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.