THE ROLE OF UPF3B AND THE NONSENSE-MEDIATED mRNA DECAY PATHWAY IN PATHOLOGY OF INTELLECTUAL DISABILITY

Lam Son Nguyen

Supervisors: Prof. Jozef Gecz & Dr. Mark Corbett

The Neurogenetics Laboratory,

Discipline of Pediatrics, School of Pediatrics & Reproductive Health,

Faculty of Health Sciences, The University of Adelaide, S.A., Australia

2013

Table of Contents

ABSTRACT

Nonsense mediated mRNA decay (NMD) functions to degrade transcripts containing premature termination codons and to regulate normal expression of the transcriptome. We identified mutations in *UPF3B*, a core member of NMD, as the cause of intellectual disability (ID) with or without other neuro-psychiatric traits and congenital anomalies. Recently, Thrombocytopenia with Absent Radius (TAR) syndrome was linked to compound heterozygous mutations in *RBM8A*, another NMD factor. About 7% of TAR patients also display ID, suggesting a common etiology underlying abnormal brain development in patients with compromised NMD.

To gauge into the role of NMD in the brain, we assessed the transcriptome wide deregulation as a consequence of compromised NMD in lymphoblastoid cell lines (LCLs) of patients with *UPF3B* mutations using RNA-Seq and exon array. We showed that up to 5% of the transcriptome was impacted on, affecting multiple genes with important neuronal functions. Among these, we demonstrated that up regulation of *ARHGAP24* isoform 1, which encodes an actin cytoskeleton remodeling protein, severely disrupted the axonal growth and hindered the survival of primary hippocampal neurons. This suggested that deregulation of *ARHGAP24* isoform 1, among other important neuronal genes, contributed directly to the patients' neuronal phenotype.

We expanded our inquiry into the role of other NMD factors in the brain. We surveyed copy number variants (CNVs) encompassing 18 NMD genes in 57,365 patients and 20,474 controls and identified 11 *de novo* CNVs encompassing *UPF2*, another core NMD factor which encodes for a direct interacting partner of UPF3B. The transcriptome deregulation due to heterozygous loss of *UPF2* in these patients was 95% similar to those of patients with *UPF3B* mutations, suggesting that *UPF2* is a novel, dosage

sensitive neuro-developmental gene. Additionally, CNVs of other four NMD genes *UPF3A*, *SMG6*, *EIF4A3* and *RNPS1* were also significantly enriched in the patients, and likely contributed to neuro-developmental disease etiology.

Overall, our work emphasizes the importance of properly functioning NMD in normal brain development. It also lays a solid foundation for future investigations into the causative and predisposing (when mutated) or modifying (when dosage imbalanced or polymorphic) role of NMD genes in a broad spectrum of neurological disorders.

DECLARATION

I certify that 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. In addition, I certify that no part of this work will, in the future, be used in a submission for any other degree or diploma in any university or other tertiary institution without the prior approval of the University of Adelaide and where applicable, any partner institution responsible for the joint-award of this degree.

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.

The author acknowledges that copyright of published works contained within this thesis resides with the copyright holder(s) of those works.

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.

Lam Son Nguyen

June 2013

ACKNOWLEDGEMENTS

I feel very fortunate to have been able to work, learn and participate in the exciting researches at the Neurogenetics Laboratory. I would like to thank my supervisors, Prof. Jozef Gecz and Dr. Mark Corbett. You have always believed in me and have always been there every step of my journey.

I would like to thank old and new members of the Neurogenetics Laboratory for useful discussions at meetings and for being such a wonderful bunch of people. In particular, I would like to thank Dr. Lachlan Jolly and Dr. Cheryl Shoubridge for contributing to the work described in Chapter 3 of this thesis.

I am grateful to the clinicians and genetic counselors for coordinating the genetic materials and cell lines needed for this work. I am also deeply thankful to the patients and their families for giving consents and for donating their unique genetic materials which helped to form the hypotheses and conclusions of many aspects of this thesis.

I would like to thank The MS McLeod Foundation for supporting my PhD scholarship, The Adelaide Centre for Neuroscience Research and The Florey Medical Research Foundation for funding my travels, and The Australian National Health and Medical Research Council for funding the project.

I would like to thank my mom, my late father, my sister and my brother for moral supports and encouragement. I would like to thank members of The Park 10 Ninjas and The Tuff Guys - Adelaide for keeping me fit. Finally, I would like to thank my partner, Kevin, for keeping me sane all these years. I am lucky to have you in my life.

PUBLICATIONS

1. LAUMONNIER, F., **NGUYEN, L. S.**, JOLLY, L., RAYNAUD, M. & GECZ, J. UPF3B Gene and Nonsense-Mediated mRNA Decay in Autism Spectrum Disorders. In: PATEL, V., PREEDY, V., MARTIN, C. (Ed.) The Comprehensive Guide to Autism: Springer Reference (www.springerreference.com). Springer-Verlag Berlin Heidelberg, 2013. DOI: 10.1007/SpringerReference_331210 2013-02-11 05:55:57 UTC.

2. **NGUYEN, L. S.**, JOLLY, L., SHOUBRIDGE, C., CHAN, W. K., HUANG, L., LAUMONNIER, F., RAYNAUD, M., HACKETT, A., FIELD, M., RODRIGUEZ, J., SRIVASTAVA, A. K., LEE, Y., LONG, R., ADDINGTON, A. M., RAPOPORT, J. L., SUREN, S., HAHN, C. N., GAMBLE, J., WILKINSON, M. F., CORBETT, M. A. & GECZ, J. 2012. Transcriptome profiling of UPF3B/NMD-deficient lymphoblastoid cells from patients with various forms of intellectual disability. *Mol Psychiatry*, 17, 1103-1115.

3. **NGUYEN, L. S.**, KIM, H. G., ROSENFELD, J. A., SHEN, Y., GUSELLA, J. F., LACASSIE, Y., LAYMAN, L. C., SHAFFER, L. G. & GECZ, J. 2013. Contribution of copy number variants involving nonsense-mediated mRNA decay pathway genes to neuro-developmental disorders. *Hum Mol Genet.* Epub ahead of print.

4. **NGUYEN, L. S.**, WILKINSON, M. & GECZ, J. The extent and relevance of inter-individual variability in NMD efficiency. *Neurosci and Biobehavioral Rev* (invited review, in submission).

5. LYNCH, S. A., **NGUYEN, L. S.**, NG, L. Y., WALDRON, M., MCDONALD, D. & GECZ, J. 2012. Broadening the phenotype associated with mutations in UPF3B: Two further cases with renal dysplasia and variable developmental delay. *Eur J Med Genet*, 55, 476-9.

ABBREVIATIONS

STANDARD TERMS

%	Percentage
^{o}C	Degrees Celsius
bp	Base pair
cDNA	Complimentary DNA
DNA	Deoxyribonucleic acid
Kb	Kilobase
KDa	Kilodalton
L	Litre
M	Molar
ml	Millilitre
mM	Milimolar
mRNA	Messenger RNA
P	Statistical significant value
RNA	Ribonucleic acid
RPM	Round per minute
s	Second
s.d.	Standard deviation
s.r.	Sample range
µg	Microgram
µl	Microliter

MATERIALS & METHODS

Ab	Antibody
aCGH	Array-comparative genomic hybridisation
BAC	Bacteria artificial chromosome
CHX	Cycloheximide

DAPI	4, 6-diamidino-2-phenylindole
DMEM	Dulbecco's Modified Eagle's Medium
DMSO	Dimethyl Sulfoxide
PBS	Phosphate buffered saline
PBST	Phosphate buffered saline with Tween20
dNTP	Deoxyribonucleotide
EBV	Ebstein-Bar virus
EDTA	Ethylene diamine tetraacetic Acid
FCS	Fetal calf sera
FISH	Fluorescence *in situ* hybridization
HEK293T	Human embryonic kidney 293T cell line
Ig	Immunoglobulin
LB	Luria broth
LCL	Lymphoblastoid cell line
L-Glu	L-Glutamine
NHS	Normal horse sera
PAGE	Poly-acrylamide gel electrophoresis
PCR	Polymerase chain reaction
PMSF	Phenylmethylsulfonyl fluoride
PolyA+	Polyadenylated
RT	Room temperature
RT-PCR	Reverse transcription PCR
RT-qPCR	Reverse transcription quantitative PCR
SDS	Sodium dodecyl sulphate
TBS	Tris buffered saline
TBST	Tris buffered saline with Tween20
Tm	Melting temperature
WBC	White blood cell

NON STANDARD TERMS

+ve	Positive

-ve	Negative
ADHD	Attention deficit hyperactivity disorder
ANOVA	Analysis of variance
AP	Alternative polyadenylation
AS	Alternative splicing
ASD	Autism spectrum disorder
CE	Cassette exon
CNV	Copy number variant
DAVID	Database for Annotation, Visualization and Integrated Discovery
DECIPHER	Database of Chromosomal Imbalance and Phenotype in Humans using Ensembl Resources
DEG	Differently expressed gene
DD	Developmental delay
DGAP	Developmental Gene Anatomy Project
DGDP	Developmental Gene Discovery Project
DGV	Database of Genomic Variants
ELAND	Efficient alignment of nucleotide databases
EJC	Exon junction complex
eQTL	Expression quantitative trait locus
GEO	Gene Expression Omnibus
HG	Human genome
ID	Intellectual disability
II	Intron isoform
IR	Intron retention
ME	Mutually exclusive
MIM	Mendelian Inheritance in Man
MMES	Minimal match on either side of exon junctions
NIGMS	The National Institute of General Medical Sciences
NMD	Nonsense-mediated mRNA decay
PTC	Premature termination codon
RNA-Seq	RNA sequencing

RPKM	Reads Per Kilobase exon Model per million mapped reads
SI	Splicing index
SNP	Single nucleotide polymorphism
STD	Standard deviation
SURF	SMG1, UPF1 and ERF1-ERF3 complex
SZ	Schizophrenia
TAR	Thrombocytopenia with Absent Radius
TERRA	Telomeric repeat-containing RNA
UCSC	University of California, Santa Cruz
uORF	Upstream open reading frame
UTR	Untranslated region
WG	Whole genome

NMD GENES AND GENE FAMILIES

CASC3	Cancer susceptibility candidate 3
CBP	Cap binding protein
DHX34	DEAH (Asp-Glu-Ala-His) box polypeptide 34
EIF4A3	Eukaryotic translation initiation factor 4A3
MAGOH	Mago-nashi homolog, proliferation-associated
NBAS	Neuroblastoma amplified sequence
RBM8A	RNA binding motif protein 8A
RNPS1	RNA binding protein S1, serine-rich domain
SMG	Suppressor of morphological defects on genitalia 1
UPF	Up frameshift suppressor homolog
WIBG	Within bgcn homolog (Drosophila)

SPECIES

D. melanogaster	*Drosophila melanogaster*
D. rerio	*Danio rerio*
E. coli	*Escherichia coli*
H. sapiens	*Homo sapiens*
S. cerevisiae	*Saccharomyces cerevisiae*