62** **P=0.010**	1.33 P=0.065	CECR1	Adenosine deaminase activity, growth factor activity
echinoid	ed	CG12676	**0.63** **P=0.002**	-0.24 P=0.027		actin cytoskeleton/cell adhesion, may form a signalling complex with Grip, important in myogenesis
CG31704	CG31704	CG31704	**0.63** **P=0.023**	0.86 P=0.197	SPINK2	Serine-type endopeptidase inhibitor
Drosomycin	Drs	CG10810	**0.64** **P=0.012**	1.09 P=0.025		Innate immune response
CG32816	CG32816	CG32816	**0.65** **P=0.048**	0.44 P=0.098		
CG33528	CG33528	CG33528	**0.68** **P=0.025**	0.11 P=0.557	SLC18A2	Neurotransmitter secretion, synaptic vesicle amine transport
CG5653	CG5653	CG5653	**0.68** **P=0.010**	0.4 P=0.192	SMOX	

Gene Title	Gene Symbol	Ensembl	log2(ratio) *elav>rCAG* to *elav>+*	log2(ratio) *elav>rCUG* to *elav>+*	Human Orthologue	Function
CG11825	CG11825	CG11825	**0.70** **P=0.019**	0.31 P=0.728		
CG13704	CG13704	CG13704	**0.73** **P=0.045**	0.68 P=0.398		
Immune induced molecule 2	IM2	CG18106	**0.75** **P=0.017**	1.01 P=0.106		Immune response
Tetraspanin 5D	Tsp5D	CG4690	**0.81** **P=0.042**	1.57 P=0.020	TSPAN9	
CG14959	CG14959	CG14959	**0.81** **P=0.012**	0.82 P=0.046		Chitin binding
CG18641	CG18641	CG18641	**0.85** **P=0.011**	0.72 P=0.036		Lipase activity
CG14528	CG14528	CG14528	**0.85** **P=0.026**	0.97 P=0.003		Metallo-endopeptidase
CG11910	CG11910	CG11910	**0.99** **P=0.047**	1.96 P=0.029		Insulin-like growth factor binding protein complex
CG9394	CG9394	CG9394	**1.00** **P=0.002**	1.8 P=0.035		Lipid metabolism, glycerophosphodiester phosphodiesterase
CG14495	CG14495	CG14495	**1.04** **P=0.002**	1.52 P=0.018		
Hormone receptor-like in 38	Hr38	CG1864	**1.05** **P=0.040**	0.55 P=0.094	NR4A1	Ligand-dependent nuclear receptor activity
CG1732	CG1732	CG1732	**1.06** **P=0.008**	2.58 P=0.075	SLC6A1	Neurotransmitter transport, GABA:sodium transporter
CG13062	CG13062	CG13062	**1.26** **P=0.029**	0.76 P=0.072		
CG2901	CG2901	CG2901	**1.27** **P=0.016**	-1.14 P=0.277		
yolk protein	Yp3	CG11129	**1.43** **P=0.004**	0.47 P=0.194		lipid metabolism
white	w	CG2759	**2.67** **P=0.002**	2.61 P=0.002	ABCG5	Eye pigment precursor transport, metabolic process

Table B10: Genes altered in *elav>rCUG* flies compared to *elav>+* in microarray experiment 2. Selected for Log2(ratio) >0.5 or <-0.5, P<0.05.

Gene Title	Gene Symbol	Ensembl	log2(ratio) *elav>rCUG* to *elav>+*	log2(ratio) *elav>rCAG* to *elav>+*	Human Orthologue	Function
Stellate orphon	Ste12DOR	CG32616	**-2.00** **P=0.003**	-2.88 P=0.002		Spermato-genesis, protein kinase regulator
CG16752	CG16752	CG16752	**-1.24** **P=0.006**	-0.14 P=0.056		Neuropeptide receptor activity
scabrous	sca	CG17579	**-1.22** **P=0.011**	-0.37 P=0.340		Fibrinogen complex, bristle morphogenesis, nervous system development, signal transduction
CG14216	CG14216	CG14216	**-1.02** **P=0.017**	-0.59 P=0.104	SSU72	mRNA processing
CG4688	CG4688	CG4688	**-0.99** **P=0.017**	-0.72 P=0.039	MARS	Glutathione transferase
CG13908	CG13908	CG13908	**-0.95** **P=0.012**	-0.62 P=0.174		
TBPH	TBPH	CG10327	**-0.94** **P=0.034**	0.04 P=0.943	TARDBP	Neuromuscular junction development, RNA binding
CG40084	CG40084	CG42595	**-0.89** **P=0.011**	-0.54 P=0.191		
CG13035	CG13035	CG13035	**-0.88** **P=0.018**	-0.41 P=0.039	NSUN7	
CG32552	CG32552	CG32552	**-0.86** **P=0.010**	-0.55 P=0.013		
CG31781	CG31781	CG31781	**-0.82** **P=0.010**	-0.13 P=0.058		
Cyp12a4	Cyp12a4	CG6042	**-0.82** **P=0.050**	-0.26 P=0.437	CYP24A1	Electron carrier activity
sequoia	seq	CG32904	**-0.81** **P=0.041**	-0.19 P=0.479		Dendrite morphogenesis
sickie	sick	CG42589	**-0.81** **P=0.015**	-0.19 P=0.482		Immune response
CG32521	CG32521	CG32521	**-0.8** **P=0.008**	-0.1 P=0.765		
D-Titin	sls	CG1915	**-0.8** **P=0.029**	-0.3 P=0.086	TTN	Mesoderm development, myoblast fusion
CG13077	CG13077	CG13077	**-0.79** **P=0.027**	-0.55 P=0.019	CYB561D2	
CG31846	CG31846	CG31846	**-0.77** **P=0.027**	-0.24 P=0429		
CG40188	CG40188	CG40188	**-0.76** **P=0.040**	-0.41 P=0.386		
scabrous	sca	CG17579	**-0.75** **P=0.039**	-0.38 P=0.212		Fibrinogen complex, bristle morphogenesis, nervous system development, signal transduction
CG5883	CG5883	CG5883	**-0.75** **P=0.005**	-0.58 P=0.237		Chitin metabolic process
Poly-glutamine tract binding protein 1	PQBP-1	CG31369	**-0.73** **P=0.013**	-0.20 P=0.455		
alan shepard	shep	CG32423	**-0.73** **0.002**	0.05 P=0.915		

Gene Title	Gene Symbol	Ensembl	log2(ratio) *elav>rCUG* to *elav>+*	log2(ratio) *elav>rCAG* to *elav>+*	Human Orthologue	Function
aquaporin	AQP /// blw	CG12251	**-0.70** **P=0.019**	-0.24 P=0.490	AQP12A	Water transporter, mitochondrial
CG5010	CG5010	CG5010	**-0.70** **P=0.026**	-0.20 P=0.702	CHCHD2	
CG7330	CG7330	CG7330	**-0.70** **P=0.024**	-0.38 P=0.068		
section 7	l(1)19Ec	CG11233	**-0.69** **P=0.037**	-0.43 P=0.371	PPP1R11	Dorsal closure, nervous system development
CG32457	CG32457	CG32457	**-0.67** **P=0.044**	0.10 P=0.272		
CG15311	CG15311	CG15311	**-0.66** **P=0.001**	-0.27 P=0.545		Diphosphatase
dynamin	shi	CG18102	**-0.65** **P=0.017**	-0.12 P=0.278	DNM1	Synaptic vesicle endocytosis, microtubule motor activity
CG40485	CG40485	CG40485	**-0.65** **P=0.003**	-0.40 P=0.070		Oxidoreductase activity
CG8213	CG8213	CG8213	**-0.64** **P=0.041**	-0.28 P=0.210	ST14	Serine-type endopeptidase
CG40485	CG40485	CG40485	**-0.63** **P=0.020**	-0.45 P=0.007		
CG14614	CG14614	CG14614	**-0.63** **P=0.017**	-0.10 P=0.393	WDR68	
Frizzled	fz	CG17697	**-0.62** **P=0.002**	-0.20 P=0.014	FZD1	Wnt receptor, establishment of cell polarity, cell adhesion
CG1463	CG1463	CG1463	**-0.62** **P=0.007**	-0.28 P=0.011		
hu-li tai shao	hts	CG9325	**-0.62** **P=0.032**	-0.44 P=0.056	ADD1	Actin assembly, ring canal formation
ETS-domain lacking	edl	CG15085	**-0.61** **P=0.024**	-0.11 P=0.781		Nuclear export, embryonic patterning
CG31814	CG31814	CG31814	**-0.61** **P=0.014**	0.13 P=0.740	HNT	
lethal (1) G0136	l(1)G0136	CG8198	**-0.61** **P=0.005**	-0.25 P=0.044	HBLD2	Iron-sulfur cluster assembly, mitochondrial
CG13025	CG13025	CG13025	**-0.60** **P=0.009**	0.14 P=0.734	RFWD3	
mindmelt	mbl	CG33197	**-0.60** **P=0.022**	-0.09 P=0.582	MBNL1	Splicing factor, muscle and nervous system development
CG11138	CG11138	CG11138	**-0.60** **P=0.021**	-0.14 P=0.714	IRF2BP2	
CG3259	CG3259	CG3259	**-0.60** **P=0.036**	-0.43 P=0.082	TRAFIP1	TRAF3 associated, cilium assembly
CG12825	CG12825	CG12825	**-0.59** **P=0.041**	-0.21 P=0.411		
CG6340	CG6340	CG6340	**-0.59** **P=0.023**	-0.21 P=0.506	RSRC2	
expanded	ex	CG4114	**-0.58** **P=0.023**	-0.11 P=0.066	FRMD6	Regulation of cell proliferation and differentiation
CheA7a	CheA7a	CG15033	**-0.58** **P=0.007**	-0.06 P=0.749		
m-spondin	mspo	CG10145	**-0.58** **P=0.013**	-0.17 P=0.627	SPON2	

Gene Title	Gene Symbol	Ensembl	log2(ratio) *elav>rCUG* to *elav>+*	log2(ratio) *elav>rCAG* to *elav>+*	Human Orthologue	Function
Cyp4s3	Cyp4s3	CG9081	**-0.58** **P=0.014**	-0.0 P=0.910		Electron carrier
TGF-beta activated kinase 1	Tak1	CG18492	**-0.57** **P=0.005**	0.01 P=0.831	MAP3K7	JNK cascade, apoptosis, immune response
CG40298	CG40298	CG40298	**-0.57** **P=0.008**	-0.25 P=0.012		
Neprilysin 1	Nep1	CG5905	**-0.57** **P=0.047**	-0.47 P=0.057	ECE2	Serine-type endopeptidase, heart development
CG18646	CG18646	CG18646	**-0.56** **P=0.047**	-0.21 P=0.681		GTPase activity
CG13117	CG13117	CG13117	**-0.56** **P=0.009**	-0.53 P=0.015		
CG4804	CG4804	CG4804	**-0.56** **P=0.027**	0.08 P=0.714		Serine-type endopeptidase inhibitor
Overflow	Dl	CG3619	**-0.56** **P=0.001**	-0.32 P=0.006	DLL1	Notch signalling pathway, neural development
***Drosophila* Tat-like**	CG31241	CG31241	**-0.56** **P=0.012**	0.00 P=0.991	TGS1	RNA methylation
quo vadis	shn	CG7734	**-0.55** **P=0.031**	-0.18 P=0.433		TGFβ signalling, patterning and organ development
CG13293	CG13293	CG13293	**-0.55** **P=0.033**	-0.26 P=0.154		
CG6752	CG6752	CG6752	**-0.54** **P=0.027**	-0.58 P=0.003	RNF123	
SET domain binding factor	Sbf	CG6939	**-0.54** **P=0.046**	-0.45 P=0.065	SBF2	DAG signalling
lethal (2) essential for life	l(2)efl	CG4533	**-0.53** **P=0.011**	-0.44 P=0.256	CRYAB	Response to heat
CG3883	CG3883	CG3883	**-0.53** **P=0.043**	-0.39 P=0.359		
Ect4	Ect4	CG7915	**-0.52** **P=0.046**	-0.06 P=0.303	SARM1	Innate immune response
bruno	aret	CG31762	**-0.52** **P=0.015**	-0.07 P=0.291	CUGBP1	Regulation of alternative splicing and translation
CG4078	CG4078	CG4078	**-0.52** **P=0.046**	-0.17 P=0.112	RTEL1	Nucleotide excision repair
Smg5	Smg5	CG8954	**-0.51** **P=0.047**	-0.22 P=0.067	SMG5	Gene silencing by miRNA, nonsense-mediated decay
CG10617	CG10617	CG10617	**-0.51** **P=0.048**	-0.32 P=0.114	SYT12	Neurotrans-mitter secretion
fettucine	cic	CG5067	**-0.51** **P=0.011**	-0.25 P=0.017	CIC	Transcription factor activity
Olfactory-specific E	Os-E	CG11422	**0.52** **P=0.046**	0.40 P=0.266		Sensory perception of chemical stimulus
semiphorin	Sema-5c	CG5661	**0.52** **P=0.016**	0.37 P=0.138	SEMA5A	Axon guidance
CG12484	CG12484	CG12484	**0.53** **P=0.031**	0.51 P=0.046		
CG34054	CG34054	CG34054	**0.53** **P=0.036**	-0.21 P=0.241		

Gene Title	Gene Symbol	Ensembl	log2(ratio) *elav>rCUG* to *elav>+*	log2(ratio) *elav>rCAG* to *elav>+*	Human Orthologue	Function
cuticle cluster 2	Ccp84Ag	CG2342	**0.54** **P=0.047**	0.90 P=0.116		Component of cuticle
CG9686	CG9686	CG9686	**0.54** **P=0.016**	0.55 P=0.014		
Equilibrative nucleoside transporter 2	Ent2	CG31911	**0.55** **P=0.012**	0.25 P=0.103	SLC29A1	
Hormone receptor-like in 38	Hr38	CG1864	**0.55** **P=0.028**	1.05 P=0.040	NR4A1	Ligand-dependent nuclear receptor activity, transcription activity
CG15209	CG15209	CG15209	**0.56** **P=0.010**	0.26 P=0.009		
CG8021	CG8021	CG8021	**0.57** **P=0.035**	0.32 P=0.366		
CG10562	CG10562	CG10562	**0.58** **P=0.011**	0.27 P=0.296		
reaper	rpr	CG4319	**0.59** **P=0.029**	0.31 P=0.378		Regulation of apoptosis
CG3603	CG3603	CG3603	**0.59** **P=0.018**	0.42 P=0.289	HSD17B8	Oxidation reduction
Inwardly rectifying potassium channel 3	Irk3	CG10369	**0.60** **P=0.030**	0.24 P=0.107		Potassium channel
CG9449	CG9449	CG9449	**0.60** **P=0.031**	0.45 P=0.124		phagocytosis
CG9134	CG9134	CG9134	**0.61** **P=0.020**	0.48 P=0.020	BCAN	
lectin-58Fg	Ugt58Fa	CG4414	**0.61** **P=0.031**	0.48 P=0.256		Glucoronosyl-transferase activity
CG9027	CG9027	CG9027	**0.62** **P=0.037**	0.50 P=0.133	SOD1	Superoxide dismutase activity
CG15456	CG15456	CG15456	**0.63** **P=0.011**	0.09 P=0.779		Cell redox homeostasis
CG5585 /// **CG6434**	CG5585 /// CG6434	CG5585 /// CG6434	**0.64** **P=0.015**	0.37 P=0.109	RBBP5	
Grip128	Grip128	CG9201	**0.66** **P=0.024**	0.38 P=0.063		Mitotic spindle arrangement
CG12179	CG12179	CG12179	**0.74** **P=0.041**	0.72 P=0.155		
CG17758	CG17758	CG17758	**0.75** **P=0.017**	0.93 P=0.265	OTOP1	
CG9186	CG9186	CG9186	**0.77** **P=0.031**	0.58 P=0.017		
Kua	Kua	CG10723	**0.91** **P=0.002**	0.56 P=0.182	KUA	
CG10006	CG10006	CG10006	**0.95** **P=0.001**	0.56 P=0.329	SLC39A6	Metal ion transporter
CG6465	CG6465	CG6465	**1.01** **P=0.014**	0.85 P=0.111	ACY1	Hydrolysis of acetylated amino acids
sugarbabe	sug	CG3850	**1.02** **P=0.040**	0.33 P=0.319	GLIS2	Transcriptional regulation

Gene Title	Gene Symbol	Ensembl	log2(ratio) *elav>rCUG* to *elav>+*	log2(ratio) *elav>rCAG* to *elav>+*	Human Orthologue	Function
lethal (1) G0155	Ykt6	CG1515	**1.05** **P=0.012**	0.61 P=0.518	YKT6	Vesicle-mediated transport, SNAP receptor activity
Odorant-binding protein 8a	Obp8a	CG12665	**1.11** **P=0.001**	1.03 P=0.511		Sensory perception of chemical stimulus
CG8974	CG8974	CG8974	**1.31** **P=0.003**	0.95 P=0.362	RNF185	
Bicoid interacting protein 4	Pk17E	CG7001	**1.36** **P=0.003**	0.34 P=0.816		Protein kinase
CG10924	CG10924	CG10924	**1.52** **P=0.043**	-0.16 P=0.108	PCK2	Phosphoenol-pyruvate carboxykinase, gluconeo-genesis
anon-fast-evolving-1H4	Rala	CG2849	**2.03** **P=0.018**	1.30 P=0458	RALA	GTPase, actin cytoskeleton reorganisation

Appendix C

Table C1: Genes altered in *elav>rAUUCU* compared to both *elav>+* and *elav>rCAA.* Selected for Log2(ratio) >0.5 or <-0.5.

Gene Title	Gene Symbol	Ensembl	log2(ratio) *elav>rAUUCU* to *elav>rCAA*	log2(ratio) *elav>rAUUCU* to *elav>+*	Human Orthologue	Function
CG3898	CG3898	CG3898	**-1.27**	**-1.78**		
CG32553	CG32553	CG32553	**-0.87**	**-1.73**		
CG6416	CG6416	CG6416	**-1.62**	**-1.66**		Mesoderm development
Stretchin	Strn-Mlck	CG18255	**-0.96**	**-1.64**		Myosin light chain kinase, component of cytoskeleton
0.9kb transcript	CG2650	CG2650	**-0.80**	**-1.62**		
arthrin	Act88F	CG5178	**-1.13**	**-1.57**		Cytoskeleton organisation
CG18646	CG18646	CG42629	**-1.03**	**-1.55**		Regulation of GTPase activity
kruppel	Kr	CG3340	**-1.04**	**-1.51**		Transcriptional repressor, neuroblast fate determination
CG15335	CG15335	CG15335	**-1.57**	**-1.43**		
DCorin	Corin	CG2105	**-1.26**	**-1.39**	CORIN	Serine-type endopeptidase
79B Actin	Act79B	CG7478	**-0.89**	**-1.36**	ACTA2	Cytoskeleton organisation
CG40484	CG40484	CG40484	**-0.82**	**-1.36**		
CG40485	CG40485	CG40485	**-0.92**	**-1.35**		Oxidoreductase activity
CG17378	CG17378	CG17378	**-0.79**	**-1.35**		
CG13131	CG13131	CG13131	**-1.21**	**-1.29**		
CG12455	CG12455	CG12455	**-0.82**	**-1.23**		Calcium channel activity
CG31174	CG31174	CG31174	**-1.60**	**-1.23**		
Male-specific RNA 84Dc	Mst84Dc	CG17945	**-0.79**	**-1.22**		Electron carrier activity
CG9194	CG9194	CG9194	**-0.75**	**-1.2**		Potassium ion transporter
CG12716	CG12716	CG12716	**-1.29**	**-1.2**		
CG32652	CG32652	CG32652	**-0.51**	**-1.19**		
CG31708	CG31708	CG31708	**-0.68**	**-1.18**		

Gene Title	Gene Symbol	Ensembl	log2(ratio) *elav>rAUUCU* to *elav>rCAA*	log2(ratio) *elav>rAUUCU* to *elav>+*	Human Orthologue	Function
CG10260	CG10260	CG10260	**-0.88**	**-1.17**	PIK4CA	1-phosphatidylinositol 4-kinase activity, signalling
CG14133	CG14133	CG14133	**-0.65**	**-1.14**		
CG31730	CG31730	CG31730	**-0.97**	**-1.13**	ARD1A	N-acetyl transferase activity
CG40323	CG40323	CG40323	**-0.88**	**-1.12**		
whacked	wkd	CG5344	**-0.65**	**-1.11**	TBC1D10A	Regulation of RAB GTPase activity
CG40139	CG40139	CG40139	**-0.73**	**-1.09**		
CG11637	CG11637	CG11637	**-0.63**	**-1.09**		Cell adhesion
CYP4-related	Cyp4ad1	CG2110	**-1.20**	**-1.09**		Electron carrier activity
CG18747	CG18747		**-1.02**	**-1.08**		
larval-opioid-receptor	FR	CG2114	**-0.87**	**-1.08**		Neuropeptide receptor
CG13060	CG13060	CG13060	**-0.79**	**-1.08**		
Complement-ation group C	Mef2	CG1429	**-0.71**	**-1.04**	MEF2C	Mesoderm and heart development, transcriptional regulation
CG7544	CG7544	CG7544	**-0.78**	**-1.04**	METT10D	Methyl-transferase activity
CG33147	Hs3st-A	CG33147	**-0.62**	**-1.03**	HS3ST5	Heparan sulphate-glucosamine 3-sulfotransferase 1 activity
CG7341	CG7341	CG7341	**-0.60**	**-1.02**		
CG18348	Cpr67Fb	CG18348	**-0.97**	**-1.01**		Structural component of cuticle
CG12524	CG34356	CG34356	**-0.82**	**-1.00**		DNA methylation
CG9264	CG9264	CG9264	**-0.88**	**-0.97**		Trans-membrane amino acid transporter
CG14669	CG14669	CG14669	**-0.99**	**-0.97**		
Anillin	scra	CG2092	**-1.28**	**-0.96**	ANLN	Microtubule binding, cytokinesis
defective transmitter release	dtr	CG31623	**-0.75**	**-0.93**		Synaptic transmission
CG17290	CG17290	CG17290	**-1.60**	**-0.93**		
Cyp313b1	Cyp313b1	CG9716	**-0.74**	**-0.92**		Electron carrier
Hemese	He	CG31770	**-1.32**	**-0.92**		Innate immune response

Gene Title	Gene Symbol	Ensembl	log2(ratio) *elav>rAUUCU* to *elav>rCAA*	log2(ratio) *elav>rAUUCU* to *elav>+*	Human Orthologue	Function
CG13762	CG13762	CG13762	**-1.34**	**-0.91**		Calcium ion transport
CG31878	CG31878	CG42367	**-0.88**	**-0.91**		
calcium binding protein	Scp1	CG15848	**-0.64**	**-0.9**		Calcium ion transport
Cyclin-dependent kinase interactor 5	CycJ	CG10308	**-0.98**	**-0.9**	CCNJ	Cyclin-dependent protein kinase regulator
khotalo	kto	CG8491	**-0.61**	**-0.89**	MED12L	Transcriptional regulation, positive regulator of wnt signalling pathway
Odorant receptor 59b	Or59b	CG3569	**-0.69**	**-0.87**		Olfactory receptor activity
hu-li tai shao	hts	CG9325	**-1.12**	**-0.87**	ADD1	Actin assembly, ring canal formation
CG12524	CG34356	CG12524	**-0.75**	**-0.87**		DNA methylation
Abnormal chemosensory jump	acj6	CG9151	**-0.89**	**-0.86**	POU4F3	Dendrite morphogenesis, transcription factor activity
CREB binding protein	nej	CG15319	**-0.66**	**-0.84**	EP300	Transcriptional co-activator
CG15214	CG34391	CG34391	**-0.97**	**-0.84**		
roughened eye	rn	CG42277	**-1.00**	**-0.83**		Transcriptional regulation
CG13235	CG13235	CG13235	**-0.64**	**-0.83**		
SP71	SP71	CG17131	**-0.90**	**-0.82**		
Shaker	Sh	CG12348	**-0.63**	**-0.81**	KCNA2	Voltage-gated potassium channel
grapes	grp	CG17161	**-0.59**	**-0.81**	CHEK1	Cell-cycle check-point
oskar	osk	CG10901	**-0.83**	**-0.81**		P-granule assembly, pole cell formation
CG9009	CG9009	CG9009	**-0.76**	**-0.8**		Long-chain-fatty-acid-CoA ligase activity
CG31988	CG31988	CG31988	**-0.64**	**-0.79**		Zinc ion binding
Rbp1-like	Rbp1-like	CG1987	**-0.64**	**-0.79**	SFRS3	Nuclear mRNA splicing
CG5048	CG5048	CG5048	**-0.62**	**-0.78**		
TpnC4	TpnC4	CG12408	**-0.92**	**-0.78**		Calcium ion binding
dpr8	dpr8	CG32600	**-0.52**	**-0.77**		

Gene Title	Gene Symbol	Ensembl	log2(ratio) *elav>rAUUCU* to *elav>rCAA*	log2(ratio) *elav>rAUUCU* to *elav>+*	Human Orthologue	Function
CG32425	CG32425	CG32425	**-0.54**	**-0.77**		
CG14151	CG14151	CG14151	**-1.05**	**-0.77**		
dumbfounded	kirre	CG3653	**-0.66**	**-0.76**		Myoblast fusion, muscle tissue development
CG16758	CG16758	CG16758	**-0.62**	**-0.75**		Nucleoside metabolism
seven-up	svp	CG11502	**-0.59**	**-0.75**	NR2F1	Synaptic transmission
fau	fau	CG6544	**-0.53**	**-0.74**		
CG34135	CG34135	CG34135	**-0.81**	**-0.73**		
Shaker	Sh	CG12348	**-0.51**	**-0.72**	KCNA2	Voltage-gated potassium channel
turtle	tutl	CG15427	**-0.67**	**-0.72**	IGSF9	Regulation of dendrite morphogenesis
CG13545	CG13545	CG13545	**-0.59**	**-0.71**		
Fancd2	Fancd2	CG17269	**-1.21**	**-0.71**	FANCD2	Protein kinase activity
crossover suppressor on 2 of Manheim	c(2)M	CG4249	**-0.99**	**-0.7**		Resolution of cross-over intermediates in meiosis
Adenylate cyclase 3	Ac3	CG1506	**-0.59**	**-0.7**	ADCY3	cAMP synthesis, signalling cascade
CG17446	CG17446	CG17446	**-0.56**	**-0.69**	CXXC1	DNA binding
CG14655	CG14655	CG14655	**-0.62**	**-0.69**		
CG40092	CG40092	CG40092	**-0.53**	**-0.69**		
CG31283	CG31283	CG31283	**-0.59**	**-0.68**		
CG40450	CG40450	CG40450	**-0.59**	**-0.67**		
CG12552	CG12552	CG12552	**-0.53**	**-0.66**		
Glycogen Synthase Kinase 3	sgg	CG2621	**-0.51**	**-0.66**	GSK3B	Protein serine threonine kinase activity, synaptic growth at neuromuscular junction
methuselah-like 13	mthl13	CG30018	**-0.69**	**-0.66**		
CG4328	CG4328	CG4328	**-0.73**	**-0.66**	LMX1B	Dendrite morphogenesis, transcription factor activity
RNA-binding protein 1	Rbp1	CG17136	**-0.70**	**-0.66**	SFRS3	Nuclear RNA splicing

Gene Title	Gene Symbol	Ensembl	log2(ratio) *elav>rAUUCU* to *elav>rCAA*	log2(ratio) *elav>rAUUCU* to *elav>+*	Human Orthologue	Function
CG9200	Atac1	CG9200	**-0.58**	**-0.65**	ZZZ3	Regulation of histone acetylation
His1:CG31617	His1:CG31617	CG31617	**-0.69**	**-0.65**	HIST1H1B	Chromatin assembly or disassembly
5' gene	Gpdh	CG9042	**-0.90**	**-0.65**	GPD1	Glycerol phosphate dehydrogenase
CG8501	CG8501	CG8501	**-0.84**	**-0.65**		
CG18304	CG18304	CG18304	**-0.52**	**-0.65**		Component of ribosome
Ady43A	Ady43A	CG1851	**-0.60**	**-0.65**		Purine ribonucleoside salvage
protein phosphatase from PCR fragment D27	CanA-14F	CG9819	**-1.21**	**-0.64**	PPP3CA	Amino acid de-phosphorylation
CG11374	CG11374	CG11374	**-0.62**	**-0.64**		Sugar binding
CG30271	CG30271	CG30271	**-0.52**	**-0.64**		
CG40449	WDY	CG40449	**-0.55**	**-0.64**		
Serotonin receptor	5-HT7	CG12073	**-0.55**	**-0.64**	ADRA2B	Serotonin receptor signalling, activation of adenylate cyclase
CG32532	CG32532	CG32532	**-0.50**	**-0.64**		Transcription factor activity
Interferon-like protein	ect	CG6611	**-0.65**	**-0.63**		
CG11188	CG11188	CG11188	**-0.73**	**-0.63**	AATF	Regulation of apoptosis, transcription factor activity
CG32026	CG32026	CG32026	**-0.58**	**-0.63**		Isocitrate dehydrogenase activity
vasa	vas	CG3506	**-0.99**	**-0.63**	DDX4	Regulation of RNA localisation and translation
CG33279	CG33279	CG42458	**-0.54**	**-0.63**		mRNA binding
CG30054	CG30054	CG30054	**-0.55**	**-0.62**		G-protein coupled signalling
eye gone	eyg	CG10488	**-0.51**	**-0.62**		Transcriptional regulation
Myocardin-related transcription factor	Mrtf	CG32296	**-0.67**	**-0.62**	MKL2	Actin cytoskeleton organisation, transcription factor activity
longitudinals absent	lola	CG12052	**-0.65**	**-0.62**	ZBTB3	Axon guidance, transcriptional regulation

Gene Title	Gene Symbol	Ensembl	log2(ratio) *elav>rAUUCU* to *elav>rCAA*	log2(ratio) *elav>rAUUCU* to *elav>+*	Human Orthologue	Function
Complement-ation group D	Hr46	CG33183	**-0.97**	**-0.61**	RORB	Ligand-dependent nuclear receptor, transcriptional regulation
CG13834	CG34375	CG34375	**-0.51**	**-0.61**		Ubiquitin-dependent catabolism
CG31513	CG31513	CG31513	**-0.66**	**-0.61**		
CG40115	CG40115	CG40115	**-0.65**	**-0.61**		
dynein-related heavy chain polypeptide	Dhc93AB	CG3723	**-0.74**	**-0.6**	DNAH9	Microtubule-based movement
CG15308	CG15308	CG15308	**-0.52**	**-0.6**		
CG17163	CG17163	CG17163	**-0.54**	**-0.6**		
CG13012	CG13012	CG13012	**-0.51**	**-0.59**		
5' gene	Gpdh	CG9042	**-0.76**	**-0.59**	GPD1	Glycerol phosphate dehydrogenase
female lethal d	fl(2)d	CG6315	**-0.59**	**-0.59**	WTAP	RNA splicing, sex determination
CG2260	CG2260	CG2260	**-0.85**	**-0.59**		
CG40137	CG40137	CG40137	**-0.54**	**-0.58**		
CG6185	CG6185	CG6185	**-0.72**	**-0.57**		Glutamate-gated ion channel activity
CG11617	CG11617	CG11617	**-0.56**	**-0.56**		Transcriptional regulation
CG6012	CG6012	CG6012	**-0.52**	**-0.56**		Oxidoreductase activity
CG10086	CG10086	CG10086	**-0.64**	**-0.55**		
CG14662	CG14662	CG14662	**-0.55**	**-0.55**		
CG30054	CG30054	CG30054	**-0.56**	**-0.54**		G-protein coupled receptor activity
Rpb4	Rpb4	CG33520	**-0.70**	**-0.53**	TADA2L	Regulation of histone acetylation, transcriptional activator adaptor protein
CG14131	CG14131	CG14131	**-0.67**	**-0.53**		
Lipase 1	Lip1	CG7279	**-0.70**	**-0.53**		Lipid metabolism
CG13038	CG13038	CG13038	**-0.69**	**-0.53**		
fru-satori	fru	CG14037	**-0.85**	**-0.53**		CNS development, transcription factor activity
CG18606	CG10476	CG10476	**-0.68**	**-0.52**		

Gene Title	Gene Symbol	Ensembl	log2(ratio) *elav>rAUUCU* to *elav>rCAA*	log2(ratio) *elav>rAUUCU* to *elav>+*	Human Orthologue	Function
Modifier67.2	mod(mdg4)	CG3249	**-0.75**	**-0.52**	AKAP1	Regulation of chromatin assembly/disassembly, regulation of apoptosis
CG8511	Cpr49Ag	CG8511	**-0.54**	**-0.52**		Component of cuticle
CG13484	CG13484		**-0.62**	**-0.52**		
DN-cadh-like	CadN2	CG7527	**-0.82**	**-0.52**		Cell adhesion
meiotic recombination 11	mre11	CG16928	**-0.85**	**-0.52**	MRE11A	Nucleotide excision repair
CG16800	CG16800	CG16800	**-0.89**	**-0.52**		
VEGF-related factor 2	Pvf2	CG13780	**0.52**	**-0.51**		Hemocyte migration
CG34114	CG34114	CG34114	**-0.55**	**-0.51**		
CG14034	CG14034	CG14034	**-0.50**	**-0.51**		Phospholipid metabolism
***Drosophila* allatostatin C receptor 1**	star1	CG7285	**-0.57**	**-0.51**	SSTR5	G-protein coupled receptor protein signalling
CG10265	CG10265	CG10265	**-0.64**	**-0.51**		
scrambled	sced	CG3273	**-0.51**	**-0.51**		Actin filament re-organisation during cell cycle
CG4040 /// **CG33224**	CG33224 /// CG4040	CG42388	**0.53**	**0.51**		
ninjurin A	NijA	CG6449	**0.63**	**0.51**	NINJ1	Axon guidance, cell adhesion
bves	bves	CG32513	**0.72**	**0.52**	BVES	Cell adhesion
miple2	miple2	CG18321	**0.71**	**0.52**		Growth factor activity
Epidermal stripes and patches	Esp	CG7005	**0.67**	**0.52**		Sulphate transmembrane transporter
Matrix metallo-proteinase 1	Mmp1	CG4859	**0.95**	**0.52**	MMP14	Metallo-endopeptidase, cell adhesion
NtR	NtR	CG6698	**0.50**	**0.53**		Ligand-gated ion channel
CG33329	Sp212	CG33329	**0.65**	**0.53**		Serine-type endopeptidase
Immune induced molecule 10	IM10	CG18279	**0.86**	**0.53**		Toll-signalling pathway, immune response
CG17264	CG17264	CG17264	**0.64**	**0.53**		

Gene Title	Gene Symbol	Ensembl	log2(ratio) *elav>rAUUCU* to *elav>rCAA*	log2(ratio) *elav>rAUUCU* to *elav>+*	Human Orthologue	Function
CG11899	CG11899	CG11899	**0.59**	**0.53**	PSAT1	Pyridoxine biosynthesis
CG3091	CG3091	CG3091	**0.57**	**0.53**		Transporter activity
fragment K	alpha-Est10	CG1131	**0.56**	**0.54**		Carboxyl-esterase
CG31778	CG31778	CG31778	**0.57**	**0.54**		Serine-type endopeptidase inhibitor
***Drosophila* cold acclimation gene**	smp-30	CG7390	**0.83**	**0.54**		Lipid metabolism
CG7840	CG7840	CG7840	**0.50**	**0.55**		
Odorant-binding protein 56a	Obp56a	CG11797	**0.52**	**0.55**		Sensory perception of chemical stimulus
CG6435	CG6435	CG6435	**0.56**	**0.55**		Lysozyme activity,antibacterial
Cytochrome P450 related BF6-2	Cyp6a22	CG10240	**0.65**	**0.55**	CYP3A5	Electron carrier activity
CG6294	CG6294	CG6294	**0.59**	**0.56**		metallopeptidase
Ugt86Da	Ugt86Da	CG18578	**0.54**	**0.56**	UGT2B10	Lipid transporter activity
net	net	CG11450	**0.67**	**0.56**	ATOH8	Transcription factor activity
CG15828	CG15828	CG15828	**0.53**	**0.56**		
homogentisate 1,2-dioxygenase	hgo	CG4779	**0.74**	**0.56**	HGD	L-phenylalanine/tyrosine catabolism
CG32695	CG32695	CG32695	**0.60**	**0.56**		
CG9312	CG9312	CG9312	**0.74**	**0.57**		
CG14117	CG14117	CG14117	**0.70**	**0.57**		
Glutathione S transferase D9	GstD9	CG10091	**0.51**	**0.57**		Glutathione transferase
glutathione-dependent formaldehyde dehydrogenase	Fdh	CG6598	**0.50**	**0.57**	ADH5	Alcohol metabolism, oxidation
Spaetzle	spz	CG6134	**0.61**	**0.57**		Immune response, motor axon guidance
CG2938	CG2938	CG2938	**0.50**	**0.58**		
CG17350	CG17350	CG17350	**0.53**	**0.58**		

Gene Title	Gene Symbol	Ensembl	log2(ratio) *elav>rAUUCU* to *elav>rCAA*	log2(ratio) *elav>rAUUCU* to *elav>+*	Human Orthologue	Function
CG17928	CG17928	CG17928	**0.57**	**0.58**		Fatty acid biosynthetic process, oxidation
Cyp6a23	Cyp6a23	CG10242	**0.61**	**0.58**		Electron carrier activity
predicted gene W	CG31217	CG31217	**0.52**	**0.59**		Innate immune response
CG18107	CG18107	CG18107	**-1.02**	**0.59**		
CG18473	CG18473	CG18473	**0.51**	**0.6**	PTER	Aryldialkylphosphatase activity, ester hydrolase activity
sterol carrier protein X-related thiolase	ScpX	CG17597	**0.53**	**0.6**	SCP2	Acetyl-CoA C-acyltransferase activity, phospholipid transport
CG6113	CG6113	CG6113	**0.77**	**0.6**	LIPA	Triglyceride lipase
CG18249	CG18249	CG18249	**0.59**	**0.6**		
Serine Protease 2	Ser7	CG2045	**0.68**	**0.6**		Serine-type endopeptidase
CG18547	CG18547	CG18547	**0.68**	**0.61**		Oxidoreductase activity
CG9377	CG9377	CG9377	**-0.63**	**0.61**		Serine-type endopeptidase
CG12262	CG12262	CG12262	**0.60**	**0.61**	ACADM	Fatty acid beta-oxidation
CG13833	CG13833	CG13833	**0.70**	**0.61**		Oxidoreductase activity
Cytochrome P450-9c1	Cyp9c1	CG3616	**0.84**	**0.61**	CYP3A4	Electron carrier activity
CG8586	CG8586	CG8586	**0.54**	**0.62**		Serine-type endopeptidase
selenocysteine methyl-transferase	CG10621	CG10621	**0.61**	**0.62**	MTR	Selenocysteine methyltransferase activity
CG3663	CG3663	CG3663	**0.65**	**0.62**		
CG31436	CG31436	CG31436	**0.68**	**0.62**		
CG18302	CG18302	CG18302	**0.72**	**0.62**	LIPA	Lipid metabolism
brahma associated protein 60 kDa	Bap60	CG4303	**0.59**	**0.62**	SMARCD1	Dendrite and muscle development, transcription factor activity
fly plexin a	Cyp9b2	CG4486	**0.61**	**0.62**	CYP3A4	Electron carrier activity
Cyp6a13	Cyp6a13	CG2397	**0.62**	**0.63**	CYP3A4	Electron carrier activity
GIP-like	Gip	CG2227	**0.52**	**0.64**	HYI	Hydroxy-pyruvate isomerase

Gene Title	Gene Symbol	Ensembl	log2(ratio) *elav>rAUUCU* to *elav>rCAA*	log2(ratio) *elav>rAUUCU* to *elav>+*	Human Orthologue	Function
CG5222	CG5222	CG5222	**0.68**	**0.64**	RC74	mRNA cleavage and polyadenylation
CG14692	CG14692	CG14692	**0.60**	**0.64**		cAMP-dependent protein kinase regulation
CG10575	CG10575	CG10575	**0.94**	**0.64**	COASY	Co-enzyme A biosynthesis
CG9631	CG9631	CG9631	**0.50**	**0.64**		Serine-type endopeptidase
CG5707	CG5707	CG5707	**0.64**	**0.64**		
CG16704	CG16704	CG16704	**0.62**	**0.64**		Serine-type endopeptidase inhibitor activity
CG5697	CG5697	CG5697	**0.93**	**0.64**		
Glutathione S transferase E4	GstE4	CG17525	**1.04**	**0.64**		Glutathione transferase
CG16926	CG16926	CG16926	**0.79**	**0.64**		
Ugt86Di	Ugt86Di	CG6658	**0.79**	**0.64**	UGT2B4	glucoronosyltransferase
CG10026	CG10026	CG10026	**0.64**	**0.65**		Vitamin E binding, transporter
Nedd2-like caspase	Nc	CG8091	**0.97**	**0.65**	CASP7	Induction of apoptosis, CNS and eye development
CG6421	CG6421	CG6421	**0.55**	**0.65**		Lysozyme activity, antibacterial
CG13912	CG13912	CG13912	**0.60**	**0.66**		
CG13283	CG13283	CG13283	**0.89**	**0.66**		metalloendopeptidase
CG9691	CG9691	CG9691	**0.67**	**0.66**		
rhythmically expressed gene 2	Reg-2	CG3200	**0.79**	**0.66**	HDH3	Phosphoglycolate phosphatase
CG31769	CG31769	CG31769	**0.90**	**0.66**		
CG8317	CG8317	CG8317	**0.79**	**0.67**		
germ cell-expressed bHLH-PAS	gce	CG6211	**0.59**	**0.67**	ARNT	Transcription factor activity
CG11919	CG11919	CG11919	**0.63**	**0.67**	PEX6	Nucleoside triphosphatase
Juvenile hormone esterase duplication	Jhedup	CG8424	**0.63**	**0.67**		Carboxylesterase activity

Gene Title	Gene Symbol	Ensembl	log2(ratio) *elav>rAUUCU* to *elav>rCAA*	log2(ratio) *elav>rAUUCU* to *elav>+*	Human Orthologue	Function
CG30503	CG30503		**0.61**	**0.67**	PLA2G3	Phospholipase activity
Transcription unit B	CG14630	CG14630	**0.63**	**0.68**		Oxidation regulation
regucalcin	regucalcin	CG1803	**0.51**	**0.68**	RGN	
CG31548	CG31548	CG31548	**0.58**	**0.68**		Oxidation regulation
Gram-negative bacteria binding protein 3	GNBP3	CG5008	**0.52**	**0.68**		Immune response
CG7339	CG7339	CG7339	**0.58**	**0.68**	POLR3H	Transcriptional regulation
lethal (2) k10201	l(2)k10201	CG13951	**0.70**	**0.69**		
CG30002	CG30002	CG30002	**0.55**	**0.69**		Serine-type endopeptidase
AIR-carboxylase-SAICAR synthetase	ade5	CG3989	**0.57**	**0.69**	PAICS	Inosine monophosphate biosynthesis
Pheromone-binding protein-related protein 4	Pbprp4	CG1176	**0.58**	**0.69**		Sensory perception of chemical stimulus
CG31414	CG31414	CG31414	**0.55**	**0.69**		
cathepsin B	CG10992	CG10992	**0.66**	**0.69**	CTSB	Autophagic cell death, cysteine-type endopeptidase
Imaginal disc growth factor1	Idgf1	CG4472	**0.74**	**0.69**		Growth factor activity
yippee interacting protein 2	yip2	CG4600	**0.54**	**0.69**	ACAA2	Fatty acid beta-oxidation
CG6426	CG6426	CG6426	**0.58**	**0.69**		Lysozyme activity, immune response
CG13641	CG13641	CG13641	**0.61**	**0.69**		
Immune induced molecule 23	IM23	CG15066	**1.11**	**0.69**		Toll signalling pathway, immune response
CG11315	CG11315	CG11315	**0.56**	**0.7**		
secretory Phospholipase A2	sPLA2	CG11124	**0.53**	**0.7**	PLA2G3	Phospholipase activity
CG13845	CG34376	CG34376	**0.57**	**0.7**		

Gene Title	Gene Symbol	Ensembl	log2(ratio) *elav>rAUUCU* to *elav>rCAA*	log2(ratio) *elav>rAUUCU* to *elav>+*	Human Orthologue	Function
CG15414	CG15414	CG15414	**0.63**	**0.7**		
CR30029	CR30029	CR30029	**0.64**	**0.7**		
CG1889	CG1889	CG1889	**0.64**	**0.7**		
no-on-transient A	nonA	CG4211	**0.84**	**0.71**	SFPQ	Nuclear mRNA splicing
Cytochrome b5-related	Cyt-b5-r	CG13279	**0.60**	**0.71**	FADS1	Fatty acid biosynthesis, oxidation regulation
CG2004	CG2004	CG2004	**0.82**	**0.71**		
MSP protein	CG33523	CG33523	**0.65**	**0.71**		
CG2444	CG2444	CG2444	**0.72**	**0.72**		
CG1397	CG1397	CG1397	**0.84**	**0.72**		Cuticle synthesis
CG10352	CG10352	CG10352	**0.80**	**0.72**		
CG11909	CG11909	CG11909	**0.84**	**0.73**	KIAA1161	Carbohydrate metabolism
serine pyruvate amino-transferase	Spat	CG3926	**0.72**	**0.73**	AGXT	Glyoxylate catabolic process
CG31292 /// **CG3303**	CG31292 /// CG3303	CG31292	**0.72**	**0.73**		Serine-type endopeptidase
CG9519	CG9519	CG9519	**0.65**	**0.73**	CHDH	Choline dehydrogenase activity
Drosomycin B	dro5	CG10812	**0.79**	**0.74**		Defense response
Pugilist	pug	CG4067	**0.64**	**0.74**		Folic acid synthesis
CG14935	CG14935	CG14935	**0.85**	**0.75**	SLC3A1	Alpha glucosidase, carbohydrate metabolism
CYP6-like	Cyp6g1	CG8453	**0.98**	**0.75**	CYP3A7	Electron carrier activity
CG14872	CG14872		**0.68**	**0.75**	CG14872	
drosomycin-F	dro4	CG32282	**1.02**	**0.75**		Defense response
CG2233	CG2233	CG2233	**0.67**	**0.75**		
neither inactivation nor afterpotential D	ninaD	CG31783	**1.12**	**0.75**		Cell adhesion, rhodopsin biosynthesis
Maternal transcript 89Bb	Mat89Bb	CG6814	**0.68**	**0.76**		
CG9396	CG9396	CG9396	**0.66**	**0.76**		
CG9455	CG9455	CG9455	**0.86**	**0.76**		Serine-type endopeptidase

Gene Title	Gene Symbol	Ensembl	log2(ratio) *elav>rAUUCU* to *elav>rCAA*	log2(ratio) *elav>rAUUCU* to *elav>+*	Human Orthologue	Function
CG5288	CG5288	CG5288	**0.58**	**0.76**	GALK2	Galactokinase, carbohydrate metabolism
CG4576	CG4576	CG4576	**0.63**	**0.77**		Amino acyl-transferase
CG11550	CG11550	CG11550	**0.57**	**0.77**		
CG30026	CG30026	CG30026	**0.58**	**0.78**		
Angiotensin-converting enzyme-related	Acer	CG10593	**0.65**	**0.78**		Heart development, peptidyl-dipeptidase
CG16713	CG16713	CG16713	**0.67**	**0.79**		Serine-type endopeptidase inhibitor
CG31643	CG31643	CG31643	**0.56**	**0.79**	FASTKD1	Regulation of apoptosis, ATP binding
CG1791	CG1791	CG1791	**0.54**	**0.79**	TNR	Signal transduction
CG12713	CG12713	CG12713	**0.70**	**0.79**		
serpin 1	Spn1	CG9456	**0.81**	**0.79**	SERPINB4	Serine-type endopeptidase inhibitor
Cyp28d1	Cyp28d1	CG10833	**0.59**	**0.79**	CYP3A4	Electron carrier activity
serpin43Ac	nec	CG1857	**0.55**	**0.79**		Serine-type endopeptidase inhibitor, immune response
Myoglianin	myoglianin	CG1838	**0.62**	**0.79**	GDF8	Growth factor activity
CG2118	CG2118	CG2118	**0.84**	**0.8**	MCCC1	Leucine metabolism
CG12269	CG12269	CG12269	**0.52**	**0.82**		Sterol carrier
CG15281	CG15281	CG15281	**0.52**	**0.82**		
CG3603	CG3603	CG3603	**0.54**	**0.82**	HSD17B8	Oxidoreductase activity
Mth-like 2	mthl2	CG17795	**0.97**	**0.82**		G-protein coupled receptor signalling, response to stress
CG9689	CG9689	CG9689	**0.93**	**0.82**		
virus-induced 1	vir-1	CG31764	**0.66**	**0.83**		Defense response to virus
CG4716	CG4716	CG4716	**0.56**	**0.83**		Methylenetetrahydrafolate dehydrogenase activity
Cyp309a2	Cyp309a2	CG18559	**0.63**	**0.83**		Electron carrier activity
CG11395	CG11395	CG11395	**0.73**	**0.85**		
lush	lush	CG8807	**0.52**	**0.87**		Sensory perception of chemical stimulus

Gene Title	Gene Symbol	Ensembl	log2(ratio) *elav>rAUUCU* to *elav>rCAA*	log2(ratio) *elav>rAUUCU* to *elav>+*	Human Orthologue	Function
CG3301	CG3301	CG3301	**0.68**	**0.87**	MGC4172	Oxidoreductase activity
CG7219	CG7219	CG7219	**0.60**	**0.87**		Serine-type endopeptidase inhibitor
CG4721	CG4721	CG4721	**0.89**	**0.88**		metalloendopeptidase
beta-galactosidase	Ect3	CG3132	**0.88**	**0.88**	GLB1L	Autophagic cell death
PGRP-SD	PGRP-SD	CG7496	**0.71**	**0.89**		
Cytochrome P450-4e1	Cyp4e1	CG2062	**0.65**	**0.89**		
acid DNase	DNaseII	CG7780	**0.53**	**0.89**	DNASE2	Deoxyribonuclease
Immune induced molecule 1	IM1	CG18108	**0.99**	**0.89**		
CG15293	CG15293	CG15293	**0.66**	**0.9**		
CG14629	CG14629	CG14629	**0.53**	**0.9**		
CG32613	CG32613	CG32613	**0.66**	**0.9**		Immune response, polysaccharide binding
CG5791	CG5791	CG5791	**0.81**	**0.91**		
CG1468	CG1468	CG1468	**0.63**	**0.91**		
Transferrin	Tsf1	CG6186	**0.98**	**0.91**		Iron ion homeostasis
Immune induced molecule 4	IM4	CG15231	**0.72**	**0.92**		Immune response
CG14400	CG14400	CG14400	**0.75**	**0.92**		
CG15067	CG15067	CG15067	**0.89**	**0.92**		
sex-specific enzyme 2	sxe2	CG4979	**0.66**	**0.93**		Lipid metabolism, phosphatidylserine-specific
takeout	to	CG11853	**0.66**	**0.94**		Circadian rhythm and feeding behaviour
CG4408	CG4408	CG4408	**0.73**	**0.94**		Metallocarboxypeptidase activity
Lipid storage droplet-1	Lsd-1	CG10374	**0.72**	**0.94**		Lipid storage
CG3246	CG3246	CG3246	**0.50**	**0.94**		
Odorant-binding protein 59a	Obp59a	CG13517	**0.71**	**0.95**		Sensory perception of chemical stimulus

Gene Title	Gene Symbol	Ensembl	log2(ratio) *elav>rAUUCU* to *elav>rCAA*	log2(ratio) *elav>rAUUCU* to *elav>+*	Human Orthologue	Function
CG17324	CG17324	CG17324	**0.79**	**0.95**	UGT1A3	UDP-glycosyltransferase activity
CG11842	CG11842	CG11842	**0.86**	**0.95**		Serine-type endopeptidase
CG17325	CG17325	CG17325	**0.95**	**0.96**		
CG34020	CG34020	CG34020	**1.08**	**0.96**		
hemolectin	Hml	CG7002	**0.57**	**0.97**	VWF	Hemostasis, cell adhesion
antdh	antdh	CG1386	**0.50**	**0.97**	MGC4172	Carbonyl reductase
CG14567	CG14567	CG14567	**0.54**	**0.98**		
CG7299	CG7299	CG7299	**0.90**	**0.98**		
CG5126	CG5126	CG5126	**0.51**	**0.99**		
CG6206	CG6206	CG6206	**0.69**	**0.99**	MAN2B1	Alpha-mannosidase activity
Ance-4	Ance-4	CG8196	**0.84**	**1**		Peptidyl dipeptidase
Lectin24Db	lectin-24Db	CG2958	**0.75**	**1**		Galactose binding
***Drosophila* cold acclimation gene**	smp-30	CG7390	**0.72**	**1.02**		
cuticle cluster 2	Ccp84Ag	CG2342	**0.56**	**1.02**		Component of cuticle
CG15065	CG15065	CG15065	**1.16**	**1.02**		
Turandot X	TotX	CG31193	**1.26**	**1.02**		Stress response
CG10357	CG10357	CG10357	**1.09**	**1.03**		Lipid metabolism, triglyceride lipase
CG1667	CG1667	CG1667	**1.16**	**1.03**		
bangles and beads	bnb	CG7088	**0.88**	**1.04**		gliogenesis
Thiolester containing protein II	TepII	CG7052	**0.83**	**1.04**	CD109	Antibacterial humoral response
CG14528	CG14528	CG14528	**0.91**	**1.04**		Metalloendopeptidase
Iris	Iris	CG4715	**0.88**	**1.05**		
CG12656	CG12656	CG12656	**0.77**	**1.05**		
CG17189	CG17189	CG17189	**0.79**	**1.06**		
CG31839	CG31839	CG31839	**0.79**	**1.06**	FLJ14712	Mesoderm development

Gene Title	Gene Symbol	Ensembl	log2(ratio) *elav>rAUUCU* to *elav>rCAA*	log2(ratio) *elav>rAUUCU* to *elav>+*	Human Orthologue	Function
CG4335	CG4335	CG4335	**1.29**	**1.06**	TMLHE	Oxidation reduction
CG17777	CG17777	CG17777	**0.96**	**1.06**		
CG4250	CG4250	CG4250	**0.67**	**1.07**		
Cecropin	CecB	CG1878	**0.79**	**1.07**		Antibacterial humoral response
CG3588	CG3588	CG3588	**0.78**	**1.07**		
CG10799	CG10799	CG10799	**1.03**	**1.08**		
female-specific independent of transformer	fit	CG17820	**0.53**	**1.09**		
CG13086	CG13086	CG13086	**0.89**	**1.09**		
CG8147	CG8147	CG8147	**0.55**	**1.09**		Alkaline phosphatase
CG14872	CG14872	CG14872	**0.97**	**1.09**		
CG15282	CG15282	CG15282	**0.87**	**1.09**		
CG11314	CG11314	CG11314	**0.59**	**1.1**		Mesoderm development
CG4950	CG4950	CG4950	**0.69**	**1.1**		
semmelweis	PGRP-SA	CG11709	**0.63**	**1.11**	PGLYRP3	Immune response
CG9511	CG9511	CG42370	**0.80**	**1.13**		
CG5428	CG5428	CG5428	**1.28**	**1.13**		sulfotransferase
CG18067	CG18067	CG18067	**0.81**	**1.17**		3',5'-cyclic nucleotide phosphodiesterase, signalling
CG2540	CG2540	CG2540	**1.00**	**1.18**		
Ance-2	Ance-2	CG16869	**1.24**	**1.2**		Peptidyl dipeptidase
PGRP-SB1	PGRP-SB1	CG9681	**0.57**	**1.2**		Immune response
CG7526	CG7526	CG7526	**0.57**	**1.21**		Calcium ion binding
CG6188	CG6188	CG6188	**1.34**	**1.23**	GNMT	Methionine metabolism
CG15068	CG15068	CG15068	**1.38**	**1.25**		
Turandot C	TotC	CG31508	**3.19**	**1.27**		Stress response
CG4716	CG4716	CG4716	**0.99**	**1.29**		
CG16836	CG16836	CG16836	**1.04**	**1.3**		
CG14495	CG14495	CG14495	**0.86**	**1.3**		

Gene Title	Gene Symbol	Ensembl	log2(ratio) *elav>rAUUCU* to *elav>rCAA*	log2(ratio) *elav>rAUUCU* to *elav>+*	Human Orthologue	Function
ornithine decarboxylase	Odc1	CG8721	**0.65**	**1.32**	ODC1	Polyamine biosynthesis
CG6067	CG6067	CG6067	**0.65**	**1.33**		Serine-type endopeptidase
CG32641 /// **CG32640**	CG32640 /// CG32641	CG32641	**1.44**	**1.34**		Protein folding
turn on sex-specificity	Obp99b	CG7592	**1.15**	**1.39**		Autophagic cell death, sensory perception of chemical stimulus
CG13335	CG13335	CG13335	**1.01**	**1.4**		
CG5966	CG5966	CG5966	**1.49**	**1.44**	PNLIP	Triglyceride lipase
Cytochrome P450-4e3	Cyp4e3	CG4105	**1.35**	**1.45**		Electron carrier activity
lethal (1) G0237	l(1)G0237		**-1.06**	**1.55**		
Turandot	TotA	CG31509	**2.66**	**1.55**		Response to stress
CG13704	CG13704	CG13704	**1.08**	**1.56**		
CG3699	CG3699	CG3699	**0.52**	**1.62**		Oxidation reduction
Vago	Vago	CG2081	**0.96**	**1.64**		
CG1732	CG1732	CG1732	**0.68**	**1.66**	SLC6A1	GABA:sodium symporter, neurotransmitter transporter
attacin	AttA	CG10146	**1.16**	**1.76**		Antibacterial humoral response
CG13422	CG13422	CG13422	**1.26**	**1.79**		Defense response
Amylase	Amy-d /// Amy-p	CG17876	**1.49**	**1.85**	AMY2A	Carbohydrate metabolism
metchnikowin	Mtk	CG8175	**1.52**	**2**		Antibacterial humoral response
Cytochrome P450 A1	Cyp4g1	CG3972	**2.02**	**2.05**		Lipid metabolic process, oxidation reduction
attacin	AttC	CG4740	**1.38**	**2.05**		Antibacterial humoral response
Turandot M	TotM	CG14027	**3.77**	**3.07**		Response to stress
white	w	CG2759	**1.30**	**3.59**	ABCG2	Eye pigment biosynthesis

Table C2: Genes altered in *elav>rAUUCU* and *elav>rCAG* flies compared to *elav>rCAA*. Selected for Log2(ratio) >0.5 or <-0.5 for *elav>rAUUCU* and *elav>rCAG* compared to *elav>rCAA*, P<0.05 for *elav>rCAG.*

Gene Title	Gene Symbol	Ensembl	Log2(ratio) *elav>rCAG* to *elav>rCAA*	Log2(ratio) *elav>rAUUCU* to *elav>rCAA*	Human orthologue	Function
Staufen	stau	CG5753	**-0.66** **P=0.029**	**-1.10**	STAU2	double-stranded RNA binding, RNA localisation involved in cell fate determination
CG32736	CG32736	CG32736	**-0.51** **P=0.025**	**-0.77**		
***Drosophila* insulin-like peptide 5**	Ilp5	CG33273	**-0.59** **P=0.001**	**-0.67**		Insulin signalling
oocyte maintenance defects	omd	CG9591	**-0.66** **P=0.001**	**-0.64**	INTS5	
gliotactin	Gli	CG3903	**0.60** **P=0.040**	**0.54**		septate junction formation, role in polarisation of cells
CG9400	CG9400	CG9400	**0.84** **P=0.042**	**0.55**		
CG9079	Cpr47Ea	CG9079	**0.51** **P=0.021**	**0.56**		
CG34104	CG34104	CG34104	**0.58** **P=2.87E-5**	**0.57**		Signal transduction, GTPase activity
G protein gamma30A	Ggamma30A	CG3694	**0.83** **P=0.016**	**0.57**	GNG13	phototransduction
cut	ct	CG11387	**0.65** **P=0.025**	**0.57**	CUTL1	transcriptional regulation, regulation of dendrite morphogenesis
Inscuteable	insc	CG11312	**0.89** **P=0.025**	**0.60**	INSC	cytoskeletal adaptor, protein and RNA localisation, localisation is dynein dependent
CG13065	CG13065	CG13065	**0.64** **P=0.046**	**0.61**	LSM7	pre-mRNA processing
CG10632	CG10632	CG10632	**0.67** **P=0.041**	**0.61**		
CG13277	CG13277	CG13277	**0.60** **P=0.045**	**0.70**		
CG32850	CG32850	CG32850	**0.80** **P=0.017**	**0.75**	RNF11	

Gene Title	Gene Symbol	Ensembl	Log2(ratio) *elav>rCAG* to *elav>rCAA*	Log2(ratio) *elav>rAUUCU* to *elav>rCAA*	Human orthologue	Function
CG7744	CG7744	CG7744	**0.87** **P=0.017**	**0.82**		
Odorant-binding protein 56f	Obp56f	CG30450	**0.58** **P=5.94E-5**	**0.82**		Sensory perception of chemical stimulus
CG1397	CG1397	CG1397	**0.59** **P=0.028**	**0.84**		
CG12998	CG12998	CG12998	**0.92** **P=0.036**	**0.88**		
CG14528	CG14528	CG14528	**0.72** **P=0.048**	**0.91**		
Matrix metallo-proteinase 1	Mmp1	CG4859	**0.62** **P=0.041**	**0.95**	MMP14	Role in ECM regulation, cell adhesion
CG4525	CG4525	CG4525	**0.63** **P=0.044**	**0.99**	TTC26	cilium assembly
CG12877	CG12877	CG12877	**0.62** **P=0.030**	**1.03**	REXO1	Transcriptional elongation, RNA exonuclease activity
CG3099	CG3099	CG3099	**0.79** **P=0.024**	**1.03**	HECW2	Ubiquitin protein ligase

Table C3: Genes altered in *elav>rAUUCU* and *elav>rCUG* flies compared to *elav>rCAA*. Selected for Log2(ratio) >0.5 or <-0.5 for *elav>rAUUCU* and *elav>rCUG* compared to *elav>rCAA*, P<0.05 for *elav>rCUG.*

Gene Title	Gene Symbol	Ensembl	Log2(ratio) *elav>rCUG* to *elav>rCAA*	Log2(ratio) *elav>rAUUCU* to *elav>rCAA*	Human orthologue	Function
CG6416	CG6416	CG6416	**-0.92** **P=0.027**	**-1.62**		mesoderm development (Zasp66) - cytoskeletal remodelling
CG17290	CG17290	CG17290	**-0.66** **P=0.031**	**-1.60**		
CG18107	CG18107	CG18107	**0.73** **P=0.016**	**-1.02**		
CG4161	CG4161	CG4161	**-1.34** **P=0.010**	**-0.91**		
CG13416	CG13416	CG13416	**-0.67** **P=0.031**	**-0.78**		
defective transmitter release	dtr	CG31623	**-1.02** **P=0.030**	**-0.75**		synaptic transmission
CG6425	CG6425	CG6425	**-0.73** **P=0.035**	**-0.73**		
Complement-ation group C	Mef2	CG1429	**-0.90** **P=0.007**	**-0.71**	MEF2C	transcription factor, muscle development
Grip71	Grip71	CG10346	**-0.82** **P=0.001**	**-0.67**		gamma-tubulin binding, cell-cycle regulation
DNA polymerase epsilon	DNApol-epsilon	CG6768	**-0.55** **P=0.007**	**-0.66**	POLE	DNA-dependent DNA polymerase
Cyp6a19	Cyp6a19	CG10243	**-1.01** **P=0.010**	**-0.65**	CYP3A7	electron carrier
CG17177	CG17177	CG17177	**-0.78** **P=0.048**	**-0.63**		
CG13300	CG13300	CG13300	**-0.59** **P=0.014**	**-0.59**		
CG7906	CG7906	CG7906	**0.95** **P=0.038**	**-0.59**		
CG9200	Atac1	CG9200	**-0.62** **P=0.008**	**-0.58**	ZZZ3	histone acetylation

Gene Title	Gene Symbol	Ensembl	Log2(ratio) *elav>rCUG* to *elav>rCAA*	Log2(ratio) *elav>rAUUCU* to *elav>rCAA*	Human orthologue	Function
CG9817	CG9817	CG9817	**-0.82** **P=0.023**	**-0.57**		
CG1961	CG1961	CG1961	**-0.68** **P=0.040**	**-0.56**		nucleotidase
CG10320	CG10320	CG10320	**-0.63** **P=0.008**	**-0.55**	NDUFB3	NADH dehydrogenase, electron transport chain, RNA import into nucleus
Plum	bw	CG17632	**-0.51** **P=0.018**	**-0.52**		eye pigment precursor transport activity
skpB	skpB	CG8881	**-0.54** **P=0.040**	**-0.52**	SKP1A	ubiqutin-dependent protein catabolism, cell cycle?
CG14034	CG14034	CG14034	**-0.78** **P=0.024**	**-0.50**		phospholipase activity, lipid metabolism
CG12269	CG12269	CG12269	**0.95** **P=0.021**	**0.52**		sterol transporter
yippee interacting protein 2	yip2	CG4600	**0.64** **P=0.015**	**0.54**	ACAA2	lipid metabolism, fatty acid beta oxidation, mitochondrial
CG9400	CG9400	CG9400	**0.97** **P=0.014**	**0.55**		peptidase inhibitor
CG9079	Cpr47Ea	CG9079	**1.34** **P=0.003**	**0.56**		structural component of cuticle
CG11400	CG11400	CG11400	**0.55** **P=0.017**	**0.56**		
fragment K	alpha-Est10	CG1131	**0.69** **P=0.041**	**0.56**	CES2	carboxyl esterase
CG17928	CG17928	CG17928	**0.62** **P=0.013**	**0.57**		
CG11550	CG11550	CG11550	**0.59** **P=0.045**	**0.57**		
CG34104	CG34104	CG34104	**0.80** **P=0.004**	**0.57**		microtubule based G-protein coupled signal transduction
CG30026	CG30026	CG30026	**0.76** **P=0.006**	**0.58**		
CG18249	CG18249	CG18249	**0.55** **P=0.026**	**0.59**		

Gene Title	Gene Symbol	Ensembl	Log2(ratio) *elav>rCUG* to *elav>rCAA*	Log2(ratio) *elav>rAUUCU* to *elav>rCAA*	Human orthologue	Function
synaptobrevin	Syb	CG12210	**0.63** **P=0.025**	**0.60**	VAMP1	neurotransmitter secretion
Inscuteable	insc	CG11312	**0.55** **P=0.050**	**0.60**	INSC	cytoskeletal adaptor, protein and RNA localisation, localisation is dynein dependent
CG12262	CG12262	CG12262	**0.64** **P=0.027**	**0.60**	ACADM	lipid metabolism, fatty acid beta oxidation, mitochondrial
CG4576	CG4576	CG4576	**0.53** **P=0.042**	**0.63**		acyl transferase
Juvenile hormone esterase duplication	Jhedup	CG8424	**0.83** **P=0.020**	**0.63**		degradation of juvenile hormone
CG15414	CG15414	CG15414	**0.64** **P=0.007**	**0.63**		
virus-induced 1	vir-1	CG31764	**0.80** **P=0.023**	**0.66**		upregulated in response to viral infection, responsive to RNAi pathway components
CG15293	CG15293	CG15293	**0.83** **P=0.007**	**0.66**		
CG9396	CG9396	CG9396	**0.86** **P=0.039**	**0.66**	BRP44	
CG9691	CG9691	CG9691	**0.7** **P=0.018**	**0.67**		
CG14872	CG14872	CG14872	**0.81** **P=0.031**	**0.68**		
CG18302	CG18302	CG18302	**0.61** **P=0.007**	**0.72**	LIPA/LIPF	lipid metabolism
lethal (2) 09851	l(2)09851	CG12792	**0.89** **P=0.003**	**0.72**	GRWD1	glutamine rich, ribosome biogenesis
CG6340	CG6340	CG6340	**0.58** **P=0.036**	**0.76**		
CG12656	CG12656	CG12656	**0.67** **P=0.026**	**0.77**		
CG3588	CG3588	CG3588	**0.86** **P=0.044**	**0.78**		

Gene Title	Gene Symbol	Ensembl	Log2(ratio) *elav>rCUG* to *elav>rCAA*	Log2(ratio) *elav>rAUUCU* to *elav>rCAA*	Human orthologue	Function
CG16926	CG16926	CG16926	**1.23** **P=0.006**	**0.79**		
CG17189	CG17189	CG17189	**0.60** **P=0.041**	**0.79**		
CG18067	CG18067	CG18067	**0.76** **P=0.006**	**0.81**		
serpin 1	Spn1	CG9456	**0.73** **P=0.031**	**0.81**	SERPINB4	serine protease inhibitor
Limpet	Lmpt	CG32171	**0.69** **P=0.019**	**0.83**	FHL2	transcription factor, heart development
beta-galactosidase	Ect3	CG3132	**0.82** **P=0.039**	**0.88**	GLBL1	autophagic cell death
CG12998	CG12998	CG12998	**0.71** **P=0.023**	**0.88**		
bangles and beads	bnb	CG7088	**0.82** **P=0.023**	**0.88**		gliogenesis
CG4721	CG4721	CG4721	**0.87** **P=0.046**	**0.89**		Metallo-endopeptidase
CG7299	CG7299	CG7299	**1.22** **P=0.032**	**0.90**		
CG14528	CG14528	CG14528	**0.84** **P=0.015**	**0.91**		Metallo-endopeptidase
CG9689	CG9689	CG9689	**1.18** **P=0.022**	**0.93**		
CG17777	CG17777	CG17777	**1.61** **P=0.009**	**0.96**		
Mth-like 2	mthl2	CG17795	**1.05** **P=0.001**	**0.97**		G-protein coupled receptor signalling pathway, extended lifespan
Transferrin	Tsf1	CG6186	**0.90** **P=0.048**	**0.98**		cellular iron ion homeostasis
CG16836	CG16836	CG16836	**1.09** **P=0.012**	**1.04**		
CG34020	CG34020	CG34020	**1.14** **P=0.049**	**1.08**		

Gene Title	Gene Symbol	Ensembl	Log2(ratio) *elav>rCUG* to *elav>rCAA*	Log2(ratio) *elav>rAUUCU* to *elav>rCAA*	Human orthologue	Function
CG5428	CG5428	CG5428	**1.30** **P=0.033**	**1.28**	SULT1E1	Sulfo-transferase
CG10191	CG10191	CG10191	**1.08** **P=0.039**	**1.32**	WDR51A	
CG6188	CG6188	CG6188	**1.40** **P=0.001**	**1.34**	GNMT	methionine metabolism
Cytochrome P450-4e3	Cyp4e3	CG4105	**1.57** **P=0.010**	**1.35**		electron carrier
CG32641 /// **CG32640**	CG32641	CG32641	**0.55** **P=0.048**	**1.44**		
CG5966	CG5966	CG5966	**1.92** **P=0.001**	**1.49**	PNLIP	triacylglycerol lipase
metchnikowin	Mtk	CG8175	**1.47** **P=0.009**	**1.52**		defense response
Turandot	TotA	CG31509	**3.14** **P=0.008**	**2.66**		stress response
Turandot C	TotC	CG31508	**4.16** **P=0.004**	**3.19**		stress response

Table C4: Genes altered in *elav>rAUUCU* and *elav>rCAG* flies compared to *elav>+*. Selected for Log2(ratio) >0.5 or <-0.5 for *elav>rAUUCU* and *elav>rCAG* compared to *elav>+*, P<0.05 for *elav>rCAG.*

Gene Title	Gene Symbol	Ensembl	Log2(ratio) *elav>rCAG* to *elav>rCAA*	Log2(ratio) *elav>rAUUCU* to *elav>rCAA*	Human orthologue	Function
Stellate orphon	Ste12DOR	CG32616	**-2.88 P=0.002**	**-2.33**		Spermatogenesis, protein kinase regulator
Odorant-binding protein 19c	Obp19c	CG15457	**-0.75 P=0.029**	**-0.9**		Sensory perception of chemical stimulus
CG32552	CG32552	CG32552	**-0.55 P=0.013**	**-0.87**		
CG13077	CG13077	CG13077	**-0.55 P=0.019**	**-0.85**	CYB561D2	
CG13117	CG13117	CG13117	**-0.53 P=0.015**	**-0.84**		
CG13895	CG13895	CG13895	**-0.62 P=0.016**	**-0.62**		
CG15545	CG15545	CG15545	**-0.52 P=0.029**	**-0.62**		
CG7031	CG7031	CG7031	**-0.52 P=0.022**	**-0.58**		
CG14959	CG14959	CG14959	**0.81 P=0.012**	**-0.56**		Chitin binding
CG11893	CG11893	CG11893	**-0.81 P=0.001**	**-0.52**		
Modifier67.2	mod(mdg4)	CG32491	**-0.76 P=0.012**	**-0.52**		Regulation of chromatin assembly
CG9686	CG9686	CG9686	**0.55 P=0.014**	**0.54**		
CG9186	CG9186	CG9186	**0.58 P=0.017**	**0.57**		
Immune induced molecule 2	IM2	CG18106	**0.75 P=0.017**	**0.66**		Immune response
CG4484	CG4484	CG4484	**0.54 P=0.039**	**0.73**	SLC45A1	Glucose trans-membrane transport
CG9657	CG9657	CG9657	**0.61 P=0.030**	**0.81**	SLC5A12	Sodium:iodide transporter
Hormone receptor-like in 38	Hr38	CG1864	**1.05 P=0.040**	**0.91**	NR4A1/ NGFI-B	Ligand-dependent nuclear receptor activity
CG14528	CG14528	CG14528	**0.85 P=0.026**	**1.04**		Metallo-endopeptidase
CG11825	CG11825	CG11825	**0.70 P=0.019**	**1.17**		
CG2540	CG2540	CG2540	**0.56 P=0.030**	**1.18**		
CG7526	CG7526	CG7526	**-1.88 P=0.024**	**1.21**		
CG14495	CG14495	CG14495	**1.04 P=0.002**	**1.3**		

Gene Title	Gene Symbol	Ensembl	Log2(ratio) *elav>rCAG* to *elav>rCAA*	Log2(ratio) *elav>rAUUCU* to *elav>rCAA*	Human orthologue	Function
CG9394	CG9394	CG9394	**1.00** **P=0.002**	**1.33**		Lipid metabolism, glycerol-phosphodiester phosphodiesterase
CG11910	CG11910	CG11910	**0.99** **P=0.047**	**1.41**		Insulin-like growth factor binding protein complex
CG13704	CG13704	CG13704	**0.73** **P=0.045**	**1.56**		
white	w	CG2759	**2.67** **P=0.002**	**3.59**	ABCG5	Eye pigment precursor transport, metabolic process

Table C5: Genes altered in *elav>rAUUCU* and *elav>rCUG* flies compared to *elav>+*. Selected for Log2(ratio) >0.5 or <-0.5 for *elav>rAUUCU* and *elav>rCUG* compared to *elav>+*, P<0.05 for *elav>rCUG*.

Gene Title	Gene Symbol	Ensembl	log2(ratio) *elav>rCUG* to *elav>+*	log2 (ratio) *elav>rAUUCU* to *elav>+*	Human Orthologue	Function
Stellate orphon	Ste12DOR	CG32616	**-2.00** **P=0.03**	**-2.33**		Spermatogenesis, protein kinase regulator
CG18646	CG18646	CG18646	**-0.56** **P=0.047**	**-1.55**		GTPase activity
CG16752	CG16752	CG16752	**-1.24** **P=0.006**	**-1.51**		Neuropeptide receptor activity
CG4078	CG4078	CG4078	**-0.52** **P=0.046**	**-1.44**	RTEL1	Nucleotide excision repair
CG40485	CG40485	CG40485	**-0.65** **P=0.003**	**-1.35**		Oxidoreductase activity
CG12825	CG12825	CG12825	**-0.59** **P=0.041**	**-1.03**		
CG40084	CG40084	CG40084	**-0.89** **P=0.011**	**-1.02**		
CG40188	CG40188	CG40188	**-0.76** **P=0.040**	**-0.94**		
CG13908	CG13908	CG13908	**-0.95** **P=0.012**	**-0.91**		
CG32552	CG32552	CG32552	**-0.86** **P=0.009**	**-0.87**		
CG13077	CG13077	CG13077	**-0.79** **P=0.027**	**-0.85**	CYB561D2	
CG13117	CG13117	CG13117	**-0.56** **P=0.009**	**-0.84**		
CG31781	CG31781	CG31781	**-0.82** **P=0.010**	**-0.81**		
sickie	sick	CG42589	**-0.81** **P=0.015**	**-0.79**		Immune response
CG5010	CG5010	CG5010	**-0.70** **P=0.026**	**-0.78**	CHCHD2	
CG31846	CG31846	CG31846	**-0.77** **P=0.027**	**-0.75**		

Gene Title	Gene Symbol	Ensembl	log2(ratio) *elav>rCUG* to *elav>+*	log2 (ratio) *elav>rAUUCU* to *elav>+*	Human Orthologue	Function
hu-li tai shao	hts	CG9325	**-0.62** **P=0.032**	**-0.74**	ADD1	Actin assembly, ring canal formation
Cyp12a4	Cyp12a4	CG6042	**-0.82** **P=0.050**	**-0.73**	CYP24A1	Electron carrier activity
CG5883	CG5883	CG5883	**-0.75** **P=0.005**	**-0.71**		Chitin metabolic process
CG31814	CG31814	CG31814	**-0.61** **P=0.014**	**-0.67**	HNT	
D-Titin	sls	CG1915	**-0.8** **P=0.029**	**-0.66**	TTN	Mesoderm development, myoblast fusion
Poly-glutamine tract binding protein 1	PQBP-1	CG31369	**-0.73** **P=0.013**	**-0.66**		
CG7330	CG7330	CG7330	**-0.70** **P=0.024**	**-0.66**		
CG13293	CG13293	CG13293	**-0.55** **P=0.033**	**-0.66**		
mindmelt	mbl	CG33197	**-0.60** **P=0.022**	**-0.66**	MBNL1	Splicing factor, muscle and nervous system development
CG15311	CG15311	CG15311	**-0.66** **P=0.001**	**-0.65**		Diphosphatase
CG32521	CG32521	CG32521	**-0.8** **P=0.008**	**-0.63**		
sequoia	seq	CG32904	**-0.81** **P=0.041**	**-0.58**		Dendrite morphogenesis
CG1463	CG1463	CG1463	**-0.62** **P=0.007**	**-0.53**		
SET domain binding factor	Sbf	CG6939	**-0.54** **P=0.046**	**-0.53**	SBF2	DAG signalling
expanded	ex	CG4114	**-0.58** **P=0.023**	**-0.52**	FRMD6	Regulation of cell proliferation and differentiation
fettucine	cic	CG5067	**-0.51** **P=0.011**	**-0.52**	CIC	Transcription factor activity

Gene Title	Gene Symbol	Ensembl	log2(ratio) *elav>rCUG* to *elav>+*	log2 (ratio) *elav>rAUUCU* to *elav>+*	Human Orthologue	Function
alan shepard	shep	CG32423	**-0.73** **P=0.002**	**-0.52**		
CG14614	CG14614	CG14614	**-0.63** **P=0.017**	**-0.51**	WDR68	
Overflow	Dl	CG3619	**-0.56** **P=0.001**	**-0.51**	DLL1	Notch signalling pathway, neural development
bruno	aret	CG31762	**-0.52** **P=0.014**	**-0.51**	CUGBP1	Regulation of alternative splicing and translation
CG9449	CG9449	CG9449	**0.60** **P=0.031**	**0.53**		phagocytosis
CG9686	CG9686	CG9686	**0.54** **P=0.016**	**0.54**		
CG8021	CG8021	CG8021	**0.57** **P=0.035**	**0.56**		
CG9186	CG9186	CG9186	**0.77** **P=0.031**	**0.57**		
CG10006	CG10006	CG10006	**0.95** **P=0.001**	**0.58**	SLC39A6	Metal ion transporter
Kua	Kua	CG10723	**0.91** **P=0.002**	**0.67**	KUA	
CG9449	CG9449	CG9449	**0.60** **P=0.031**	**0.76**		phagocytosis
Olfactory-specific E	Os-E	CG11422	**0.52** **P=0.046**	**0.78**		Sensory perception of chemical stimulus
CG3603	CG3603	CG3603	**0.59** **P=0.018**	**0.82**	HSD17B8	Oxidation reduction
CG17758	CG17758	CG17758	**0.75** **P=0.017**	**0.83**	OTOP1	
CG32581	CG32581	CG32581	**1.31** **P=0.003**	**0.86**	RNF185	
Hormone receptor-like in 38	Hr38	CG1864	**0.55** **P=0.028**	**0.91**	NR4A1/ NGFI-B	Ligand-dependent nuclear receptor activity, transcription activity
CG6465	CG6465	CG6465	**1.01** **P=0.014**	**0.98**	ACY1	Hydrolysis of acetylated amino acids

Gene Title	Gene Symbol	Ensembl	log2(ratio) *elav>rCUG* to *elav>+*	log2 (ratio) *elav>rAUUCU* to *elav>+*	Human Orthologue	Function
cuticle cluster 2	Ccp84Ag	CG2342	**0.54** **P=0.047**	**1.02**		Component of cuticle
CG12179	CG12179	CG12179	**0.74** **P=0.041**	**1.12**		
lethal (1) G0155	Ykt6	CG1515	**1.05** **P=0.012**	**1.15**	YKT6	Vesicle-mediated transport, SNAP receptor activity
Bicoid interacting protein 4	Pk17E	CG7001	**1.36** **P=0.003**	**1.42**		Protein kinase
Odorant-binding protein 8a	Obp8a	CG12665	**1.11** **P=0.001**	**1.68**		Sensory perception of chemical stimulus
anon-fast-evolving-1H4	Rala	CG2849	**2.03** **P=0.018**	**2.19**	RALA	GTPase, actin cytoskeleton reorganisation

References

1. Kremer, E. J., Pritchard, M., Lynch, M., Yu, S., Holman, K., Baker, E., Warren, S. T., Schlessinger, D., Sutherland, G. R., and Richards, R. I., Mapping of DNA instability at the fragile X to a trinucleotide repeat sequence p(CCG)n, *Science., 252*, 1711 (1991).
2. La Spada, A. R., Wilson, E. M., Lubahn, D. B., Harding, A. E., and Fischbeck, K. H., Androgen receptor gene mutations in X-linked spinal and bulbar muscular atrophy, *Nature, 352*, 77 (1991).
3. Ranum, L. P. W., and Cooper, T. A., RNA-mediated neuromuscular disorders, *Annual Review of Neuroscience, 29*, 259 (2006).
4. Cummings, C. J., and Zoghbi, H. Y., Fourteen and counting: unraveling trinucleotide repeat diseases, *Hum. Mol. Genet., 9*, 909 (2000).
5. Finsterer, J., Myotonic dystrophy type 2, *European Journal of Neurology, 9*, 441 (2002).
6. Marz, P., Probst, A., Lang, S., Schwager, M., Rose-John, S., Otten, U., and Ozbek, S., Ataxin-10, the Spinocerebellar Ataxia Type 10 Neurodegenerative Disorder Protein, Is Essential for Survival of Cerebellar Neurons, *J. Biol. Chem., 279*, 35542 (2004).
7. Richards, R. I., Dynamic mutations: a decade of unstable expanded repeats in human genetic disease, *Hum. Mol. Genet., 10*, 2187 (2001).
8. Tsuji, S., Molecular genetics of triplet repeats: unstable expansion of triplet repeats as a new mechanism for neurodegenerative diseases, *Intern Med, 36*, 3 (1997).
9. Sieradzan, K. A., and Mann, D. M. A., The selective vulnerability of nerve cells in Huntington's disease, *Neuropathology and Applied Neurobiology, 27*, 1 (2001).
10. Gutekunst, C., Levey, A., Heilman, C., Whaley, W., Yi, H., Nash, N., Rees, H., Madden, J., and Hersch, S., Identification and Localization of Huntingtin in Brain and Human Lymphoblastoid Cell Lines with Anti-Fusion Protein Antibodies, *PNAS, 92*, 8710 (1995).
11. Velier, J., Kim, M., Schwarz, C., Kim, T. W., Sapp, E., Chase, K., Aronin, N., and DiFiglia, M., Wild-Type and Mutant Huntingtins Function in Vesicle Trafficking in the Secretory and Endocytic Pathways, *Experimental Neurology, 152*, 34 (1998).

12. Trushina, E., Dyer, R. B., Badger, J. D., II, Ure, D., Eide, L., Tran, D. D., Vrieze, B. T., Legendre-Guillemin, V., McPherson, P. S., Mandavilli, B. S., Van Houten, B., Zeitlin, S., McNiven, M., Aebersold, R., Hayden, M., Parisi, J. E., Seeberg, E., Dragatsis, I., Doyle, K., Bender, A., Chacko, C., and McMurray, C. T., Mutant Huntingtin Impairs Axonal Trafficking in Mammalian Neurons In Vivo and In Vitro, *Mol. Cell. Biol., 24*, 8195 (2004).
13. Sinadinos, C., Burbidge-King, T., Soh, D., Thompson, L. M., Marsh, J. L., Wyttenbach, A., and Mudher, A. K., Live axonal transport disruption by mutant huntingtin fragments in *Drosophila* motor neuron axons, *Neurobiol Dis., 34*, 389 (2009).
14. Gunawardena, S., Her, L. S., Brusch, R. G., Laymon, R. A., Niesman, I. R., Gordesky-Gold, B., Sintasath, L., Bonini, N. M., and Goldstein, L. S., Disruption of axonal transport by loss of huntingtin or expression of pathogenic polyQ proteins in *Drosophila*, *Neuron., 40*, 25 (2003).
15. Landwehrmeyer, G. B., McNeil, S. M., Dure, L. S. t., Ge, P., Aizawa, H., Huang, Q., Ambrose, C. M., Duyao, M. P., Bird, E. D., Bonilla, E., and et al., Huntington's disease gene: regional and cellular expression in brain of normal and affected individuals, *Ann Neurol, 37*, 218 (1995).
16. Fusco, F. R., Chen, Q., Lamoreaux, W. J., Figueredo-Cardenas, G., Jiao, Y., Coffman, J. A., Surmeier, D. J., Honig, M. G., Carlock, L. R., and Reiner, A., Cellular Localization of Huntingtin in Striatal and Cortical Neurons in Rats: Lack of Correlation with Neuronal Vulnerability in Huntington's Disease, *J. Neurosci., 19*, 1189 (1999).
17. Hardy, J., and Orr, H., The genetics of neurodegenerative diseases, *Journal of Neurochemistry, 97*, 1690 (2006).
18. Chamberlain, N. L., Driver, E. D., and Miesfeld, R. L., The length and location of CAG trinucleotide repeats in the androgen receptor N-terminal domain affect transactivation function, *Nucleic Acids Res., 22*, 3181 (1994).
19. MacLean, H. E., Warne, G. L., and Zajac, J. D., Spinal and bulbar muscular atrophy: androgen receptor dysfunction caused by a trinucleotide repeat expansion, *Journal of the Neurological Sciences, 135*, 149 (1996).
20. Thomas, P. S., Jr, Fraley, G. S., Damien, V., Woodke, L. B., Zapata, F., Sopher, B. L., Plymate, S. R., and La Spada, A. R., Loss of endogenous androgen receptor protein accelerates motor neuron degeneration and

accentuates androgen insensitivity in a mouse model of X-linked spinal and bulbar muscular atrophy, *Hum. Mol. Genet., 15*, 2225 (2006).

21. Uyama, E., Kondo, I., Uchino, M., Fukushima, T., Murayama, N., Kuwano, A., Inokuchi, N., Ohtani, Y., and Ando, M., Dentatorubral-pallidoluysian atrophy (DRPLA): clinical, genetic, and neuroradiologic studies in a family, *Journal of the Neurological Sciences, 130*, 146 (1995).
22. Burke, J. R., Wingfield, M. S., Lewis, K. E., Roses, A. D., Lee, J. E., Hulette, C., Pericak-Vance, M. A., and Vance, J. M., The Haw River Syndrome: Dentatorubropallidoluysian atrophy (DRPLA) in an African-American family, *Nat Genet, 7*, 521 (1994).
23. Nagafuchi, S., Yanagisawa, H., Ohsaki, E., Shirayama, T., Tadokoro, K., Inoue, T., and Yamada, M., Structure and expression of the gene responsible for the triplet repeat disorder, dentatorubral and pallidoluysian atrophy (DRPLA), *Nat Genet, 8*, 177 (1994).
24. Margolis, R. L., Li, S.-H., Scott Young, W., Wagster, M. V., Colin Stine, O., Kidwai, A. S., Ashworth, R. G., and Ross, C. A., DRPLA gene (Atrophin-1) sequence and mRNA expression in human brain, *Molecular Brain Research, 36*, 219 (1996).
25. Wang, L., Rajan, H., Pitman, J. L., McKeown, M., and Tsai, C.-C., Histone deacetylase-associating Atrophin proteins are nuclear receptor corepressors, *Genes Dev., 20*, 525 (2006).
26. Schilling, G., Wood, J. D., Duan, K., Slunt, H. H., Gonzales, V., Yamada, M., Cooper, J. K., Margolis, R. L., Jenkins, N. A., Copeland, N. G., Takahashi, H., Tsuji, S., Price, D. L., Borchelt, D. R., and Ross, C. A., Nuclear accumulation of truncated atrophin-1 fragments in a transgenic mouse model of DRPLA, *Neuron., 24*, 275 (1999).
27. Yu, J., Ying, M., Zhuang, Y., Xu, T., Han, M., Wu, X., and Xu, R., C-terminal deletion of the atrophin-1 protein results in growth retardation but not neurodegeneration in mice, *Dev Dyn., 238*, 2471 (2009).
28. van de Warrenburg, B. P. C., Notermans, N. C., Schelhaas, H. J., van Alfen, N., Sinke, R. J., Knoers, N. V. A. M., Zwarts, M. J., and Kremer, B. P. H., Peripheral Nerve Involvement in Spinocerebellar Ataxias, *Arch Neurol, 61*, 257 (2004).

29. Carlson, K. M., Andresen, J. M., and Orr, H. T., Emerging pathogenic pathways in the spinocerebellar ataxias, *Curr Opin Genet Dev., 19*, 247 (2009).
30. Yue, S., Serra, H. G., Zoghbi, H. Y., and Orr, H. T., The spinocerebellar ataxia type 1 protein, ataxin-1, has RNA-binding activity that is inversely affected by the length of its polyglutamine tract, *Hum Mol Genet., 10*, 25 (2001).
31. Klement, I. A., Skinner, P. J., Kaytor, M. D., Yi, H., Hersch, S. M., Clark, H. B., Zoghbi, H. Y., and Orr, H. T., Ataxin-1 Nuclear Localization and Aggregation: Role in Polyglutamine-Induced Disease in SCA1 Transgenic Mice, *Cell, 95*, 41 (1998).
32. Sharma, D., Sharma, S., Pasha, S., and Brahmachari, S. K., Peptide models for inherited neurodegenerative disorders: conformation and aggregation properties of long polyglutamine peptides with and without interruptions, *FEBS Letters, 456*, 181 (1999).
33. Imbert, G., Saudou, F., Yvert, G., Devys, D., Trottier, Y., Garnier, J.-M., Weber, C., Mandel, J.-L., Cancel, G., Abbas, N., Durr, A., Didierjean, O., Stevanin, G., Agid, Y., and Brice, A., Cloning of the gene for spinocerebellar ataxia 2 reveals a locus with high sensitivity to expanded CAG/glutamine repeats, *Nat Genet, 14*, 285 (1996).
34. Satterfield, T. F., and Pallanck, L. J., Ataxin-2 and its *Drosophila* homolog, ATX2, physically assemble with polyribosomes, *Hum. Mol. Genet., 15*, 2523 (2006).
35. Ralser, M., Albrecht, M., Nonhoff, U., Lengauer, T., Lehrach, H., and Krobitsch, S., An Integrative Approach to Gain Insights into the Cellular Function of Human Ataxin-2, *Journal of Molecular Biology, 346*, 203 (2005).
36. Satterfield, T. F., Jackson, S. M., and Pallanck, L. J., A *Drosophila* homolog of the polyglutamine disease gene SCA2 is a dosage-sensitive regulator of actin filament formation, *Genetics., 162*, 1687 (2002).
37. Huynh, D. P., Figueroa, K., Hoang, N., and Pulst, S.-M., Nuclear localization or inclusion body formation of ataxin-2 are not necessary for SCA2 pathogenesis in mouse or human, *Nat Genet, 26*, 44 (2000).
38. Klockgether, T., Schols, L., Abele, M., Burk, K., Topka, H., Andres, F., Amoiridis, G., Ludtke, R., Riess, O., Laccone, F., and Dichgans, J., Age related axonal neuropathy in spinocerebellar ataxia type 3/Machado-Joseph disease (SCA3/MJD), *J Neurol Neurosurg Psychiatry, 66*, 222 (1999).

39. Wang, G.-h., Sawai, N., Kotliarova, S., Kanazawa, I., and Nukina, N., Ataxin-3, the MJD1 gene product, interacts with the two human homologs of yeast DNA repair protein RAD23, HHR23A and HHR23B, *Hum. Mol. Genet., 9*, 1795 (2000).
40. Doss-Pepe, E. W., Stenroos, E. S., Johnson, W. G., and Madura, K., Ataxin-3 Interactions with Rad23 and Valosin-Containing Protein and Its Associations with Ubiquitin Chains and the Proteasome Are Consistent with a Role in Ubiquitin-Mediated Proteolysis, *Mol. Cell. Biol., 23*, 6469 (2003).
41. Chen, X., Tang, T. S., Tu, H., Nelson, O., Pook, M., Hammer, R., Nukina, N., and Bezprozvanny, I., Deranged calcium signaling and neurodegeneration in spinocerebellar ataxia type 3, *J Neurosci., 28*, 12713 (2008).
42. Jia, N. L., Fei, E. K., Ying, Z., Wang, H. F., and Wang, G. H., PolyQ-expanded ataxin-3 interacts with full-length ataxin-3 in a polyQ length-dependent manner, *Neurosci Bull., 24*, 201 (2008).
43. Frontali, M., Spinocerebellar ataxia type 6: channelopathy or glutamine repeat disorder?, *Brain Research Bulletin, 56*, 227 (2001).
44. Gazulla, J., and Tintore, M. A., The P/Q-type voltage-dependent calcium channel as pharmacological target in spinocerebellar ataxia type 6: gabapentin and pregabalin may be of therapeutic benefit, *Med Hypotheses., 68*, 131 (2007).
45. Watase, K., Barrett, C. F., Miyazaki, T., Ishiguro, T., Ishikawa, K., Hu, Y., Unno, T., Sun, Y., Kasai, S., Watanabe, M., Gomez, C. M., Mizusawa, H., Tsien, R. W., and Zoghbi, H. Y., Spinocerebellar ataxia type 6 knockin mice develop a progressive neuronal dysfunction with age-dependent accumulation of mutant CaV2.1 channels, *Proc Natl Acad Sci U S A., 105*, 11987 (2008).
46. Michalik, A., Martin, J. J., and Van Broeckhoven, C., Spinocerebellar ataxia type 7 associated with pigmentary retinal dystrophy, *Eur J Hum Genet, 12*, 2 (2004).
47. Strom, A.-L., Forsgren, L., and Holmberg, M., A role for both wild-type and expanded ataxin-7 in transcriptional regulation, *Neurobiology of Disease, 20*, 646 (2005).
48. Palhan, V. B., Chen, S., Peng, G.-H., Tjernberg, A., Gamper, A. M., Fan, Y., Chait, B. T., La Spada, A. R., and Roeder, R. G., Polyglutamine-expanded ataxin-7 inhibits STAGA histone acetyltransferase activity to produce retinal degeneration, *PNAS, 102*, 8472 (2005).

49. Chou, A. H., Chen, C. Y., Chen, S. Y., Chen, W. J., Chen, Y. L., Weng, Y. S., and Wang, H. L., Polyglutamine-expanded ataxin-7 causes cerebellar dysfunction by inducing transcriptional dysregulation, *Neurochem Int., 56*, 329 (2010).
50. van Roon-Mom, W. M. C., Reid, S. J., Faull, R. L. M., and Snell, R. G., TATA-binding protein in neurodegenerative disease, *Neuroscience, 133*, 863 (2005).
51. Shah, A. G., Friedman, M. J., Huang, S., Roberts, M., Li, X. J., and Li, S., Transcriptional dysregulation of TrkA associates with neurodegeneration in spinocerebellar ataxia type 17, *Hum Mol Genet., 18*, 4141 (2009).
52. DiFiglia, M., Sapp, E., Chase, K. O., Davies, S. W., Bates, G. P., Vonsattel, J. P., and Aronin, N., Aggregation of huntingtin in neuronal intranuclear inclusions and dystrophic neurites in brain, *Science., 277*, 1990 (1997).
53. Wells, R. D., Ashizawa, T., *Genetic instabilities and neurological diseases*, Academic Press (2006).
54. Li, H., Li, S.-H., Johnston, H., Shelbourne, P. F., and Li, X.-J., Amino-terminal fragments of mutant huntingtin show selective accumulation in striatal neurons and synaptic toxicity*, 25*, 385 (2000).
55. Saudou, F., Finkbeiner, S., Devys, D., and Greenberg, M. E., Huntingtin Acts in the Nucleus to Induce Apoptosis but Death Does Not Correlate with the Formation of Intranuclear Inclusions, *Cell, 95*, 55 (1998).
56. Michalik, A., and Van Broeckhoven, C., Pathogenesis of polyglutamine disorders: aggregation revisited, *Hum. Mol. Genet., 12*, R173 (2003).
57. Lavoie, H., Debeane, F., Trinh, Q.-D., Turcotte, J.-F., Corbeil-Girard, L.-P., Dicaire, M.-J., Saint-Denis, A., Page, M., Rouleau, G. A., and Brais, B., Polymorphism, shared functions and convergent evolution of genes with sequences coding for polyalanine domains, *Hum. Mol. Genet., 12*, 2967 (2003).
58. Toulouse, A., Au-Yeung, F., Gaspar, C., Roussel, J., Dion, P., and Rouleau, G. A., Ribosomal frameshifting on MJD-1 transcripts with long CAG tracts, *Hum. Mol. Genet., 14*, 2649 (2005).
59. Davies, J. E., and Rubinsztein, D. C., Polyalanine and polyserine frameshift products in Huntington's disease, *J Med Genet., 43*, 893 (2006).
60. Berger, Z., Davies, J. E., Luo, S., Pasco, M. Y., Majoul, I., O'Kane, C. J., and Rubinsztein, D. C., Deleterious and protective properties of an aggregate-prone protein with a polyalanine expansion, *Hum. Mol. Genet., 15*, 453 (2006).

61. McLeod, C. J., O'Keefe, L. V., and Richards, R. I., The pathogenic agent in *Drosophila* models of 'polyglutamine' diseases, *Hum. Mol. Genet., 14*, 1041 (2005).
62. Gecz, J., Gedeon, A. K., Sutherland, G. R., and Mulley, J. C., Identification of the gene FMR2, associated with FRAXE mental retardation*, 13*, 105 (1996).
63. Bidichandani, S. I., Ashizawa, T., and Patel, P. I., The GAA triplet-repeat expansion in Friedreich ataxia interferes with transcription and may be associated with an unusual DNA structure, *Am J Hum Genet, 62*, 111 (1998).
64. Ranum, L. P., and Day, J. W., Myotonic dystrophy: RNA pathogenesis comes into focus, *Am J Hum Genet, 74*, 793 (2004).
65. Giorgio, A., Dotti, M. T., Battaglini, M., Marino, S., Mortilla, M., Stromillo, M. L., Bramanti, P., Orrico, A., Federico, A., and De Stefano, N., Cortical damage in brains of patients with adult-form of myotonic dystrophy type 1 and no or minimal MRI abnormalities, *J Neurol., 253*, 1471 (2006).
66. Sistiaga, A., Urreta, I., Jodar, M., Cobo, A. M., Emparanza, J., Otaegui, D., Poza, J. J., Merino, J. J., Imaz, H., Marti-Masso, J. F., and Lopez de Munain, A., Cognitive/personality pattern and triplet expansion size in adult myotonic dystrophy type 1 (DM1): CTG repeats, cognition and personality in DM1, *Psychol Med., 40*, 487 (2010).
67. Seznec, H., Agbulut, O., Sergeant, N., Savouret, C., Ghestem, A., Tabti, N., Willer, J.-C., Ourth, L., Duros, C., Brisson, E., Fouquet, C., Butler-Browne, G., Delacourte, A., Junien, C., and Gourdon, G., Mice transgenic for the human myotonic dystrophy region with expanded CTG repeats display muscular and brain abnormalities, *Hum. Mol. Genet., 10*, 2717 (2001).
68. Kirkwood E. Personius, J. N., Sita Reddy,, Myotonia and muscle contractile properties in mice with SIX5 deficiency, *Muscle & Nerve, 31*, 503 (2005).
69. Klesert, T. R., Cho, D. H., Clark, J. I., Maylie, J., Adelman, J., Snider, L., Yuen, E. C., Soriano, P., and Tapscott, S. J., Mice deficient in Six5 develop cataracts: implications for myotonic dystrophy, *Nat Genet, 25*, 105 (2000).
70. Mankodi, A., Logigian, E., Callahan, L., McClain, C., White, R., Henderson, D., Krym, M., and Thornton, C. A., Myotonic Dystrophy in Transgenic Mice Expressing an Expanded CUG Repeat, *Science, 289*, 1769 (2000).
71. Ho, T. H., Savkur, R. S., Poulos, M. G., Mancini, M. A., Swanson, M. S., and Cooper, T. A., Colocalization of muscleblind with RNA foci is separable from

mis-regulation of alternative splicing in myotonic dystrophy, *J Cell Sci, 118*, 2923 (2005).

72. Jiang, H., Mankodi, A., Swanson, M. S., Moxley, R. T., and Thornton, C. A., Myotonic dystrophy type 1 is associated with nuclear foci of mutant RNA, sequestration of muscleblind proteins and deregulated alternative splicing in neurons, *Hum. Mol. Genet., 13*, 3079 (2004).
73. Kimura, T., Nakamori, M., Lueck, J. D., Pouliquin, P., Aoike, F., Fujimura, H., Dirksen, R. T., Takahashi, M. P., Dulhunty, A. F., and Sakoda, S., Altered mRNA splicing of the skeletal muscle ryanodine receptor and sarcoplasmic/endoplasmic reticulum Ca2+-ATPase in myotonic dystrophy type 1, *Hum. Mol. Genet., 14*, 2189 (2005).
74. Mankodi, A., Takahashi, M. P., Jiang, H., Beck, C. L., Bowers, W. J., Moxley, R. T., Cannon, S. C., and Thornton, C. A., Expanded CUG Repeats Trigger Aberrant Splicing of ClC-1 Chloride Channel Pre-mRNA and Hyperexcitability of Skeletal Muscle in Myotonic Dystrophy, *Molecular Cell, 10*, 35 (2002).
75. Savkur, R. S., Philips, A. V., Cooper, T. A., Dalton, J. C., Moseley, M. L., Ranum, L. P., and Day, J. W., Insulin receptor splicing alteration in myotonic dystrophy type 2, *Am J Hum Genet, 74*, 1309 (2004).
76. Ami Mankodi, P. T.-U., Matt Krym, Don Henderson, Maurice Swanson, Charles A. Thornton,, Ribonuclear inclusions in skeletal muscle in myotonic dystrophy types 1 and 2, *Annals of Neurology, 54*, 760 (2003).
77. Gerbasi, V. R., and Link, A. J., The myotonic dystrophy type 2 protein ZNF9 is part of an ITAF complex that promotes cap-independent translation, *Mol Cell Proteomics., 6*, 1049 (2007).
78. Huichalaf, C., Schoser, B., Schneider-Gold, C., Jin, B., Sarkar, P., and Timchenko, L., Reduction of the rate of protein translation in patients with myotonic dystrophy 2, *J Neurosci., 29*, 9042 (2009).
79. Chen, W., Wang, Y., Abe, Y., Cheney, L., Udd, B., and Li, Y. P., Haploinsuffciency for Znf9 in Znf9+/- mice is associated with multiorgan abnormalities resembling myotonic dystrophy, *J Mol Biol., 368*, 8 (2007).
80. Sammons, M. A., Antons, A. K., Bendjennat, M., Udd, B., Krahe, R., and Link, A. J., ZNF9 activation of IRES-mediated translation of the human ODC mRNA is decreased in myotonic dystrophy type 2, *PLoS One., 5*, e9301. (2010).
81. Mankodi, A., Urbinati, C. R., Yuan, Q.-P., Moxley, R. T., Sansone, V., Krym, M., Henderson, D., Schalling, M., Swanson, M. S., and Thornton, C. A.,

Muscleblind localizes to nuclear foci of aberrant RNA in myotonic dystrophy types 1 and 2, *Hum. Mol. Genet., 10*, 2165 (2001).

82. Salvatori, S., Furlan, S., Fanin, M., Picard, A., Pastorello, E., Romeo, V., Trevisan, C. P., and Angelini, C., Comparative transcriptional and biochemical studies in muscle of myotonic dystrophies (DM1 and DM2), *Neurol Sci., 30*, 185 (2009).
83. Botta, A., Vallo, L., Rinaldi, F., Bonifazi, E., Amati, F., Biancolella, M., Gambardella, S., Mancinelli, E., Angelini, C., Meola, G., and Novelli, G., Gene expression analysis in myotonic dystrophy: indications for a common molecular pathogenic pathway in DM1 and DM2, *Gene Expr., 13*, 339 (2007).
84. Ho, T. H., Bundman, D., Armstrong, D. L., and Cooper, T. A., Transgenic mice expressing CUG-BP1 reproduce splicing mis-regulation observed in myotonic dystrophy, *Hum. Mol. Genet., 14*, 1539 (2005).
85. Ladd, A. N., Taffet, G., Hartley, C., Kearney, D. L., and Cooper, T. A., Cardiac tissue-specific repression of CELF activity disrupts alternative splicing and causes cardiomyopathy, *Mol Cell Biol., 25*, 6267 (2005).
86. de Haro, M., Al-Ramahi, I., De Gouyon, B., Ukani, L., Rosa, A., Faustino, N. A., Ashizawa, T., Cooper, T. A., and Botas, J., MBNL1 and CUGBP1 modify expanded CUG-induced toxicity in a *Drosophila* model of myotonic dystrophy type 1, *Hum Mol Genet., 15*, 2138 (2006).
87. Hagerman, R. J., Lessons from fragile X regarding neurobiology, autism, and neurodegeneration, *J Dev Behav Pediatr, 27*, 63 (2006).
88. Hagerman, R. J., Ono, M. Y., and Hagerman, P. J., Recent advances in fragile X: a model for autism and neurodegeneration, *Curr Opin Psychiatry, 18*, 490 (2005).
89. Iwahashi, C. K., Yasui, D. H., An, H.-J., Greco, C. M., Tassone, F., Nannen, K., Babineau, B., Lebrilla, C. B., Hagerman, R. J., and Hagerman, P. J., Protein composition of the intranuclear inclusions of FXTAS, *Brain, 129*, 256 (2006).
90. Hashem, V., Galloway, J. N., Mori, M., Willemsen, R., Oostra, B. A., Paylor, R., and Nelson, D. L., Ectopic expression of CGG containing mRNA is neurotoxic in mammals, *Hum Mol Genet., 18*, 2443 (2009).
91. Day, J. W., Schut, L. J., Moseley, M. L., Durand, A. C., and Ranum, L. P. W., Spinocerebellar ataxia type 8: Clinical features in a large family, *Neurology, 55*, 649 (2000).

92. Gupta, A., and Jankovic, J., Spinocerebellar ataxia 8: variable phenotype and unique pathogenesis, *Parkinsonism Relat Disord., 15*, 621 (2009).
93. Nemes, J. P., Benzow, K. A., Moseley, M. L., Ranum, L. P., and Koob, M. D., The SCA8 transcript is an antisense RNA to a brain-specific transcript encoding a novel actin-binding protein (KLHL1), *Hum Mol Genet., 9*, 1543 (2000).
94. Aromolaran, K. A., Benzow, K. A., Cribbs, L. L., Koob, M. D., and Piedras-Renteria, E. S., T-type current modulation by the actin-binding protein Kelch-like 1, *Am J Physiol Cell Physiol., 298*, C1353 (2010).
95. Chen, W. L., Lin, J. W., Huang, H. J., Wang, S. M., Su, M. T., Lee-Chen, G. J., Chen, C. M., and Hsieh-Li, H. M., SCA8 mRNA expression suggests an antisense regulation of KLHL1 and correlates to SCA8 pathology, *Brain Res., 1233*, 176 (2008).
96. He, Y., Zu, T., Benzow, K. A., Orr, H. T., Clark, H. B., and Koob, M. D., Targeted deletion of a single Sca8 ataxia locus allele in mice causes abnormal gait, progressive loss of motor coordination, and Purkinje cell dendritic deficits, *J Neurosci., 26*, 9975 (2006).
97. Moseley, M. L., Zu, T., Ikeda, Y., Gao, W., Mosemiller, A. K., Daughters, R. S., Chen, G., Weatherspoon, M. R., Clark, H. B., Ebner, T. J., Day, J. W., and Ranum, L. P., Bidirectional expression of CUG and CAG expansion transcripts and intranuclear polyglutamine inclusions in spinocerebellar ataxia type 8, *Nat Genet., 38*, 758 (2006).
98. Mutsuddi, M., Marshall, C. M., Benzow, K. A., Koob, M. D., and Rebay, I., The spinocerebellar ataxia 8 noncoding RNA causes neurodegeneration and associates with staufen in *Drosophila*, *Curr Biol., 14*, 302 (2004).
99. Lin, X., and Ashizawa, T., Recent progress in spinocerebellar ataxia type10 (SCA10), *The Cerebellum, 4*, 37 (2005).
100. Handa, V., Yeh, H. J., McPhie, P., and Usdin, K., The AUUCU repeats responsible for spinocerebellar ataxia type 10 form unusual RNA hairpins, *J Biol Chem, 280*, 29340 (2005).
101. Keren, B., Jacquette, A., Depienne, C., Leite, P., Durr, A., Carpentier, W., Benyahia, B., Ponsot, G., Soubrier, F., Brice, A., and Heron, D., Evidence against haploinsuffiency of human ataxin 10 as a cause of spinocerebellar ataxia type 10, *Neurogenetics., 11*, 273 (2010).

102. Petersen, P., Chou, D. M., You, Z., Hunter, T., Walter, J. C., and Walter, G., Protein phosphatase 2A antagonizes ATM and ATR in a Cdk2- and Cdc7-independent DNA damage checkpoint, *Mol Cell Biol., 26*, 1997 (2006).
103. Dagda, R. K., Merrill, R. A., Cribbs, J. T., Chen, Y., Hell, J. W., Usachev, Y. M., and Strack, S., The spinocerebellar ataxia 12 gene product and protein phosphatase 2A regulatory subunit Bbeta2 antagonizes neuronal survival by promoting mitochondrial fission, *J Biol Chem., 283*, 36241 (2008).
104. Holmes, S. E., Hearn, E. O. r., Ross, C. A., and Margolis, R. L., SCA12: an unusual mutation leads to an unusual spinocerebellar ataxia, *Brain Research Bulletin, 56*, 397 (2001).
105. Holmes, S. E., O'Hearn, E., Rosenblatt, A., Callahan, C., Hwang, H. S., Ingersoll-Ashworth, R. G., Fleisher, A., Stevanin, G., Brice, A., Potter, N. T., Ross, C. A., and Margolis, R. L., A repeat expansion in the gene encoding junctophilin-3 is associated with Huntington disease-like 2*, 29*, 377 (2001).
106. Margolis, R. L., O'Hearn, E., Rosenblatt, A., Willour, V., Holmes, S. E., Franz, M. L., Callahan, C., Hwang, H. S., Troncoso, J. C., and Ross, C. A., A disorder similar to Huntington's disease is associated with a novel CAG repeat expansion, *Ann Neurol, 50*, 373 (2001).
107. Rudnicki, D. D., Holmes, S. E., Lin, M. W., Thornton, C. A., Ross, C. A., and Margolis, R. L., Huntington's disease--like 2 is associated with CUG repeat-containing RNA foci, *Ann Neurol., 61*, 272 (2007).
108. Stevanin, G., Fujigasaki, H., Lebre, A.-S., Camuzat, A., Jeannequin, C., Dode, C., Takahashi, J., San, C., Bellance, R., Brice, A., and Durr, A., Huntington's disease-like phenotype due to trinucleotide repeat expansions in the TBP and JPH3 genes, *Brain, 126*, 1599 (2003).
109. Marsh, J. L., Walker, H., Theisen, H., Zhu, Y.-Z., Fielder, T., Purcell, J., and Thompson, L. M., Expanded polyglutamine peptides alone are intrinsically cytotoxic and cause neurodegeneration in *Drosophila*, *Hum. Mol. Genet., 9*, 13 (2000).
110. Raspe, M., Gillis, J., Krol, H., Krom, S., Bosch, K., van Veen, H., and Reits, E., Mimicking proteasomal release of polyglutamine peptides initiates aggregation and toxicity, *J Cell Sci., 122*, 3262 (2009).
111. Nakayama, H., Hamada, M., Fujikake, N., Nagai, Y., Zhao, J., Hatano, O., Shimoke, K., Isosaki, M., Yoshizumi, M., and Ikeuchi, T., ER stress is the initial

response to polyglutamine toxicity in PC12 cells, *Biochem Biophys Res Commun., 377*, 550 (2008).

112. Wellington, C. L., Ellerby, L. M., Hackam, A. S., Margolis, R. L., Trifiro, M. A., Singaraja, R., McCutcheon, K., Salvesen, G. S., Propp, S. S., Bromm, M., Rowland, K. J., Zhang, T., Rasper, D., Roy, S., Thornberry, N., Pinsky, L., Kakizuka, A., Ross, C. A., Nicholson, D. W., Bredesen, D. E., and Hayden, M. R., Caspase cleavage of gene products associated with triplet expansion disorders generates truncated fragments containing the polyglutamine tract, *J Biol Chem., 273*, 9158 (1998).
113. Sobczak, K., de Mezer, M., Michlewski, G., Krol, J., and Krzyzosiak, W. J., RNA structure of trinucleotide repeats associated with human neurological diseases, *Nucleic Acids Res., 31*, 5469 (2003).
114. Amrane, S., and Mergny, J. L., Length and pH-dependent energetics of (CCG)n and (CGG)n trinucleotide repeats, *Biochimie., 88*, 1125 (2006).
115. Zumwalt, M., Ludwig, A., Hagerman, P. J., and Dieckmann, T., Secondary structure and dynamics of the r(CGG) repeat in the mRNA of the fragile X mental retardation 1 (FMR1) gene, *RNA Biol., 4*, 93 (2007).
116. Yuan, Y., Compton, S. A., Sobczak, K., Stenberg, M. G., Thornton, C. A., Griffith, J. D., and Swanson, M. S., Muscleblind-like 1 interacts with RNA hairpins in splicing target and pathogenic RNAs, *Nucleic Acids Res., 35*, 5474 (2007).
117. Li, L. B., Yu, Z., Teng, X., and Bonini, N. M., RNA toxicity is a component of ataxin-3 degeneration in *Drosophila*, *Nature., 453*, 1107 (2008).
118. Doumanis, J., Wada, K., Kino, Y., Moore, A. W., and Nukina, N., RNAi screening in *Drosophila* cells identifies new modifiers of mutant huntingtin aggregation, *PLoS One., 4*, e7275. (2009).
119. Lee, W. C., Yoshihara, M., and Littleton, J. T., Cytoplasmic aggregates trap polyglutamine-containing proteins and block axonal transport in a *Drosophila* model of Huntington's disease, *Proc Natl Acad Sci U S A., 101*, 3224 (2004).
120. Ravikumar, B., Vacher, C., Berger, Z., Davies, J. E., Luo, S., Oroz, L. G., Scaravilli, F., Easton, D. F., Duden, R., O'Kane, C. J., and Rubinsztein, D. C., Inhibition of mTOR induces autophagy and reduces toxicity of polyglutamine expansions in fly and mouse models of Huntington disease, *Nat Genet., 36*, 585 (2004).

121. Kaltenbach, L. S., Romero, E., Becklin, R. R., Chettier, R., Bell, R., Phansalkar, A., Strand, A., Torcassi, C., Savage, J., Hurlburt, A., Cha, G. H., Ukani, L., Chepanoske, C. L., Zhen, Y., Sahasrabudhe, S., Olson, J., Kurschner, C., Ellerby, L. M., Peltier, J. M., Botas, J., and Hughes, R. E., Huntingtin interacting proteins are genetic modifiers of neurodegeneration, *PLoS Genet., 3*, e82. (2007).
122. Suzuki, E., Zhao, Y., Ito, S., Sawatsubashi, S., Murata, T., Furutani, T., Shirode, Y., Yamagata, K., Tanabe, M., Kimura, S., Ueda, T., Fujiyama, S., Lim, J., Matsukawa, H., Kouzmenko, A. P., Aigaki, T., Tabata, T., Takeyama, K., and Kato, S., Aberrant E2F activation by polyglutamine expansion of androgen receptor in SBMA neurotoxicity, *Proc Natl Acad Sci U S A., 106*, 3818 (2009).
123. Takeyama, K., Ito, S., Yamamoto, A., Tanimoto, H., Furutani, T., Kanuka, H., Miura, M., Tabata, T., and Kato, S., Androgen-dependent neurodegeneration by polyglutamine-expanded human androgen receptor in *Drosophila*, *Neuron., 35*, 855 (2002).
124. Nisoli, I., Chauvin, J. P., Napoletano, F., Calamita, P., Zanin, V., Fanto, M., and Charroux, B., Neurodegeneration by polyglutamine Atrophin is not rescued by induction of autophagy, *Cell Death Differ, 26* (2010).
125. Fernandez-Funez, P., Nino-Rosales, M. L., de Gouyon, B., She, W. C., Luchak, J. M., Martinez, P., Turiegano, E., Benito, J., Capovilla, M., Skinner, P. J., McCall, A., Canal, I., Orr, H. T., Zoghbi, H. Y., and Botas, J., Identification of genes that modify ataxin-1-induced neurodegeneration, *Nature., 408*, 101 (2000).
126. Chen, H. K., Fernandez-Funez, P., Acevedo, S. F., Lam, Y. C., Kaytor, M. D., Fernandez, M. H., Aitken, A., Skoulakis, E. M., Orr, H. T., Botas, J., and Zoghbi, H. Y., Interaction of Akt-phosphorylated ataxin-1 with 14-3-3 mediates neurodegeneration in spinocerebellar ataxia type 1, *Cell., 113*, 457 (2003).
127. Lam, Y. C., Bowman, A. B., Jafar-Nejad, P., Lim, J., Richman, R., Fryer, J. D., Hyun, E. D., Duvick, L. A., Orr, H. T., Botas, J., and Zoghbi, H. Y., ATAXIN-1 interacts with the repressor Capicua in its native complex to cause SCA1 neuropathology, *Cell., 127*, 1335 (2006).
128. Bonini, N. M., A genetic model for human polyglutamine-repeat disease in *Drosophila* melanogaster, *Philos Trans R Soc Lond B Biol Sci., 354*, 1057 (1999).

129. Latouche, M., Lasbleiz, C., Martin, E., Monnier, V., Debeir, T., Mouatt-Prigent, A., Muriel, M. P., Morel, L., Ruberg, M., Brice, A., Stevanin, G., and Tricoire, H., A conditional pan-neuronal *Drosophila* model of spinocerebellar ataxia 7 with a reversible adult phenotype suitable for identifying modifier genes, *J Neurosci., 27*, 2483 (2007).
130. Jin, P., Duan, R., Qurashi, A., Qin, Y., Tian, D., Rosser, T. C., Liu, H., Feng, Y., and Warren, S. T., Pur alpha binds to rCGG repeats and modulates repeat-mediated neurodegeneration in a *Drosophila* model of fragile X tremor/ataxia syndrome, *Neuron., 55*, 556 (2007).
131. Sofola, O. A., Jin, P., Qin, Y., Duan, R., Liu, H., de Haro, M., Nelson, D. L., and Botas, J., RNA-binding proteins hnRNP A2/B1 and CUGBP1 suppress fragile X CGG premutation repeat-induced neurodegeneration in a *Drosophila* model of FXTAS, *Neuron., 55*, 565 (2007).
132. Sofola, O. A., Jin, P., Botas, J., and Nelson, D. L., Argonaute-2-dependent rescue of a *Drosophila* model of FXTAS by FRAXE premutation repeat, *Hum Mol Genet., 16*, 2326 (2007).
133. McLeod, C., Investigation of the pathogenic agent in a *Drosophila* model of polyglutamine disease, in *Discipline of Genetics*, Vol. PhD, University of Adelaide, Adelaide (2006).
134. Dickson, B., and Hafen, E., Genetic dissection of eye development in *Drosophila*, in *The Development of Drosophila melanogaster.*, Vol. 2, Bate, M., and Arias, A., Eds., Cold Spring Harbour Laboratory Press, pp. 1327 (1993).
135. Brand, A. H., and Perrimon, N., Targeted gene expression as a means of altering cell fates and generating dominant phenotypes, *Development., 118*, 401 (1993).
136. Dorsman, J. C., Pepers, B., Langenberg, D., Kerkdijk, H., Ijszenga, M., den Dunnen, J. T., Roos, R. A., and van Ommen, G. J., Strong aggregation and increased toxicity of polyleucine over polyglutamine stretches in mammalian cells, *Hum Mol Genet., 11*, 1487 (2002).
137. Rezaval, C., Werbajh, S., and Ceriani, M. F., Neuronal death in *Drosophila* triggered by GAL4 accumulation, *Eur J Neurosci., 25*, 683 (2007).
138. Groth, A. C., Fish, M., Nusse, R., and Calos, M. P., Construction of transgenic *Drosophila* by using the site-specific integrase from phage phiC31, *Genetics., 166*, 1775 (2004).

139. Ma, E., Gu, X. Q., Wu, X., Xu, T., and Haddad, G. G., Mutation in pre-mRNA adenosine deaminase markedly attenuates neuronal tolerance to O2 deprivation in *Drosophila* melanogaster, *J Clin Invest., 107*, 685 (2001).
140. Palladino, M. J., Keegan, L. P., O'Connell, M. A., and Reenan, R. A., A-to-I Pre-mRNA Editing in *Drosophila* Is Primarily Involved in Adult Nervous System Function and Integrity, *Cell, 102*, 437 (2000).
141. Sambrook, J., Fritsch, E.F. and Maniatis, T., *Molecular cloning: A laboratory manual*, Cold Spring Harbour Laboratory Press (1989).
142. Laccone, F., Maiwald, R., and Bingemann, S., A fast polymerase chain reaction-mediated strategy for introducing repeat expansions into CAG-repeat containing genes, *Hum Mutat., 13*, 497 (1999).
143. Marsh, J. L., Walker, H., Theisen, H., Zhu, Y. Z., Fielder, T., Purcell, J., and Thompson, L. M., Expanded polyglutamine peptides alone are intrinsically cytotoxic and cause neurodegeneration in *Drosophila*, *Hum Mol Genet., 9*, 13 (2000).
144. Peel, A. L., Rao, R. V., Cottrell, B. A., Hayden, M. R., Ellerby, L. M., and Bredesen, D. E., Double-stranded RNA-dependent protein kinase, PKR, binds preferentially to Huntington's disease (HD) transcripts and is activated in HD tissue, *Hum. Mol. Genet., 10*, 1531 (2001).
145. Tian, B., White, R. J., Xia, T., Welle, S., Turner, D. H., Mathews, M. B., and Thornton, C. A., Expanded CUG repeat RNAs form hairpins that activate the double-stranded RNA-dependent protein kinase PKR, *RNA, 6*, 79 (2000).
146. Balachandran, S., Roberts, P. C., Brown, L. E., Truong, H., Pattnaik, A. K., Archer, D. R., and Barber, G. N., Essential Role for the dsRNA-Dependent Protein Kinase PKR in Innate Immunity to Viral Infection, *Immunity, 13*, 129 (2000).
147. Mankodi, A., Lin, X., Blaxall, B. C., Swanson, M. S., and Thornton, C. A., Nuclear RNA Foci in the Heart in Myotonic Dystrophy, *Circ Res, 97*, 1152 (2005).
148. Peter A. Beal, Duplex RNA-Binding Enzymes: Headliners from Neurobiology, Virology, and Development, *ChemBioChem, 6*, 257 (2005).
149. Krol, J., Fiszer, A., Mykowska, A., Sobczak, K., de Mezer, M., and Krzyzosiak, W. J., Ribonuclease dicer cleaves triplet repeat hairpins into shorter repeats that silence specific targets, *Mol Cell., 25*, 575 (2007).

150. Handa, V., Saha, T., and Usdin, K., The fragile X syndrome repeats form RNA hairpins that do not activate the interferon-inducible protein kinase, PKR, but are cut by Dicer, *Nucleic Acids Res., 31*, 6243 (2003).
151. Bratt, E., and Ohman, M., Coordination of editing and splicing of glutamate receptor pre-mRNA, *RNA, 9*, 309 (2003).
152. Ohlson, J., Pedersen, J. S., Haussler, D., and Ohman, M., Editing modifies the GABA(A) receptor subunit alpha3, *RNA, 13*, 698 (2007).
153. Scadden, A. D., The RISC subunit Tudor-SN binds to hyper-edited double-stranded RNA and promotes its cleavage, *Nat Struct Mol Biol., 12*, 489 (2005).
154. Keegan, L., Leroy, A., Sproul, D., and O'Connell, M., Adenosine deaminases acting on RNA (ADARs): RNA-editing enzymes, *Genome Biology, 5*, 209 (2004).
155. Kallman, A. M., Sahlin, M., and Ohman, M., ADAR2 A->I editing: site selectivity and editing efficiency are separate events, *Nucl. Acids Res., 31*, 4874 (2003).
156. DeCerbo, J., and Carmichael, G. G., Retention and repression: fates of hyperedited RNAs in the nucleus, *Current Opinion in Cell Biology, 17*, 302 (2005).
157. Emili, A., Shales, M., McCracken, S., Xie, W., Tucker, P. W., Kobayashi, R., Blencowe, B. J., and Ingles, C. J., Splicing and transcription-associated proteins PSF and p54nrb/nonO bind to the RNA polymerase II CTD, *RNA, 8*, 1102 (2002).
158. Wang, Q., Zhang, Z., Blackwell, K., and Carmichael, G. G., Vigilins Bind to Promiscuously A-to-I-Edited RNAs and Are Involved in the Formation of Heterochromatin, *Current Biology, 15*, 384 (2005).
159. Scadden, A. D. J., and O'Connell, M. A., Cleavage of dsRNAs hyper-edited by ADARs occurs at preferred editing sites, *Nucl. Acids Res., 33*, 5954 (2005).
160. O'Connell, M. A., and Keegan, L. P., Drosha versus ADAR: wrangling over pri-miRNA, *Nat Struct Mol Biol, 13*, 3 (2006).
161. Li, C. L., Yang, W. Z., Chen, Y. P., and Yuan, H. S., Structural and functional insights into human Tudor-SN, a key component linking RNA interference and editing, *Nucleic Acids Res., 36*, 3579 (2008).
162. Yang, W., Chendrimada, T. P., Wang, Q., Higuchi, M., Seeburg, P. H., Shiekhattar, R., and Nishikura, K., Modulation of microRNA processing and

expression through RNA editing by ADAR deaminases, *Nat Struct Mol Biol, 13*, 13 (2006).

163. Hang, P. N., Tohda, M., and Matsumoto, K., Developmental changes in expression and self-editing of adenosine deaminase type 2 pre-mRNA and mRNA in rat brain and cultured cortical neurons, *Neurosci Res., 61*, 398 (2008).
164. Maas, S., Patt, S., Schrey, M., and Rich, A., Underediting of glutamate receptor GluR-B mRNA in malignant gliomas, *PNAS, 98*, 14687 (2001).
165. Higuchi, M., Maas, S., Single, F. N., Hartner, J., Rozov, A., Burnashev, N., Feldmeyer, D., Sprengel, R., and Seeburg, P. H., Point mutation in an AMPA receptor gene rescues lethality in mice deficient in the RNA-editing enzyme ADAR2*, 406*, 78 (2000).
166. Kwak, S., and Kawahara, Y., Deficient RNA editing of GluR2 and neuronal death in amyotropic lateral sclerosis, *Journal of Molecular Medicine, 83*, 110 (2005).
167. Kwok, K. H. H., Wong, R. N. S., and Yung, K. K. L., Depletion of glutamate GluR2 receptor-containing neurons in the rat neostriatum by specific immunotoxin, *Neuroscience, 96*, 537 (2000).
168. Akbarian, S., Smith, M. A., and Jones, E. G., Editing for an AMPA receptor subunit RNA in prefrontal cortex and striatum in Alzheimer's disease, Huntington's disease and schizophrenia, *Brain Res., 699*, 297 (1995).
169. Levanon, E. Y., Hallegger, M., Kinar, Y., Shemesh, R., Djinovic-Carugo, K., Rechavi, G., Jantsch, M. F., and Eisenberg, E., Evolutionarily conserved human targets of adenosine to inosine RNA editing, *Nucl. Acids Res., 33*, 1162 (2005).
170. Schenck, A., Bardoni, B., Langmann, C., Harden, N., Mandel, J.-L., and Giangrande, A., CYFIP/Sra-1 Controls Neuronal Connectivity in *Drosophila* and Links the Rac1 GTPase Pathway to the Fragile X Protein, *Neuron, 38*, 887 (2003).
171. Palladino, M. J., Keegan, L. P., O'Connell, M. A., and Reenan, R. A., dADAR, a *Drosophila* double-stranded RNA-specific adenosine deaminase is highly developmentally regulated and is itself a target for RNA editing, *RNA, 6*, 1004 (2000).

172. Ma, E., Gu, X.-Q., Wu, X., Xu, T., and Haddad, G. G., Mutation in pre-mRNA adenosine deaminase markedly attenuates neuronal tolerance to O2 deprivation in *Drosophila* melanogaster, *J. Clin. Invest., 107*, 685 (2001).
173. Xia, S., Yang, J., Su, Y., Qian, J., Ma, E., and Haddad, G. G., Identification of new targets of *Drosophila* pre-mRNA adenosine deaminase, *Physiol. Genomics, 20*, 195 (2005).
174. Chen, L., Rio, D. C., Haddad, G. G., and Ma, E., Regulatory role of dADAR in ROS metabolism in *Drosophila* CNS, *Molecular Brain Research, 131*, 93 (2004).
175. Sansam, C. L., Wells, K. S., and Emeson, R. B., Modulation of RNA editing by functional nucleolar sequestration of ADAR2, *PNAS, 100*, 14018 (2003).
176. Handa, V., Saha, T., and Usdin, K., The fragile X syndrome repeats form RNA hairpins that do not activate the interferon-inducible protein kinase, PKR, but are cut by Dicer, *Nucl. Acids Res., 31*, 6243 (2003).
177. Morse, D. P., Aruscavage, P. J., and Bass, B. L., RNA hairpins in noncoding regions of human brain and Caenorhabditis elegans mRNA are edited by adenosine deaminases that act on RNA, *PNAS, 99*, 7906 (2002).
178. Lee, L. C., Chen, C. M., Chen, F. L., Lin, P. Y., Hsiao, Y. C., Wang, P. R., Su, M. T., Hsieh-Li, H. M., Hwang, J. C., Wu, C. H., Lee, G. C., Singh, S., Lin, Y., Hsieh, S. Y., Lee-Chen, G. J., and Lin, J. Y., Altered expression of HSPA5, HSPA8 and PARK7 in spinocerebellar ataxia type 17 identified by 2-dimensional fluorescence difference in gel electrophoresis, *Clin Chim Acta., 400*, 56 (2009).
179. Lam, W., Chan, W. M., Lo, T. W., Wong, A. K., Wu, C. C., and Chan, H. Y., Human receptor for activated protein kinase C1 associates with polyglutamine aggregates and modulates polyglutamine toxicity, *Biochem Biophys Res Commun., 377*, 714 (2008).
180. Sorolla, M. A., Reverter-Branchat, G., Tamarit, J., Ferrer, I., Ros, J., and Cabiscol, E., Proteomic and oxidative stress analysis in human brain samples of Huntington disease, *Free Radic Biol Med., 45*, 667 (2008).
181. Walsh, B., Pearl, A., Suchy, S., Tartaglio, J., Visco, K., and Phelan, S. A., Overexpression of Prdx6 and resistance to peroxide-induced death in Hepa1-6 cells: Prdx suppression increases apoptosis, *Redox Rep., 14*, 275 (2009).
182. Fatma, N., Kubo, E., Sen, M., Agarwal, N., Thoreson, W. B., Camras, C. B., and Singh, D. P., Peroxiredoxin 6 delivery attenuates TNF-alpha-and

glutamate-induced retinal ganglion cell death by limiting ROS levels and maintaining Ca2+ homeostasis, *Brain Res., 1233*, 63 (2008).

183. Power, J. H., Asad, S., Chataway, T. K., Chegini, F., Manavis, J., Temlett, J. A., Jensen, P. H., Blumbergs, P. C., and Gai, W. P., Peroxiredoxin 6 in human brain: molecular forms, cellular distribution and association with Alzheimer's disease pathology, *Acta Neuropathol., 115*, 611 (2008).
184. Manevich, Y., Shuvaeva, T., Dodia, C., Kazi, A., Feinstein, S. I., and Fisher, A. B., Binding of peroxiredoxin 6 to substrate determines differential phospholipid hydroperoxide peroxidase and phospholipase A(2) activities, *Arch Biochem Biophys., 485*, 139 (2009).
185. Chowdhury, I., Mo, Y., Gao, L., Kazi, A., Fisher, A. B., and Feinstein, S. I., Oxidant stress stimulates expression of the human peroxiredoxin 6 gene by a transcriptional mechanism involving an antioxidant response element, *Free Radic Biol Med., 46*, 146 (2009).
186. Metaxakis, A., Oehler, S., Klinakis, A., and Savakis, C., Minos as a genetic and genomic tool in *Drosophila* melanogaster, *Genetics., 171*, 571 (2005).
187. Dauer, W. T., and Worman, H. J., The nuclear envelope as a signaling node in development and disease, *Dev Cell., 17*, 626 (2009).
188. Basel-Vanagaite, L., Muncher, L., Straussberg, R., Pasmanik-Chor, M., Yahav, M., Rainshtein, L., Walsh, C. A., Magal, N., Taub, E., Drasinover, V., Shalev, H., Attia, R., Rechavi, G., Simon, A. J., and Shohat, M., Mutated nup62 causes autosomal recessive infantile bilateral striatal necrosis, *Ann Neurol., 60*, 214 (2006).
189. Sheffield, L. G., Miskiewicz, H. B., Tannenbaum, L. B., and Mirra, S. S., Nuclear pore complex proteins in Alzheimer disease, *J Neuropathol Exp Neurol., 65*, 45 (2006).
190. Suhr, S. T., Senut, M. C., Whitelegge, J. P., Faull, K. F., Cuizon, D. B., Gage, F. H., G√ºttinger, S., M√ºhlh√§usser, P., Koller-Eichhorn, R., Brennecke, J., and Kutay, U., Identities of Sequestered Proteins in Aggregates from Cells with Induced Polyglutamine Expression, *J Cell Biol., 153*, 283 (2001).
191. Kinoshita, Y., Ito, H., Hirano, A., Fujita, K., Wate, R., Nakamura, M., Kaneko, S., Nakano, S., and Kusaka, H., Nuclear contour irregularity and abnormal transporter protein distribution in anterior horn cells in amyotrophic lateral sclerosis, *J Neuropathol Exp Neurol., 68*, 1184 (2009).

192. Crampton, N., Kodiha, M., Shrivastava, S., Umar, R., and Stochaj, U., Oxidative stress inhibits nuclear protein export by multiple mechanisms that target FG nucleoporins and Crm1, *Mol Biol Cell., 20*, 5106 (2009).
193. Monshausen, M., Gehring, N. H., and Kosik, K. S., The mammalian RNA-binding protein Staufen2 links nuclear and cytoplasmic RNA processing pathways in neurons, *Neuromolecular Med., 6*, 127 (2004).
194. Kiebler, M. A., Jansen, R.-P., Dahm, R., and Macchi, P., A putative nuclear function for mammalian Staufen, *Trends in Biochemical Sciences, 30*, 228 (2005).
195. Dargemont, C., Schmidt-Zachmann, M. S., and Kuhn, L. C., Direct interaction of nucleoporin p62 with mRNA during its export from the nucleus, *J Cell Sci., 108*, 257 (1995).
196. Han, I., Roos, M. D., and Kudlow, J. E., Interaction of the transcription factor Sp1 with the nuclear pore protein p62 requires the C-terminal domain of p62, *J Cell Biochem., 68*, 50 (1998).
197. Capelson, M., Liang, Y., Schulte, R., Mair, W., Wagner, U., and Hetzer, M. W., Chromatin-bound nuclear pore components regulate gene expression in higher eukaryotes, *Cell., 140*, 372 (2010).
198. Kalverda, B., Pickersgill, H., Shloma, V. V., and Fornerod, M., Nucleoporins directly stimulate expression of developmental and cell-cycle genes inside the nucleoplasm, *Cell., 140*, 360 (2010).
199. Dansithong, W., Wolf, C. M., Sarkar, P., Paul, S., Chiang, A., Holt, I., Morris, G. E., Branco, D., Sherwood, M. C., Comai, L., Berul, C. I., and Reddy, S., Cytoplasmic CUG RNA foci are insufficient to elicit key DM1 features, *PLoS One., 3*, e3968. Epub 2008 Dec 18. (2008).
200. Mastroyiannopoulos, N. P., Chrysanthou, E., Kyriakides, T. C., Uney, J. B., Mahadevan, M. S., and Phylactou, L. A., The effect of myotonic dystrophy transcript levels and location on muscle differentiation, *Biochem Biophys Res Commun., 377*, 526 (2008).
201. Kim, D. H., Langlois, M. A., Lee, K. B., Riggs, A. D., Puymirat, J., and Rossi, J. J., HnRNP H inhibits nuclear export of mRNA containing expanded CUG repeats and a distal branch point sequence, *Nucleic Acids Res., 33*, 3866 (2005).
202. Van Dusen, C. M., Yee, L., McNally, L. M., and McNally, M. T., A glycine-rich domain of hnRNP H/F promotes nucleocytoplasmic shuttling and nuclear

import through an interaction with transportin 1, *Mol Cell Biol., 30*, 2552 (2010).

203. Luthi-Carter, R., Strand, A. D., Hanson, S. A., Kooperberg, C., Schilling, G., La Spada, A. R., Merry, D. E., Young, A. B., Ross, C. A., Borchelt, D. R., and Olson, J. M., Polyglutamine and transcription: gene expression changes shared by DRPLA and Huntington's disease mouse models reveal context-independent effects, *Hum Mol Genet., 11*, 1927 (2002).
204. Evert, B. O., Vogt, I. R., Vieira-Saecker, A. M., Ozimek, L., de Vos, R. A., Brunt, E. R., Klockgether, T., and Wullner, U., Gene expression profiling in ataxin-3 expressing cell lines reveals distinct effects of normal and mutant ataxin-3, *J Neuropathol Exp Neurol., 62*, 1006 (2003).
205. Sipione, S., Rigamonti, D., Valenza, M., Zuccato, C., Conti, L., Pritchard, J., Kooperberg, C., Olson, J. M., and Cattaneo, E., Early transcriptional profiles in huntingtin-inducible striatal cells by microarray analyses, *Hum Mol Genet., 11*, 1953 (2002).
206. Chou, A. H., Chen, C. Y., Chen, S. Y., Chen, W. J., Chen, Y. L., Weng, Y. S., and Wang, H. L., Polyglutamine-expanded ataxin-7 causes cerebellar dysfunction by inducing transcriptional dysregulation, *Neurochem Int, 10*, 10 (2009).
207. Chou, A. H., Yeh, T. H., Ouyang, P., Chen, Y. L., Chen, S. Y., and Wang, H. L., Polyglutamine-expanded ataxin-3 causes cerebellar dysfunction of SCA3 transgenic mice by inducing transcriptional dysregulation, *Neurobiol Dis., 31*, 89 (2008).
208. Runne, H., Regulier, E., Kuhn, A., Zala, D., Gokce, O., Perrin, V., Sick, B., Aebischer, P., Deglon, N., and Luthi-Carter, R., Dysregulation of gene expression in primary neuron models of Huntington's disease shows that polyglutamine-related effects on the striatal transcriptome may not be dependent on brain circuitry, *J Neurosci., 28*, 9723 (2008).
209. Lin, D. M., and Goodman, C. S., Ectopic and increased expression of Fasciclin II alters motoneuron growth cone guidance, *Neuron., 13*, 507 (1994).
210. Luo, L., Liao, Y. J., Jan, L. Y., and Jan, Y. N., Distinct morphogenetic functions of similar small GTPases: *Drosophila* Drac1 is involved in axonal outgrowth and myoblast fusion, *Genes Dev., 8*, 1787 (1994).
211. Panov, A. V., Gutekunst, C. A., Leavitt, B. R., Hayden, M. R., Burke, J. R., Strittmatter, W. J., and Greenamyre, J. T., Early mitochondrial calcium defects

in Huntington's disease are a direct effect of polyglutamines, *Nat Neurosci., 5*, 731 (2002).

212. Ocampo, A., Zambrano, A., and Barrientos, A., Suppression of polyglutamine-induced cytotoxicity in Saccharomyces cerevisiae by enhancement of mitochondrial biogenesis, *Faseb J, 14*, 14 (2009).
213. Yu, Z., Wang, A. M., Robins, D. M., and Lieberman, A. P., Altered RNA splicing contributes to skeletal muscle pathology in Kennedy disease knock-in mice, *Dis Model Mech., 2*, 500 (2009).
214. Morfini, G., Pigino, G., Szebenyi, G., You, Y., Pollema, S., and Brady, S. T., JNK mediates pathogenic effects of polyglutamine-expanded androgen receptor on fast axonal transport, *Nat Neurosci., 9*, 907 (2006).
215. Li, X., Standley, C., Sapp, E., Valencia, A., Qin, Z. H., Kegel, K. B., Yoder, J., Comer-Tierney, L. A., Esteves, M., Chase, K., Alexander, J., Masso, N., Sobin, L., Bellve, K., Tuft, R., Lifshitz, L., Fogarty, K., Aronin, N., and DiFiglia, M., Mutant huntingtin impairs vesicle formation from recycling endosomes by interfering with Rab11 activity, *Mol Cell Biol., 29*, 6106 (2009).
216. Nonis, D., Schmidt, M. H., van de Loo, S., Eich, F., Dikic, I., Nowock, J., and Auburger, G., Ataxin-2 associates with the endocytosis complex and affects EGF receptor trafficking, *Cell Signal., 20*, 1725 (2008).
217. Kegel, K. B., Sapp, E., Alexander, J., Valencia, A., Reeves, P., Li, X., Masso, N., Sobin, L., Aronin, N., and DiFiglia, M., Polyglutamine expansion in huntingtin alters its interaction with phospholipids, *J Neurochem., 110*, 1585 (2009).
218. Romero, E., Cha, G. H., Verstreken, P., Ly, C. V., Hughes, R. E., Bellen, H. J., and Botas, J., Suppression of neurodegeneration and increased neurotransmission caused by expanded full-length huntingtin accumulating in the cytoplasm, *Neuron., 57*, 27 (2008).
219. Joshi, P. R., Wu, N. P., Andre, V. M., Cummings, D. M., Cepeda, C., Joyce, J. A., Carroll, J. B., Leavitt, B. R., Hayden, M. R., Levine, M. S., and Bamford, N. S., Age-dependent alterations of corticostriatal activity in the YAC128 mouse model of Huntington disease, *J Neurosci., 29*, 2414 (2009).
220. Osborne, R. J., Lin, X., Welle, S., Sobczak, K., O'Rourke, J. R., Swanson, M. S., and Thornton, C. A., Transcriptional and post-transcriptional impact of toxic RNA in myotonic dystrophy, *Hum Mol Genet., 18*, 1471 (2009).

221. Yue, L., and Spradling, A. C., hu-li tai shao, a gene required for ring canal formation during *Drosophila* oogenesis, encodes a homolog of adducin, *Genes Dev., 6*, 2443 (1992).
222. Zaccai, M., and Lipshitz, H. D., Role of Adducin-like (hu-li tai shao) mRNA and protein localization in regulating cytoskeletal structure and function during *Drosophila* Oogenesis and early embryogenesis, *Dev Genet., 19*, 249 (1996).
223. Gubb, D., Green, C., Huen, D., Coulson, D., Johnson, G., Tree, D., Collier, S., and Roote, J., The balance between isoforms of the prickle LIM domain protein is critical for planar polarity in *Drosophila* imaginal discs, *Genes Dev., 13*, 2315 (1999).
224. De Gregorio, E., Spellman, P. T., Tzou, P., Rubin, G. M., and Lemaitre, B., The Toll and Imd pathways are the major regulators of the immune response in *Drosophila*, *Embo J., 21*, 2568 (2002).
225. Kraut, R., Menon, K., and Zinn, K., A gain-of-function screen for genes controlling motor axon guidance and synaptogenesis in *Drosophila*, *Curr Biol., 11*, 417 (2001).
226. Chen-Plotkin, A. S., Sadri-Vakili, G., Yohrling, G. J., Braveman, M. W., Benn, C. L., Glajch, K. E., DiRocco, D. P., Farrell, L. A., Krainc, D., Gines, S., MacDonald, M. E., and Cha, J. H., Decreased association of the transcription factor Sp1 with genes downregulated in Huntington's disease, *Neurobiol Dis., 22*, 233 (2006).
227. Dunah, A. W., Jeong, H., Griffin, A., Kim, Y. M., Standaert, D. G., Hersch, S. M., Mouradian, M. M., Young, A. B., Tanese, N., and Krainc, D., Sp1 and TAFII130 transcriptional activity disrupted in early Huntington's disease, *Science., 296*, 2238 (2002).
228. Astrinidis, A., Kim, J., Kelly, C. M., Olofsson, B. A., Torabi, B., Sorokina, E. M., and Azizkhan-Clifford, J., The transcription factor SP1 regulates centriole function and chromosomal stability through a functional interaction with the mammalian target of rapamycin/raptor complex, *Genes, 49*, 282 (2010).
229. Zaid, A., Li, R., Luciakova, K., Barath, P., Nery, S., and Nelson, B. D., On the role of the general transcription factor Sp1 in the activation and repression of diverse mammalian oxidative phosphorylation genes, *J Bioenerg Biomembr., 31*, 129 (1999).

230. Blesa, J. R., Prieto-Ruiz, J. A., Abraham, B. A., Harrison, B. L., Hegde, A. A., and Hernandez-Yago, J., NRF-1 is the major transcription factor regulating the expression of the human TOMM34 gene, *Biochem Cell Biol., 86*, 46 (2008).
231. Du, X. L., Edelstein, D., Rossetti, L., Fantus, I. G., Goldberg, H., Ziyadeh, F., Wu, J., and Brownlee, M., Hyperglycemia-induced mitochondrial superoxide overproduction activates the hexosamine pathway and induces plasminogen activator inhibitor-1 expression by increasing Sp1 glycosylation, *Proc Natl Acad Sci U S A., 97*, 12222 (2000).
232. Deniaud, E., Baguet, J., Chalard, R., Blanquier, B., Brinza, L., Meunier, J., Michallet, M. C., Laugraud, A., Ah-Soon, C., Wierinckx, A., Castellazzi, M., Lachuer, J., Gautier, C., Marvel, J., and Leverrier, Y., Overexpression of transcription factor Sp1 leads to gene expression perturbations and cell cycle inhibition, *PLoS One., 4*, e7035. (2009).
233. Krainc, D., Bai, G., Okamoto, S., Carles, M., Kusiak, J. W., Brent, R. N., and Lipton, S. A., Synergistic activation of the N-methyl-D-aspartate receptor subunit 1 promoter by myocyte enhancer factor 2C and Sp1, *J Biol Chem., 273*, 26218 (1998).
234. Liu, A., Zhuang, Z., Hoffman, P. W., and Bai, G., Functional analysis of the rat N-methyl-D-aspartate receptor 2A promoter: multiple transcription starts points, positive regulation by Sp factors, and translational regulation, *J Biol Chem., 278*, 26423 (2003).
235. Corti, C., Xuereb, J. H., Corsi, M., and Ferraguti, F., Identification and characterization of the promoter region of the GRM3 gene, *Biochem Biophys Res Commun., 286*, 381 (2001).
236. Ricketts, C., Zatyka, M., and Barrett, T., The characterisation of the human Wolfram syndrome gene promoter demonstrating regulation by Sp1 and Sp3 transcription factors, *Biochim Biophys Acta., 1759*, 367 (2006).
237. Chen, Y. M., Gerwin, C., and Sheng, Z. H., Dynein light chain LC8 regulates syntaphilin-mediated mitochondrial docking in axons, *J Neurosci., 29*, 9429 (2009).
238. Fejtova, A., Davydova, D., Bischof, F., Lazarevic, V., Altrock, W. D., Romorini, S., Schone, C., Zuschratter, W., Kreutz, M. R., Garner, C. C., Ziv, N. E., and Gundelfinger, E. D., Dynein light chain regulates axonal trafficking and synaptic levels of Bassoon, *J Cell Biol., 185*, 341 (2009).

239. Tsai, N. P., Tsui, Y. C., and Wei, L. N., Dynein motor contributes to stress granule dynamics in primary neurons, *Neuroscience., 159*, 647 (2009).
240. Satoh, D., Sato, D., Tsuyama, T., Saito, M., Ohkura, H., Rolls, M. M., Ishikawa, F., and Uemura, T., Spatial control of branching within dendritic arbors by dynein-dependent transport of Rab5-endosomes, *Nat Cell Biol., 10*, 1164 (2008).
241. Phillis, R., Statton, D., Caruccio, P., and Murphey, R. K., Mutations in the 8 kDa dynein light chain gene disrupt sensory axon projections in the *Drosophila* imaginal CNS, *Development., 122*, 2955 (1996).
242. Batlevi, Y., Martin, D. N., Pandey, U. B., Simon, C. R., Powers, C. M., Taylor, J. P., and Baehrecke, E. H., Dynein light chain 1 is required for autophagy, protein clearance, and cell death in *Drosophila*, *Proc Natl Acad Sci U S A, 107*, 742 (2009).
243. Dick, T., Ray, K., Salz, H. K., and Chia, W., Cytoplasmic dynein (ddlc1) mutations cause morphogenetic defects and apoptotic cell death in *Drosophila* melanogaster, *Mol Cell Biol., 16*, 1966 (1996).
244. Nelson, B., Nishimura, S., Kanuka, H., Kuranaga, E., Inoue, M., Hori, G., Nakahara, H., and Miura, M., Isolation of gene sets affected specifically by polyglutamine expression: implication of the TOR signaling pathway in neurodegeneration, *Cell Death Differ., 12*, 1115 (2005).
245. Piccioni, F., Pinton, P., Simeoni, S., Pozzi, P., Fascio, U., Vismara, G., Martini, L., Rizzuto, R., and Poletti, A., Androgen receptor with elongated polyglutamine tract forms aggregates that alter axonal trafficking and mitochondrial distribution in motor neuronal processes, *Faseb J., 16*, 1418 (2002).
246. Nadezhdina, E. S., Lomakin, A. J., Shpilman, A. A., Chudinova, E. M., and Ivanov, P. A., Microtubules govern stress granule mobility and dynamics, *Biochim Biophys Acta, 28*, 28 (2009).
247. Loschi, M., Leishman, C. C., Berardone, N., and Boccaccio, G. L., Dynein and kinesin regulate stress-granule and P-body dynamics, *J Cell Sci., 122*, 3973 (2009).
248. Repicky, S., and Broadie, K., Metabotropic glutamate receptor-mediated use-dependent down-regulation of synaptic excitability involves the fragile X mental retardation protein, *J Neurophysiol., 101*, 672 (2009).

249. Pan, L., and Broadie, K. S., *Drosophila* fragile X mental retardation protein and metabotropic glutamate receptor A convergently regulate the synaptic ratio of ionotropic glutamate receptor subclasses, *J Neurosci., 27*, 12378 (2007).
250. Shalizi, A., Gaudilliere, B., Yuan, Z., Stegmuller, J., Shirogane, T., Ge, Q., Tan, Y., Schulman, B., Harper, J. W., and Bonni, A., A calcium-regulated MEF2 sumoylation switch controls postsynaptic differentiation, *Science, 311*, 1012 (2006).
251. Levesque, D., and Rouillard, C., Nur77 and retinoid X receptors: crucial factors in dopamine-related neuroadaptation, *Trends Neurosci., 30*, 22 (2007).
252. Kanzleiter, T., Preston, E., Wilks, D., Ho, B., Benrick, A., Reznick, J., Heilbronn, L. K., Turner, N., and Cooney, G. J., Overexpression of the orphan receptor Nur77 alters glucose metabolism in rat muscle cells and rat muscle in vivo, *Diabetologia., 53*, 1174 (2010).
253. Liberski, P. P., and Budka, H., Neuroaxonal pathology in Creutzfeldt-Jakob disease, *Acta Neuropathol., 97*, 329 (1999).
254. Strom, A. L., Gal, J., Shi, P., Kasarskis, E. J., Hayward, L. J., and Zhu, H., Retrograde axonal transport and motor neuron disease, *J Neurochem., 106*, 495 (2008).
255. Ittner, L. M., Fath, T., Ke, Y. D., Bi, M., van Eersel, J., Li, K. M., Gunning, P., and Gotz, J., Parkinsonism and impaired axonal transport in a mouse model of frontotemporal dementia, *Proc Natl Acad Sci U S A., 105*, 15997 (2008).
256. Torroja, L., Chu, H., Kotovsky, I., and White, K., Neuronal overexpression of APPL, the *Drosophila* homologue of the amyloid precursor protein (APP), disrupts axonal transport, *Curr Biol., 9*, 489 (1999).
257. Poon, W. W., Blurton-Jones, M., Tu, C. H., Feinberg, L. M., Chabrier, M. A., Harris, J. W., Jeon, N. L., and Cotman, C. W., beta-Amyloid impairs axonal BDNF retrograde trafficking, *Neurobiol Aging, 18* (2009).
258. Hollenbeck, P. J., and Saxton, W. M., The axonal transport of mitochondria, *J Cell Sci., 118*, 5411 (2005).
259. Yonekawa, Y., Harada, A., Okada, Y., Funakoshi, T., Kanai, Y., Takei, Y., Terada, S., Noda, T., and Hirokawa, N., Defect in synaptic vesicle precursor transport and neuronal cell death in KIF1A motor protein-deficient mice, *J Cell Biol., 141*, 431 (1998).
260. Dupuis, L., Fergani, A., Braunstein, K. E., Eschbach, J., Holl, N., Rene, F., Gonzalez De Aguilar, J. L., Zoerner, B., Schwalenstocker, B., Ludolph, A. C.,

and Loeffler, J. P., Mice with a mutation in the dynein heavy chain 1 gene display sensory neuropathy but lack motor neuron disease, *Exp Neurol., 215*, 146 (2009).

261. DiProspero, N. A., Chen, E. Y., Charles, V., Plomann, M., Kordower, J. H., and Tagle, D. A., Early changes in Huntington's disease patient brains involve alterations in cytoskeletal and synaptic elements, *J Neurocytol., 33*, 517 (2004).
262. Kaminosono, S., Saito, T., Oyama, F., Ohshima, T., Asada, A., Nagai, Y., Nukina, N., and Hisanaga, S., Suppression of mutant Huntingtin aggregate formation by Cdk5/p35 through the effect on microtubule stability, *J Neurosci., 28*, 8747 (2008).
263. Sergeant, N., Sablonniere, B., Schraen-Maschke, S., Ghestem, A., Maurage, C. A., Wattez, A., Vermersch, P., and Delacourte, A., Dysregulation of human brain microtubule-associated tau mRNA maturation in myotonic dystrophy type 1, *Hum Mol Genet., 10*, 2143 (2001).
264. Wu, L. L., Fan, Y., Li, S., Li, X. J., and Zhou, X. F., Huntingtin-associated protein-1 interacts with pro-brain-derived neurotrophic factor and mediates its transport and release, *J Biol Chem., 285*, 5614 (2010).
265. Chang, D. T., Rintoul, G. L., Pandipati, S., and Reynolds, I. J., Mutant huntingtin aggregates impair mitochondrial movement and trafficking in cortical neurons, *Neurobiol Dis., 22*, 388 (2006).
266. Reddy, P. H., Mao, P., and Manczak, M., Mitochondrial structural and functional dynamics in Huntington's disease, *Brain Res Rev., 61*, 33 (2009).
267. Her, L. S., and Goldstein, L. S., Enhanced sensitivity of striatal neurons to axonal transport defects induced by mutant huntingtin, *J Neurosci., 28*, 13662 (2008).
268. Katsuno, M., Adachi, H., Minamiyama, M., Waza, M., Tokui, K., Banno, H., Suzuki, K., Onoda, Y., Tanaka, F., Doyu, M., and Sobue, G., Reversible disruption of dynactin 1-mediated retrograde axonal transport in polyglutamine-induced motor neuron degeneration, *J Neurosci., 26*, 12106 (2006).
269. Buchner, K., Roth, P., Schotta, G., Krauss, V., Saumweber, H., Reuter, G., and Dorn, R., Genetic and molecular complexity of the position effect variegation modifier mod(mdg4) in *Drosophila*, *Genetics, 155*, 141 (2000).

270. Daughters, R. S., Tuttle, D. L., Gao, W., Ikeda, Y., Moseley, M. L., Ebner, T. J., Swanson, M. S., and Ranum, L. P., RNA gain-of-function in spinocerebellar ataxia type 8, *PLoS Genet., 5*, e1000600. Epub 2009 Aug 14. (2009).
271. Leroy, O., Wang, J., Maurage, C. A., Parent, M., Cooper, T., Buee, L., Sergeant, N., Andreadis, A., and Caillet-Boudin, M. L., Brain-specific change in alternative splicing of Tau exon 6 in myotonic dystrophy type 1, *Biochim Biophys Acta., 1762*, 460 (2006).
272. Flavell, S. W., Kim, T. K., Gray, J. M., Harmin, D. A., Hemberg, M., Hong, E. J., Markenscoff-Papadimitriou, E., Bear, D. M., and Greenberg, M. E., Genome-wide analysis of MEF2 transcriptional program reveals synaptic target genes and neuronal activity-dependent polyadenylation site selection, *Neuron, 60*, 1022 (2008).
273. Artero, R., Prokop, A., Paricio, N., Begemann, G., Pueyo, I., Mlodzik, M., Perez-Alonso, M., and Baylies, M. K., The muscleblind gene participates in the organization of Z-bands and epidermal attachments of *Drosophila* muscles and is regulated by Dmef2, *Dev Biol, 195*, 131 (1998).
274. Pazyra-Murphy, M. F., Hans, A., Courchesne, S. L., Karch, C., Cosker, K. E., Heerssen, H. M., Watson, F. L., Kim, T., Greenberg, M. E., and Segal, R. A., A retrograde neuronal survival response: target-derived neurotrophins regulate MEF2D and bcl-w, *J Neurosci, 29*, 6700 (2009).
275. Pfeiffer, B. E., Zang, T., Wilkerson, J. R., Taniguchi, M., Maksimova, M. A., Smith, L. N., Cowan, C. W., and Huber, K. M., Fragile X mental retardation protein is required for synapse elimination by the activity-dependent transcription factor MEF2, *Neuron., 66*, 191 (2010).
276. Tian, X., Kai, L., Hockberger, P. E., Wokosin, D. L., and Surmeier, D. J., MEF-2 regulates activity-dependent spine loss in striatopallidal medium spiny neurons, *Mol Cell Neurosci., 44*, 94 (2010).
277. Bolger, T. A., Zhao, X., Cohen, T. J., Tsai, C. C., and Yao, T. P., The neurodegenerative disease protein ataxin-1 antagonizes the neuronal survival function of myocyte enhancer factor-2, *J Biol Chem, 282*, 29186 (2007).
278. Perdoni, F., Malatesta, M., Cardani, R., Giagnacovo, M., Mancinelli, E., Meola, G., and Pellicciari, C., RNA/MBNL1-containing foci in myoblast nuclei from patients affected by myotonic dystrophy type 2: an immunocytochemical study, *Eur J Histochem., 53*, 151 (2009).

279. Lin, X., Miller, J. W., Mankodi, A., Kanadia, R. N., Yuan, Y., Moxley, R. T., Swanson, M. S., and Thornton, C. A., Failure of MBNL1-dependent post-natal splicing transitions in myotonic dystrophy, *Hum Mol Genet., 15*, 2087 (2006).
280. Matsuura, T., Yamagata, T., Burgess, D. L., Rasmussen, A., Grewal, R. P., Watase, K., Khajavi, M., McCall, A. E., Davis, C. F., Zu, L., Achari, M., Pulst, S. M., Alonso, E., Noebels, J. L., Nelson, D. L., Zoghbi, H. Y., and Ashizawa, T., Large expansion of the ATTCT pentanucleotide repeat in spinocerebellar ataxia type 10, *Nat Genet, 26*, 191 (2000).
281. Almeida, T., Alonso, I., Martins, S., Ramos, E. M., Azevedo, L., Ohno, K., Amorim, A., Saraiva-Pereira, M. L., Jardim, L. B., Matsuura, T., Sequeiros, J., and Silveira, I., Ancestral origin of the ATTCT repeat expansion in spinocerebellar ataxia type 10 (SCA10), *PLoS ONE, 4*, e4553 (2009).
282. Lin, X., and Ashizawa, T., SCA10 and ATTCT repeat expansion: clinical features and molecular aspects, *Cytogenet Genome Res, 100*, 184 (2003).
283. Matsuura, T., Fang, P., Lin, X., Khajavi, M., Tsuji, K., Rasmussen, A., Grewal, R. P., Achari, M., Alonso, M. E., Pulst, S. M., Zoghbi, H. Y., Nelson, D. L., Roa, B. B., and Ashizawa, T., Somatic and germline instability of the ATTCT repeat in spinocerebellar ataxia type 10, *Am J Hum Genet, 74*, 1216 (2004).
284. Potaman, V. N., Bissler, J. J., Hashem, V. I., Oussatcheva, E. A., Lu, L., Shlyakhtenko, L. S., Lyubchenko, Y. L., Matsuura, T., Ashizawa, T., Leffak, M., Benham, C. J., and Sinden, R. R., Unpaired structures in SCA10 (ATTCT)n.(AGAAT)n repeats, *J Mol Biol, 326*, 1095 (2003).
285. Liu, G., Bissler, J. J., Sinden, R. R., and Leffak, M., Unstable spinocerebellar ataxia type 10 (ATTCT*(AGAAT) repeats are associated with aberrant replication at the ATX10 locus and replication origin-dependent expansion at an ectopic site in human cells, *Mol Cell Biol., 27*, 7828 (2007).
286. Hagerman, K. A., Ruan, H., Edamura, K. N., Matsuura, T., Pearson, C. E., and Wang, Y. H., The ATTCT repeats of spinocerebellar ataxia type 10 display strong nucleosome assembly which is enhanced by repeat interruptions, *Gene., 434*, 29 (2009).
287. Colin, E., Zala, D., Liot, G., Rangone, H., Borrell-Pages, M., Li, X. J., Saudou, F., and Humbert, S., Huntingtin phosphorylation acts as a molecular switch for anterograde/retrograde transport in neurons, *Embo J., 27*, 2124 (2008).

288. Wakamiya, M., Matsuura, T., Liu, Y., Schuster, G. C., Gao, R., Xu, W., Sarkar, P. S., Lin, X., and Ashizawa, T., The role of ataxin 10 in the pathogenesis of spinocerebellar ataxia type 10, *Neurology, 67*, 607 (2006).
289. Taneja, K. L., McCurrach, M., Schalling, M., Housman, D., and Singer, R. H., Foci of trinucleotide repeat transcripts in nuclei of myotonic dystrophy cells and tissues, *J Cell Biol., 128*, 995 (1995).
290. Davis, B. M., McCurrach, M. E., Taneja, K. L., Singer, R. H., and Housman, D. E., Expansion of a CUG trinucleotide repeat in the 3' untranslated region of myotonic dystrophy protein kinase transcripts results in nuclear retention of transcripts, *PNAS, 94*, 7388 (1997).
291. St Johnston, D., Beuchle, D., and Nusslein-Volhard, C., Staufen, a gene required to localize maternal RNAs in the *Drosophila* egg, *Cell., 66*, 51 (1991).
292. Li, P., Yang, X., Wasser, M., Cai, Y., and Chia, W., Inscuteable and Staufen mediate asymmetric localization and segregation of prospero RNA during *Drosophila* neuroblast cell divisions, *Cell., 90*, 437 (1997).
293. Dubnau, J., Chiang, A. S., Grady, L., Barditch, J., Gossweiler, S., McNeil, J., Smith, P., Buldoc, F., Scott, R., Certa, U., Broger, C., and Tully, T., The staufen/pumilio pathway is involved in *Drosophila* long-term memory, *Curr Biol., 13*, 286 (2003).
294. Macchi, P., Brownawell, A. M., Grunewald, B., DesGroseillers, L., Macara, I. G., and Kiebler, M. A., The brain-specific double-stranded RNA-binding protein Staufen2: nucleolar accumulation and isoform-specific exportin-5-dependent export, *J Biol Chem., 279*, 31440 (2004).
295. Sandmann, T., Jensen, L. J., Jakobsen, J. S., Karzynski, M. M., Eichenlaub, M. P., Bork, P., and Furlong, E. E., A temporal map of transcription factor activity: mef2 directly regulates target genes at all stages of muscle development, *Dev Cell, 10*, 797 (2006).
296. Wang, X., She, H., and Mao, Z., Phosphorylation of neuronal survival factor MEF2D by glycogen synthase kinase 3beta in neuronal apoptosis, *J Biol Chem, 284*, 32619 (2009).
297. Bassell, G. J., and Warren, S. T., Fragile X syndrome: loss of local mRNA regulation alters synaptic development and function, *Neuron, 60*, 201 (2008).
298. Liang, B., Song, X., Liu, G., Li, R., Xie, J., Xiao, L., Du, M., Zhang, Q., Xu, X., Gan, X., and Huang, D., Involvement of TR3/Nur77 translocation to the

endoplasmic reticulum in ER stress-induced apoptosis, *Exp Cell Res., 313*, 2833 (2007).

299. Lam, B. Y., Zhang, W., Ng, D. C., Maruthappu, M., Roderick, H. L., and Chawla, S., CREB-dependent Nur77 induction following depolarization in PC12 cells and neurons is modulated by MEF2 transcription factors, *J Neurochem., 112*, 1065 (2010).
300. McCampbell, A., Taylor, J. P., Taye, A. A., Robitschek, J., Li, M., Walcott, J., Merry, D., Chai, Y., Paulson, H., Sobue, G., and Fischbeck, K. H., CREB-binding protein sequestration by expanded polyglutamine, *Hum Mol Genet, 9*, 2197 (2000).
301. Iijima-Ando, K., Wu, P., Drier, E. A., Iijima, K., and Yin, J. C. P., cAMP-response element-binding protein and heat-shock protein 70 additively suppress polyglutamine-mediated toxicity in *Drosophila*, *PNAS, 102*, 10261 (2005).
302. Jiang, H., Poirier, M. A., Liang, Y., Pei, Z., Weiskittel, C. E., Smith, W. W., DeFranco, D. B., and Ross, C. A., Depletion of CBP is directly linked with cellular toxicity caused by mutant huntingtin, *Neurobiol Dis, 23*, 543 (2006).
303. Grimes, C. A., and Jope, R. S., CREB DNA binding activity is inhibited by glycogen synthase kinase-3 beta and facilitated by lithium, *J Neurochem, 78*, 1219 (2001).
304. Min, W. W., Yuskaitis, C. J., Yan, Q., Sikorski, C., Chen, S., Jope, R. S., and Bauchwitz, R. P., Elevated glycogen synthase kinase-3 activity in Fragile X mice: key metabolic regulator with evidence for treatment potential, *Neuropharmacology, 56*, 463 (2009).
305. Hernandez, F., Gomez de Barreda, E., Fuster-Matanzo, A., Lucas, J. J., and Avila, J., GSK3: A possible link between beta amyloid peptide and tau protein, *Exp Neurol* (2009).
306. Lee, J., and Kim, M. S., The role of GSK3 in glucose homeostasis and the development of insulin resistance, *Diabetes Res Clin Pract, 77 Suppl 1*, S49 (2007).
307. Karim, R., Tse, G., Putti, T., Scolyer, R., and Lee, S., The significance of the Wnt pathway in the pathology of human cancers, *Pathology, 36*, 120 (2004).
308. Dominguez, I., and Green, J. B., Missing links in GSK3 regulation, *Dev Biol., 235*, 303 (2001).

309. Takashima, A., GSK-3 is essential in the pathogenesis of Alzheimer's disease, *J Alzheimers Dis., 9*, 309 (2006).
310. Palazzolo, I., Burnett, B. G., Young, J. E., Brenne, P. L., La Spada, A. R., Fischbeck, K. H., Howell, B. W., and Pennuto, M., Akt blocks ligand binding and protects against expanded polyglutamine androgen receptor toxicity, *Hum Mol Genet., 16*, 1593 (2007).
311. Palazzolo, I., Stack, C., Kong, L., Musaro, A., Adachi, H., Katsuno, M., Sobue, G., Taylor, J. P., Sumner, C. J., Fischbeck, K. H., and Pennuto, M., Overexpression of IGF-1 in muscle attenuates disease in a mouse model of spinal and bulbar muscular atrophy, *Neuron., 63*, 316 (2009).
312. Branco, J., Al-Ramahi, I., Ukani, L., Perez, A. M., Fernandez-Funez, P., Rincon-Limas, D., and Botas, J., Comparative analysis of genetic modifiers in *Drosophila* points to common and distinct mechanisms of pathogenesis among polyglutamine diseases, *Hum Mol Genet., 17*, 376 (2008).
313. Hernandez-Hernandez, O., Bermudez-de-Leon, M., Gomez, P., Velazquez-Bernardino, P., Garcia-Sierra, F., and Cisneros, B., Myotonic dystrophy expanded CUG repeats disturb the expression and phosphorylation of tau in PC12 cells, *J Neurosci Res., 84*, 841 (2006).
314. Jorgensen, N. D., Andresen, J. M., Lagalwar, S., Armstrong, B., Stevens, S., Byam, C. E., Duvick, L. A., Lai, S., Jafar-Nejad, P., Zoghbi, H. Y., Clark, H. B., and Orr, H. T., Phosphorylation of ATXN1 at Ser776 in the cerebellum, *J Neurochem., 110*, 675 (2009).
315. Pastori, V., Sangalli, E., Coccetti, P., Pozzi, C., Nonnis, S., Tedeschi, G., and Fusi, P., CK2 and GSK3 phosphorylation on S29 controls wild-type ATXN3 nuclear uptake, *Biochim Biophys Acta, 27*, 27 (2010).
316. Xilouri, M., and Papazafiri, P., Induction of Akt by endogenous neurosteroids and calcium sequestration in P19 derived neurons, *Neurotox Res., 13*, 209 (2008).
317. Almeida, R. D., Manadas, B. J., Melo, C. V., Gomes, J. R., Mendes, C. S., Graos, M. M., Carvalho, R. F., Carvalho, A. P., and Duarte, C. B., Neuroprotection by BDNF against glutamate-induced apoptotic cell death is mediated by ERK and PI3-kinase pathways, *Cell Death Differ., 12*, 1329 (2005).
318. Nguyen, N., Lee, S. B., Lee, Y. S., Lee, K. H., and Ahn, J. Y., Neuroprotection by NGF and BDNF against neurotoxin-exerted apoptotic death in neural stem

cells are mediated through Trk receptors, activating PI3-kinase and MAPK pathways, *Neurochem Res., 34*, 942 (2009).

319. Ishiuchi, S., Yoshida, Y., Sugawara, K., Aihara, M., Ohtani, T., Watanabe, T., Saito, N., Tsuzuki, K., Okado, H., Miwa, A., Nakazato, Y., and Ozawa, S., Ca2+-permeable AMPA receptors regulate growth of human glioblastoma via Akt activation, *J Neurosci., 27*, 7987 (2007).
320. Pekarsky, Y., Hallas, C., Palamarchuk, A., Koval, A., Bullrich, F., Hirata, Y., Bichi, R., Letofsky, J., and Croce, C. M., Akt phosphorylates and regulates the orphan nuclear receptor Nur77, *Proc Natl Acad Sci U S A., 98*, 3690 (2001).
321. Yuskaitis, C. J., Mines, M. A., King, M. K., Sweatt, J. D., Miller, C. A., and Jope, R. S., Lithium ameliorates altered glycogen synthase kinase-3 and behavior in a mouse model of fragile X syndrome, *Biochem Pharmacol, 79*, 632 (2010).
322. Ribeiro, F. M., Paquet, M., Ferreira, L. T., Cregan, T., Swan, P., Cregan, S. P., and Ferguson, S. S., Metabotropic glutamate receptor-mediated cell signaling pathways are altered in a mouse model of Huntington's disease, *J Neurosci., 30*, 316 (2010).
323. Milnerwood, A. J., Gladding, C. M., Pouladi, M. A., Kaufman, A. M., Hines, R. M., Boyd, J. D., Ko, R. W., Vasuta, O. C., Graham, R. K., Hayden, M. R., Murphy, T. H., and Raymond, L. A., Early increase in extrasynaptic NMDA receptor signaling and expression contributes to phenotype onset in Huntington's disease mice, *Neuron., 65*, 178 (2010).
324. Giralt, A., Rodrigo, T., Martin, E. D., Gonzalez, J. R., Mila, M., Cena, V., Dierssen, M., Canals, J. M., and Alberch, J., Brain-derived neurotrophic factor modulates the severity of cognitive alterations induced by mutant huntingtin: involvement of phospholipaseCgamma activity and glutamate receptor expression, *Neuroscience., 158*, 1234 (2009).
325. Song, C., Perides, G., and Liu, Y. F., Expression of full-length polyglutamine-expanded Huntingtin disrupts growth factor receptor signaling in rat pheochromocytoma (PC12) cells, *J Biol Chem., 277*, 6703 (2002).
326. Ortega, F., Perez-Sen, R., Morente, V., Delicado, E. G., and Miras-Portugal, M. T., P2X7, NMDA and BDNF receptors converge on GSK3 phosphorylation and cooperate to promote survival in cerebellar granule neurons, *Cell Mol Life Sci., 67*, 1723 (2010).

327. Tongiorgi, E., and Baj, G., Functions and mechanisms of BDNF mRNA trafficking, *Novartis Found Symp., 289*, 136 (2008).
328. Franco, B., Bogdanik, L., Bobinnec, Y., Debec, A., Bockaert, J., Parmentier, M. L., and Grau, Y., Shaggy, the homolog of glycogen synthase kinase 3, controls neuromuscular junction growth in *Drosophila*, *J Neurosci, 24*, 6573 (2004).
329. Chiang, A., Priya, R., Ramaswami, M., Vijayraghavan, K., and Rodrigues, V., Neuronal activity and Wnt signaling act through Gsk3-beta to regulate axonal integrity in mature *Drosophila* olfactory sensory neurons, *Development, 136*, 1273 (2009).
330. Stocker, H., and Hafen, E., Genetic control of cell size, *Curr Opin Genet Dev, 10*, 529 (2000).
331. Ghosh, S., and Feany, M. B., Comparison of pathways controlling toxicity in the eye and brain in *Drosophila* models of human neurodegenerative diseases, *Hum Mol Genet., 13*, 2011 (2004).
332. Ladd, P. D., Smith, L. E., Rabaia, N. A., Moore, J. M., Georges, S. A., Hansen, R. S., Hagerman, R. J., Tassone, F., Tapscott, S. J., and Filippova, G. N., An antisense transcript spanning the CGG repeat region of FMR1 is upregulated in premutation carriers but silenced in full mutation individuals, *Hum. Mol. Genet., 16*, 3174 (2007).

Amendments:

The following changes and additions have been made to this thesis:

1. ***Line 1, Page 155*** *should read* "The principal aim of this study was…"

2. *The following text should be inserted at the end of the Introduction* ***(Page 26)*** *to clarify the "working hypothesis" and specific hypotheses tested in this study:*

"This thesis investigates the role of hairpin-forming RNA species in pathogenesis of the expanded repeat diseases. It specifically tests the intrinsic toxicity of untranslated expanded repeat sequences and their ability to induce cellular dysfunction in a *Drosophila* model. The "working hypothesis" of this work is that expression of hairpin-forming RNA species causes cellular dysfunction which contributes to pathogenesis in both the polyglutamine diseases and the untranslated repeat diseases.

Several separate hypotheses have been tested in this thesis:

- In Chapter 3, a role for RNA editing in pathogenesis of CAG repeat RNAs is investigated. This hypothesis is based on the observation that RNA editing is essential for survival of a subset of neurons which are amongst those most affected in some of the expanded repeat diseases, as well as the prediction that the structure formed by CAG repeat RNA may be a target for editing.
- In Chapter 4 and 5, an investigation of cellular effects of expression of CAG and CUG repeat RNAs was performed by proteomic and microarray analyses respectively, based on the hypothesis that the similar structures of these RNA species may result in similar cellular outcomes. The ability of candidates from the microarray analyses to interact with expanded repeats in this model was then validated, as described in Chapter 6.
- Since the expanded repeat responsible for SCA10 is also predicted to form a similar hairpin secondary structure, a *Drosophila* model of expression of the associated repeat sequence, an AUUCU repeat, was also generated in the course of this study. Chapter 7 describes this model and investigates cellular effects of expressing this sequence.

- As a result of the outcomes of investigation of cellular effects of expression of different expanded repeat sequences, the hypothesis that CAG, CUG and AUUCU repeat RNAs perturb the Akt/Gsk3β pathway was tested. This data is described in Chapter 8 of this thesis."

3. *The following text should be inserted before the Discussion* ***(Page 155)*** *to reaffirm the outcomes of the thesis:*

"This study has produced data which supports the hypothesis that expression of expanded repeat RNAs is able to cause cellular dysfunction. This appears to be an intrinsic property of the expanded repeat sequences.

- A role for RNA editing in CAG repeat pathogenesis was not supported by the results described in Chapter 3.
- A number of candidates involved in expanded repeat pathogenesis were identified in Chapter 4 and 5. Several of these also showed an interaction with expanded repeat RNA, as described in Chapter 4 and 6.
- The SCA10 AUUCU repeat sequence was demonstrated to act in a similar manner to the CAG and CUG repeat sequences, both in its ability to form cellular foci and in the sort of cellular changes which expression produced. This data is described in Chapter 7.
- A role for the Akt/Gsk3β pathway in expanded repeat pathogenesis was supported by genetic validation in Chapter 8."

4. *The following text should be added to the end of the Discussion* ***(Page 160):***

"While this study describes a number of cellular outcomes of expanded repeat expression in *Drosophila*, no evidence of neurodegeneration was observed in this model. The absence of severe effects in this model may suggest that these sequences are not highly toxic in *Drosophila*, possibly as a result of short lifespan. It may therefore be appropriate to further investigate neurodegenerative phenotypes, including ataxia, resulting from expression of hairpin RNAs in a mouse model."