Double-stranded RNA as a pathogenic agent in a *Drosophila* model of dominant expanded repeat diseases

A thesis submitted for the degree of Doctor of Philosophy

November 2012

Saumya Erandathi Samaraweera, BSc. (Hons.)

School of Molecular and Biomedical Science, Discipline of Genetics

The University of Adelaide

Table of contents

Index of Figures and Tables ..... **VII**
Declaration ..... **XI**
Acknowledgements ..... **XIII**
Abbreviations ..... **XV**
Nomenclature ..... **XVIII**
Abstract ..... **XIX**

CHAPTER 1: Introduction **1**
1.1 Repeat expansion diseases 1
1.2 Translated repeat diseases 2
1.2.1 Polyglutamine diseases 2
1.2.2 Pathogenesis of polyglutamine diseases 6
1.3 Untranslated repeat diseases 7
1.3.1 Dominantly inherited untranslated repeat diseases 7
1.4 RNA as a common pathogenic agent 12
1.4.1 RNA structure of expanded repeats 12
1.4.2 The role of RNA binding proteins 13
1.5 Dicer mediated RNA processing 14
1.6 Immunity and expanded repeat diseases 16
1.7 Bi-directional transcription 16
1.8 *Drosophila* as a model organism for neurodegenerative diseases 20
1.9 A *Drosophila* expanded repeat disease model 21
1.10 A model for bi-directional transcription of expanded repeats 24
CHAPTER 2: Materials and Methods **27**
2.1 Materials 27
2.2 Methods 36
CHAPTER 3: Characterisation of double-stranded RNA as a pathogenic agent in *Drosophila* **47**
3.1 Expanded repeat-containing double-stranded RNA is toxic in *Drosophila* 49
3.2 Double-stranded RNA causes an age-dependent motor defect in *Drosophila* 52
3.3 Characterisation of double-stranded RNA mediated cell death 56
3.4 Autophagy restricts double-stranded RNA toxicity 62
3.5 Chapter discussion 64

CHAPTER 4: The role of small RNA pathways in double-stranded RNA mediated toxicity 67
4.1 Double-stranded RNA mediated toxicity is dependent on Dicer-2 70
4.2 Double-stranded RNA mediated toxicity is dependent on Argonaute-2 activity 75
4.3 R2D2 plays a role in double-stranded RNA mediated toxicity 77
4.4 Dicer-1 and Loquacious have opposing effects on double-stranded RNA mediated toxicity 80
4.5 Argonaute-1 is not rate-limiting for double-stranded RNA mediated pathogenesis 82
4.6 Chapter discussion 84
CHAPTER 5: Determining the small RNA signature and its role in double-stranded RNA induced toxicity 87
5.1 Complementary expanded repeat RNA is processed to CAG 21mer 89
5.2 Identification of potential downstream targets of repeat-derived small RNAs .. 96
5.3 Investigating the possible targeting of repeat-containing transcripts by repeat-derived small RNA 98
5.4. Investigating the pathogenic effect of repeat-derived small RNA on downstream target genes 104
5.5 Expression of double-stranded RNA alters the endogenous miRNA profile .. 110
5.6 Investigating the effect of miRNA changes on double-stranded RNA mediated pathogenesis in *Drosophila* 114
5.7 The role of potential miRNA target proteins in double-stranded RNA mediated toxicity 117
5.8 Preliminary small RNA profiling of Huntington's disease (HD) patient samples 119
5.9 Chapter discussion 123
CHAPTER 6: Identifying novel pathways of pathogenesis in dominant expanded repeat diseases by microarray analysis 127
6.1 Determine transcriptional changes in neuronal cells expressing double-stranded RNA 128
6.2 Significant transcriptional changes caused by double-stranded RNA 129
6.3 Genetic verification of significantly changed candidates 133
6.4 Significant transcriptional changes of repeat-containing genes 136
6.5 Chapter discussion 139

CHAPTER 7: Immune activation in double-stranded RNA induced toxicity 141
7.1 The Toll pathway is required for double-stranded RNA mediated toxicity 143
7.2 Pharmacological alteration of the Toll pathway effects double-stranded RNA toxicity 147
7.3 Glial expression of double-stranded RNA is toxic 151
7.4 Chapter discussion 153
CHAPTER 8: Discussion.....155
8.1 Summary of results.....155
8.2 Implications for dominant expanded repeat diseases.....157
8.3 Limitations of this study 158
8.4 Future directions.....159
Appendices.....161
Appendix A 161
Appendix B 163
Appendix C.....168
Appendix D.....199
References.....207

Index of Figures and Tables

Chapter 1

Figure 1.1: Schematic representation of expanded repeat diseases..... 2
Table 1.1: Translated repeat diseases 3
Table 1.2: Dominantly inherited untranslated repeat diseases..... 8
Figure 1.2: Predicted secondary structure of expanded CNG repeat RNA 13
Figure 1.3: Dicer mediated RNA processing and gene regulation in *Drosophila*..... 15
Figure 1.4: Bi-directional transcription..... 17
Figure 1.5: Bi-directional transcription at the SCA8 locus. 18
Figure 1.6: Expression of translated and untranslated expanded repeat sequences 23
Figure 1.7: Ubiquitous expression of rCAG$_{\sim 100}$ or rCUG$_{\sim 100}$ causes a tergite phenotype 23
Figure 1.8: GAL4/UAS expression of repeat constructs..... 24

Chapter 3

Figure 3.1: Expression of double-stranded RNA is toxic in the *Drosophila* eye 50
Figure 3.2: Range of double-stranded RNA toxicity with independent rCAG.rCUG$_{\sim 100}$ lines..... 51
Figure 3.3: Double-stranded RNA toxicity is length dependent..... 51
Figure 3.4: qRT-PCR levels of repeat transcript in rCAG.rCUG$_{\sim 50}$ lines 52
Figure 3.5: Apparatus for climbing assays 53
Figure 3.6: Expression of double-stranded RNA results in age-dependent neuronal pathogenesis..... 54
Table 3.1: Average number of flies that ‘fail to climb’ with repeat expression using the *elavII-GAL4* driver. 55
Table 3.2: Average number of flies that ‘fail to climb’ with repeat expression using the *elav^{C155}-GAL4* driver. 56
Figure 3.7: Expression of double-stranded RNA caused age-dependent neuronal pathogenesis with an independent neuronal driver. 56
Figure 3.8: Caspase dependent cell death in *Drosophila*. 57
Figure 3.9: Loss of Dronc has no effect on double-stranded RNA mediated toxicity. 58
Figure 3.10: Drice contributes to double-stranded RNA mediated toxicity 59

Figure 3.11: DIAP1 has an effect on double-stranded RNA mediated toxicity of milder lines..... 60
Figure 3.12: P35 has different effects on double-stranded RNA mediated toxicity suggesting a toxicity threshold. 61
Figure 3.13: Autophagy restricts double-stranded RNA toxicity. 63

Chapter 4

Figure 4.1: Drosophila small RNA biogenesis and RNA interference pathways 68
Figure 4.2: Double-stranded RNA mediated toxicity is dependent on Dicer-2 levels. 71
Figure 4.3: Heterozygous loss of Dicer-2 restores internal patterning of the eye..... 72
Figure 4.4: Dicer-2 dependent toxicity in weak rCAG.rCUG$_{\sim 100}$ lines 73
Figure 4.5: Dicer-2 mediated toxicity of double-stranded RNA is length dependent . 74
Figure 4.6: Over-expression of Dicer-2 has no effect on single-stranded RNA expressing eyes 75
Figure 4.7: Argonaute-2 is a rate-limiting factor in double-stranded RNA mediated toxicity. 76
Figure 4.8: R2D2 plays a significant role in double-stranded RNA mediated toxicity 78
Figure 4.9: Dicer-1 and Loquacious have opposite effects on double-stranded RNA mediated toxicity. 81
Figure 4.10: Argonaute-1 has no effect on double-stranded RNA mediated toxicity. 83

Chapter 5

Table 5.1: Trinucleotide repeat reads observed in the small RNA fractions of *Drosophila* deep sequencing experiments 90
Table 5.2: Tandem CAG and CTG repeat reads- *Drosophila* Experiment 1 92
Table 5.3: Tandem CAG repeat reads- *Drosophila* Experiment 2..... 94
Table 5.4: Summary of CAG and CTG repeat reads in *Drosophila* experiments 96
Table 5.5: Genes identified by bioinformatics to containing eight or more tandem CUG or CAG repeats 98
Table 5.6: Expression levels of CUG repeat containing transcripts 100
Table 5.7: Expression levels of CAG repeat containing transcripts..... 101
Table 5.8: Significantly altered genes containing CAG repeats identified by microarray analysis 103
Table 5.9: Expression levels of CAG repeat containing transcripts identified by microarray analyses 103

Figure 5.1: RNAi knockdown of tandem CUG-repeat containing genes has little effect on double-stranded RNA toxicity. 105
Figure 5.2: RNAi knockdown of tandem CAG-repeat containing genes. 107
Figure 5.3: RNAi knockdown of tandem CAG-repeat containing genes significantly altered in microarray analysis 108
Table 5.10: List of *Drosophila* miRNAs showing significant changes 112
Table 5.11: List of *Drosophila* miRNAs showing significant changes 113
Figure 5.4: Altered miRNAs have a minimal effect on double-stranded RNA toxicity 116
Table 5.12: Genes targeted by at least two of the altered miRNAs 118
Figure 5.5: Altered miRNA targets have a minimal effect on double-stranded RNA toxicity 119
Table 5.13: Summary of CAG or CTG repeat reads in human patient samples 120
Table 5.14: List of human miRNAs showing significant changes 122

Chapter 6

Figure 6.1: qRT-PCR levels of repeat transcripts in rCAG.rCUG$_{\sim 100}$ samples used in the microarray 129
Figure 6.2: Comparison of microarray data 129
Figure 6.3: Gene ontology analysis of significantly altered genes in flies expressing dsRNA pan-neuronally 130
Table 6.1: Top 10 significantly increased genes 131
Table 6.2: Top 10 significantly decreased genes 132
Figure 6.4: RNAi knockdown of genes showing a significant increase in expression in rCAG.rCUG$_{\sim 100}$ flies 134
Figure 6.5: RNAi knockdown of genes showing a significant decrease in expression in rCAG.rCUG$_{\sim 100}$ flies 135
Table 6.3: Significantly altered genes containing a tandem repeat tract 137

Chapter 7

Figure 7.1: Schematic representation of the *Drosophila* Toll pathway. 143
Figure 7.2: qRT-PCR levels of Drosomycin 144
Figure 7.3: Knockdown of *Drosophila* Toll receptors reduces double-stranded RNA toxicity. 145
Figure 7.4: Double-stranded RNA toxicity is dependent on the Toll pathway. 146

Figure 7.5: Exposure to pharmaceutical compounds has no effect on the viability of rCAG.rCUG$_{\sim 100}$ flies. 149
Figure 7.6: Typical phenotypes observed due to supplementation with pharmaceutical compounds. 149
Figure 7.7: Proportion of double-stranded RNA expressing flies showing a phenotype in addition to a rough eye following exposure to pharmaceutical compounds..... 150
Table 7.1: Average number of flies that ‘fail to climb’ with repeat expression using the *repo-GAL4* driver. 152
Figure 7.8: Expression of double-stranded RNA in glial cells results in a neurodegenerative phenotype 152

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 to Saumya Samaraweera 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 made 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, and also through web search engines, unless permission has been granted by the University to restrict access for a period of time.

Saumya Samaraweera Date

Acknowledgements

I would like to thank my supervisor, Rob Richards, for advice and support throughout this project. I am also extremely grateful to my co-supervisor Louise O'Keefe, for excellent mentorship, expert advice, constant encouragement and confidence in me. Thank you to all the members of the Richards lab, past and present, and members of the Discipline of Genetics, for support during this project and for creating an excellent work environment. Special thanks to Clare van Eyk and Amanda Choo for their friendship and good humour throughout the years.

I also wish to thank Lisa Nicholas for her friendship and understanding, as well as the many discussions over coffee, both scientific and otherwise. To Nivasinee Sreetharan and Paul Cavuoto, thank you for being my first port of call, and for the selfless effort of proofing this thesis. Finally I am grateful to my family for their love and encouragement and for teaching me to strive to be the best that I can.

Abbreviations

°C: degrees Celsius
µg: microgram
µl: microlitre
µm: micrometre
18W: 18-wheeler
ADAR: Adenosine deaminase acting on RNA
AGO: Argonaute
ALS: Amyotrophic lateral sclerosis
AR: Androgen receptor
Ark: Apaf1-related-killer
Atg: autophagy-specific gene
ATN: Atrophin
Atx: Ataxin (*Drosophila*)
ATXN: Ataxin
BEAN: brain expressed, associated with Nedd4
cact: cactus
c-cup: calcutta cup
cDNA: complementary DNA
CLCN-1: Chloride channel 1
cpo: couch potato
CTCF: CCCTC-binding factor
CUG-BP: CUG binding protein
Dcr: Dicer
DEPC: diethyl pyrocarbonate
DIAP: *Drosophila* inhibitor of apoptosis
Dif: dorsal-related immunity factor
dl: dorsal
DM: Myotonic dystrophy
DMPK: Dystrophia myotonica protein kinase
dnc: dunce
DRPLA: Dentatorubral-pallidoluysian atrophy
Dscam: Down syndrome cell adhesion molecule

dsRNA: double-stranded RNA
EDTA: ethylene diamine tetra-acetic acid
eIF2: eukaryotic initiation factor 2
elav: embryonic lethal abnormal vision
endo-siRNA: endogenous siRNA
exo-siRNA: exogenous siRNA
FMR1: Fragile X mental retardation 1
FRAXE: Fragile XE syndrome
FTD: Frontotemporal dementia
FXS: Fragile X syndrome
FXTAS: Fragile X tremor-ataxia syndrome
GABA: gamma-aminobutyric acid
Gen: XPG-like endonuclease
GFP: green fluorescence protein
GluR: glutamate receptor
Glut1: glucose transporter-1
GMR: glass multimer reporter
GSK3: Glycogen synthase kinase 3
HD: Huntington's disease
HDL2: Huntington's disease-like-2
hid: head involusion factor
hnRNP: Heterogenous nuclear ribonucleoprotein
HTT: Huntingtin
IL: Interleukin
IM: immune-induced molecule
IR: Insulin receptor
jagn: jagunal
KLHL1: Kelch-like 1
LB: Luria broth
Loqs: Loquacious
M: molar
MBNL: Muscleblind-like
miRNA: micro-RNA
mRNA: messenger RNA
mg: milligram

ml:millilitre

mM: millimolar

mm: millimetre

ng: nanogram

nM: nanomolar

MQ: MilliQ™ purified water

NAT: natural antisense transcript

Nc: Nedd-like caspase

ORF: open reading frame

pax: paxillin

PKR: RNA regulated protein kinase

pll: pelle

pmol: picomole

PP2R2B: PP2A regulatory subunit 2B

pre-miR: precursor miRNA

Proc R: Proctolin receptor

pros: prospero

pum: pumilio

qRT-PCR: quantitative real-time polymerase chain reaction

rcf: relative centrifugal force

rdgA: retinal degeneration A

rdx: roadkill

repo: reversed polarity

RISC: RNA-induced silencing complex

RNAi: RNA interference

Rp49: Ribosomal protein 49

rpr: reaper

RT-PCR: reverse transcription polymerase chain reaction

rut: rutabaga

SAP: shrimp alkaline phosphatise

SCA: Spinocerebellar ataxia

SDS: sodium dodecyl sulphate

SMBA: Spinal bulbar muscular atrophy

siRNA: small interference-RNA

spz: spaetzle

sqa: spaghetti squash activator

stau: staufen

T: thymine

TAE: tris-acetate EDTA

TBP: TATA-box binding protein

TDP43: TAR-DNA binding protein-43

TK2: thymidine kinase 2

Tl: Toll receptor

TLR: Toll-like receptor

TNR: trinucleotide repeat

UAS: upstream activation sequence

UTR: untranslated region

VGLUT1: vesicular glutamate transporter-1

w/v: weight/volume

Nomenclature

Throughout this thesis, *Drosophila* nomenclature is denoted by conventional notation as found in the *Drosophila* database, FlyBase (www.flybase.org). Genes are represented in italicised text (e.g. '*htt*') and proteins are represented in non-italicised text (e.g. 'htt').

Abstract

The expansion of tandem repeat sequences beyond a pathogenic threshold is responsible for a series of neurodegenerative diseases known as dominantly inherited expanded repeat diseases. A number of these diseases are caused by the expansion of a CAG repeat tract in the coding region of various genes and are termed polyglutamine diseases. In these cases the polyglutamine tract is thought to contribute to pathogenesis. Several other clinically indistinguishable diseases however, are caused by the expansion of various repeat sequences in untranslated regions of genes. As expanded repeat RNAs are present in each of these cases, these RNAs have been proposed as a common pathogenic agent. Increasing evidence now exists for bi-directional transcription across the expanded repeat sequence of disease genes, leading to toxicity. The products of bi-directional transcription are predicted to form complementary double-stranded RNA. This study uses a *Drosophila* model of bi-directional transcription to determine the physical properties of the RNA and the downstream pathways that could contribute to pathogenesis in these diseases.

Expression of complementary repeat sequences predicted to form double-stranded RNA was toxic in this model and caused age-dependent neurodegeneration. This toxicity was dependent on several components of the RNA biogenesis pathway, including Dicer-2. The abundance of rCAG 21mer RNA and an altered miRNA profile were identified as biomarkers of this pathway. Microarray analysis identified genes involved in redox regulation, immune activation, cellular signalling and neurotransmission as novel candidates of pathogenesis. Furthermore, activation and signalling via the Toll pathway was required for pathogenesis indicating that an elevated immune response contributes to toxicity. Glial expression of double-stranded RNA caused severe neurodegeneration suggestive of non-autonomous toxicity as the cause of neuronal dysfunction. The identification of pathogenic pathways and molecular biomarkers are a critical step in developing therapeutic interventions for these diseases.