The genetic determinants of cerebral palsy

A thesis submitted for the degree of Doctor of Philosophy (PhD) to the University of Adelaide

By
Gai McMichael

Supervisors: Professors Jozef Gecz and Eric Haan
The University of Adelaide, Robinson Institute
School of Medicine
Faculty of Health Science

May 2016
**Statement of 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.

Gai Lisette McMichael

January 2016
### Table of contents

**Statement of declaration** i

**Table of contents** ii

**Acknowledgements** ix

**Publications** xi

**HUGO Gene Nomenclature gene symbol and gene name** xiii

**Abbreviations** xvi

**URLs** xix

**Chapter 1 Introduction** 1

1.1 Definition of cerebral palsy 2

1.2 Clinical classification of cerebral palsy 3
   1.2.1 Gross motor function classification system 5

1.3 Neuroimaging 7

1.4 Incidence and economic cost of cerebral palsy 8

1.5 Known clinical risk factors for cerebral palsy 9
   1.5.1 Preterm birth 9
   1.5.2 Low birth weight 9
   1.5.3 Multiple birth 10
   1.5.4 Male gender 10

1.6 Other known clinical risk factors 11
   1.6.1 Birth asphyxia 11

1.7 Other possible risk factors 12

1.8 Evidence for a genetic contribution to cerebral palsy causation 13
   1.8.1 Sibling risks and twin studies 13
   1.8.2 Effect of consanguinity 14
   1.8.3 Associated congenital anomalies 15
   1.8.4 Families providing evidence for Mendelian inheritance 16

1.9 Genetic disorders that can be mistaken for cerebral palsy 17

1.10 Approaches to identification of genes involved in cerebral palsy 19
   1.10.1 Linkage analysis and homozygosity mapping 19
   1.10.2 GADI 20
   1.10.3 KANK1 20
   1.10.4 Chromosome 9p12-q12 21

1.11 Candidate-gene association studies 22

1.12 Copy number variants 23

1.13 Massively parallel sequencing 28

1.14 Whole-exome sequencing 28
Chapter 4  
Whole exome sequencing points to considerable genetic  
heterogeneity of cerebral palsy

4.1 Abstract
4.2 Introduction
4.3 Materials and methods
  4.3.1 Study cohort
  4.3.2 DNA extraction
  4.3.3 Illumina library construction
  4.3.4 Illumina exome capture
  4.3.5 Illumina sequencing
  4.3.6 Analysis
  4.3.7 DNA variant and gene prioritization
4.4 Results
  4.4.1 Case-parent trios (n=98)
    4.4.1.1 De novo mutations
    4.4.1.2 Inherited X-chromosome and recessive variants
  4.4.2 Whole-exome sequencing
  4.4.3 Pathway analysis
  4.4.4 Clinical associations
4.5 Discussion
4.6 References

Chapter 5  
A family diagnosed with cerebral palsy and intellectual disability
with a X-linked mutation in ZC4H2

5.1 Introduction
5.2 Clinical reports
  5.2.1 The family
  5.2.2 The parents (I-1 and I-2)
  5.2.3 The son with ‘cerebral palsy’ and intellectual disability
    (II-1)
  5.2.4 The son with ‘cerebral palsy’ and intellectual disability
    (II-5)
  5.2.5 The daughter with developmental delay (II-7)
  5.2.6 Remaining siblings (II2, II3, II-4, II-6 and II-8)
5.3 Methods
  5.3.1 DNA isolation
  5.3.2 Whole-exome sequencing and analysis
  5.3.3 Sanger sequencing
8.3 Familial cases

8.3.1 Family 1 with ZC4H2 mutation

8.3.2 Family 2 with NKX2-1 mutation

8.3.3 Family 3 with SCN2A mutation

8.4 Limitations of this study

8.5 Future prospects

8.5.1 Whole-exome and whole-genome sequencing

8.5.2 Custom gene-panels

8.5.3 Central data repository

8.5.4 Genome-wide expression profiling (RNA sequencing)

8.5.5 Epigenetics and cerebral palsy

8.6 Conclusion

8.7 References

Appendix Supplementary data

9.1 References

List of Tables

Table 1.1 Summary of potentially pathogenic CNVs encompassing genes previously associated with cerebral palsy

Table 1.2 Summary of the different solution-based commercial exome capture kits

Table 1.3 Summary of novel single gene mutations, in known (OMIM) genes, in cerebral palsy cases

Table 2.1 Type of cerebral palsy in 183 sporadic cases

Table 2.2 Gestational age distribution of 183 sporadic cases with cerebral palsy compared with South Australian (SA) pregnancy outcome data

Table 2.3 Birth weight distribution in 183 sporadic cases with cerebral palsy compared with South Australian (SA) pregnancy outcome data

Table 2.4 Distribution of co-morbidities in 183 sporadic cases with cerebral palsy

Table 2.5 Gender distributions in 183 sporadic cases with cerebral palsy

Table 2.6 Summary of the clinical characteristics for three families With more than one individual with a diagnosis of cerebral Palsy

Table 3.1 Primer sequences for validating candidate variants from whole blood (The University of Adelaide)

Table 4.1 Clinical characteristics of 183 individuals with cerebral palsy

Table 4.2 Novel variants in known (OMIM) genes in CP cases identified in 98 trios: CADD score $>$ 10 and RVIS % $<$ 50
Table 4.3 Novel candidate genes in CP cases identified in 98 trios; CADD score > 20 and RVIS % < 50 124
Table 5.1 Multispecies alignment around p.Arg213 of ZC4H2 141
Table 7.1 Clinical characteristics, SCN2A variants and segregation pattern and pathogenicity predictions and scores 178
Table 7.2 Fragments of sequence chromatograms from 5’-3’ 178
Table 9.1 Supplementary Table A 224
Table 9.2 Supplementary Table B 242
Table 9.3 Supplementary Table C 243
Table 9.4 Supplementary Table D 247
Table 9.5 Supplementary Table E 248
Table 9.6 Supplementary Table F 249
Table 9.7 Supplementary Table G 250
Table 9.8 Supplementary Table H 253
Table 9.9 Supplementary Table I 254
Table 9.10 Supplementary Table J 255
Table 9.11 Supplementary Table K 258
Table 9.12 Supplementary Table L 259

List of Figures
Figure 1.1 Different types of cerebral palsy and the regions of the body affected 5
Figure 1.2 Prevalence of cerebral palsy rates per 1000 live births compared with caesarean rates over the past 50 years 12
Figure 1.3 The prevalence and type of co-morbidities seen in individuals diagnosed with cerebral palsy 16
Figure 1.4 Allele frequency correlated with effect size 23
Figure 2.1 Type of cerebral palsy in 183 sporadic cases 71
Figure 2.2 Gestational age distribution of 183 sporadic cases with cerebral palsy compared with South Australian (SA) pregnancy outcome data 73
Figure 2.3 Birth weight distribution in 183 sporadic cases with cerebral palsy compared with South Australian (SA) pregnancy outcome data 75
Figure 2.4 Distribution of co-morbidities in 183 sporadic cases with cerebral palsy 76
Figure 2.5 Gender distributions in 183 sporadic cases with cerebral palsy 77
Figure 2.6-2.8 Overview of the pedigree details for three families with a diagnosis of cerebral palsy 78
Figure 5.1 Two generation pedigree with two brothers (II-1 and II-5) diagnosed with ‘cerebral palsy’ and intellectual disability and a sister (II-7) significantly behind in language, cognitive and personal/social skills 132
Figure 5.2  Representative sequence chromatograms of ZC4H2 wildtype, mutations and heterozygous variants identified in nine out of the ten family members

Figure 6.1  Three generation pedigree with de novo mutation in II-1 and subsequent autosomal dominant transmission in III-2 and III-3

Figure 6.2(A)  Fragments of sequence chromatograms from an affected individual (heterozygous mutation) and an unaffected individual (normal homozygous) from 5’–3’ and corresponding amino acid sequences

Figure 6.2(B)  Comparison of amino acid sequences of wildtype and mutant NKX2-1 proteins

Figure 7.1  Two-generation pedigree

Figure 7.2  Fragments of sequence chromatograms from 5’-3’

Figure 9.1  Supplementary Figure A

Figure 9.2  Supplementary Figure B

Figure 9.3  Supplementary Figure C

Figure 9.4  Supplementary Figure D

Figure 9.5  Supplementary Figure E

Figure 9.6  Supplementary Figure F

Figure 9.7  Supplementary Figure G

Figure 9.8  Supplementary Figure H

Figure 9.9  Supplementary Figure I

Figure 9.10  Supplementary Figure J
Acknowledgments

I extend my thanks to my supervisors, Professor Jozef Gecz and Professor Eric Haan for their support, wisdom and guidance over the duration of this study. Their mentoring and encouragement to me personally and in my career thus far has provided me with many valuable and exciting opportunities for which I am extremely thankful. It has been an absolute pleasure and privilege to work alongside Professor Jozef Gecz and Professor Eric Haan during both my PhD and my career in cerebral palsy research.

I gratefully extend my thanks to Professor Alastair MacLennan for his support for my career and studies in cerebral palsy research, in particular for supervising me during the first two years of my PhD. I very much appreciate the support from the Cerebral Palsy Research Group, University of Adelaide, Robinson Institute. I gratefully extend my sincere thanks to Corinne Reynolds for her support and friendship throughout my PhD. I extend sincere thanks to Professor Richard Gibbs for providing the opportunity to collaborate and work at Baylor College of Medicine, The Human Genome Sequencing Center (HGSC), Houston, Texas, USA.

I am very appreciative and am thankful for the support from the Department of Paediatric Rehabilitation Medicine at Women’s and Children’s Hospital, Adelaide in particular Dr Ray Russo, Dr James Rice and Dr Andrew Tidemann for allowing us to collect blood samples whilst our participants were under general anaesthetic. I also wish to thank Dr Jane Valentine and Peta Watts for facilitating blood collections at Princess Margaret Hospital, Western Australia.
Thank you to the staff at the University of Adelaide Department of Obstetrics and Gynaecology, both at the Women’s and Children’s Hospital and the Medical School for supporting me through this candidature. Thank you to Dr Kathie Friend and staff at the Department of Diagnostic Molecular Genetics, SA Pathology, Adelaide and the Australian Genome Research Facility (Adelaide node) for support with DNA extractions.

I sincerely thank my colleagues at the Adelaide Neurogenetics Research Group, The University of Adelaide, Robinson Institute for their support, advice, mentoring and above all their valued friendship. In particular thank you to Dr Raman Sharma, Dr Duyen Pham, Stanley Tan, Dr Mark Corbett and Alison Gardner it was a pleasure to work with you.

This work has been funded in part by the Cerebral Palsy Foundation, Tennix, University of Adelaide, Women’s and Children’s Research Foundation and National Health and Medical Research Council.
Publications


HUGO Gene Nomenclature gene symbol and gene name

ADD3 – Adducin 3 (Gamma)

AGAP1 – ArfGAP with GTPase Domain, Ankyrin Repeat and PH Domain 1

AHI1 – Abelson Helper Integration Site 1

ANKRD15 – KN Motif And Ankyrin Repeat Domains 1 (KANK1)

AP-4 – AP-4 Complex

AP4B1 – Adaptor-related Protein Complex 4, β1 subunit

AP4E1 – Adaptor-related Protein Complex 4, ε1 subunit

AP4M1 – Adaptor-related Protein Complex 4, μ1 subunit

AP4S1 – Adaptor-related Protein Complex 4, σ1 subunit

CDK17 – Cyclin-Dependent Kinase 17

CD99L2 – CD99 Molecule-Like 2

CEP290 – Centrosomal Protein 290KDa

COPS3 – COP9 Signalosome Subunit 3

CTNND2 – Catenin (Cadherin-Associated Protein), Delta 2

CUL4B – Cullin 4B

ENPP4 – Ectonucleotide Pyrophosphatase/Phosphodiesterase 4 (Putative)

FLNB – Filamin B (Beta)

GAD1 – Glutamate Decarboxylase 1

GCH1 – GTP Cyclohydrolase 1

HSPA4 – Heat Shock 70kDa Protein 4

HSPG2 – Heparan Sulfate Proteoglycan 2

ITPR1 – Inositol 1,4,5-Trisphosphate Receptor, Type 1

JHDM1D – Lysine (K)-Specific Demethylase 7A

KANK1 – KN Motif And Ankyrin Repeat Domains 1 (ANKRD15)
HUGO Gene Nomenclature gene symbol and gene name (continued)

**KDM5C** – Lysine (K)-Specific Demethylase 5C

**KCNC3** – Voltage-Gated Potassium Channel Subunit Kv3.3

**LICAM** – L1 Cell Adhesion Molecule

**LTN1** – Listerin E3 Ubiquitin Protein Ligase 1

**MAOA** – Monoamine Oxidase A

**MAOB** – Monoamine Oxidase B

**MAST1** – Microtubule Associated Serine/Threonine Kinase 1

**MCPH1** – Microcephalin 1

**MED17** – Mediator Complex Subunit 17

**MEF2C** – Myocyte Enhancer Factor 2C

**MIIP** – Migration and Invasion Inhibitory Protein

**MUM1L1** – Melanoma Associated Antigen (Mutated) 1-Like 1

**MYH14** – Myosin, Heavy Chain 14, Non-Muscle

**NAA35** – N(Alpha)-Acetyltransferase 35, NatC Auxiliary Subunit

**NEMF** – Nuclear Export Mediator Factor

**NPHP1** – Nephronophthisis 1 (Juvenile)

**NKX2-1** – NK2 Homeobox 1

**PAK3** – P21 Protein (Cdc42/Rac)-Activated Kinase 3

**PACRG** – PARK2 Co-Regulated

**PARK2** – Parkin RBR E3 Ubiquitin Protein Ligase

**PAX5** – Paired Box 5

**PLAC4** – Placenta-Specific 4

**PLP1** – Proteolipid Protein 1

**PCDH11X** – Protocadherin 11 X-Linked
HUGO Gene Nomenclature gene symbol and gene name (continued)

*RFX2* – Factor X, 2 (Influences HLA Class II Expression)

*SCN2A* – Sodium Channel, Voltage Gated, Type II Alpha Subunit

*SCN8A* – Sodium Channel, Voltage Gated, Type VIII Alpha Subunit

*SLC11A2* – Solute Carrier Family 11 (Proton-Coupled Divalent Metal Ion Transporter), Member 2

*SPAST* – Spastin

*SPR* – Sepiapterin Reductase (7,8-Dihydrobiopterin:NADP+ Oxidoreductase)

*SPTBN2* – Spectrin, beta, non-erythrocytic 2

*SSPO* – SCO-Spondin

*SYNGAP1* – Synaptic Ras GTPase Activating Protein 1

*TBC1D24* – TBC1 Domain Family, Member 24

*TENM1* – Teneurin Transmembrane Protein 1

*TGM6* – Transglutaminase 6

*TUBA1A* – Tubulin, Alpha 1a

*UBE3A* – Ubiquitin Protein Ligase E3A

*UBQLN3* – Ubiquilin 3

*WDR45* – WD Repeat Domain 45

*WIPI2* – WD Repeat Domain, Phosphoinositide Interacting 2

*ZC4H2* – Zinc Finger, C4H2 Domain Containing

*ZNF674* – Zinc Finger Protein 674
Abbreviations

ACD – Acid citrate dextrose
AMC – Arthrogryposis multiplex congenita
ASD – Autism spectrum disorder
ATLAS – SNP
BCM – Baylor College of Medicine
BHC – Benign Hereditary Chorea
BWA – Burrows-Wheeler Aligner
CADD – Combined Annotation-Dependent Depletion
CGH – Comparative Genomic Hybridisation
ChIP – Chromatin immunoprecipitation
CNVs – Copy Number Variants
CNS – Central nervous system
CP – Cerebral palsy
DNA – Deoxyribonucleic acid
EDTA – Ethylenediaminetetraacetic acid
EVS – Exome Variant Server
ExAC – Exome Aggregation Consortium
GA – General anaesthetic
GABA – Gamma-aminobutyric acid
GATK – Genome Analysis Toolkit
GERP – Genomic Evolutionary Rate Profiling
GMFCS – Gross Motor Function Classification System
GRA – Genetic Repositories Australia
HGMD – Human Genome Mutation Database
Abbreviations (continued)

HGSC – Human Genome Sequencing Center
HSP – Hereditary Spastic Paraplegia
ID – Intellectual disability
IQ – Intelligence quotient
IUGR – Intrauterine growth restriction
IVF – *In vitro* fertilization
IVH – Intraventricular haemorrhage
LCLs – Lymphoblastoid cell lines
LM-PCR – Ligation mediated-polymerase chain reaction
LOF – Loss of function
MAF – Minor allele frequency
MIM – Mendelian Inheritance in Man
MPS – Massively parallel sequencing
mRNA – Messenger ribonucleic acid
miRNA – Micro ribonucleic acid
MRI – Magnetic resonance imaging
ncRNA – Non-coding ribonucleic acid
NGS – Next generation sequencing
NHLBI – National Heart, Lung, and Blood Institute
OMIM – Online Mendelian Inheritance in Man
OR – Odds ratio
PCR – Polymerase chain reaction
PE – Paired-end
PMD – Pelizaeus-Merzbacher disease
**Abbreviations (continued)**

POSU – Pregnancy Outcome in South Australia

PVL – Periventricular leukomalacia

RNA – Ribonucleic acid

rRNA – Ribosomal ribonucleic acid

tRNA – Transfer ribonucleic acid

RVIS – Residual Variation Intolerance Score

SHRs – Standardized hospitalisation ratios

SIFT – Scale-invariant feature transform

SNPs – Single nucleotide polymorphisms

SPG4 – Spastic paraplegia 4

UCSC – University of California, Santa Cruz Genome Browser

UTRs – Untranslated regions

WES – Whole-exome sequencing

WGS – Whole-genome sequencing

XLID – X-linked intellectual disability
URLs

Ensembl – www.ensembl.org/
EVS – http://evs.gs.washington.edu/EVS/
ExAC database – exac.broadinstitute.org/
OMIM – www.omim.org/
Partek – http://www.partek.com/
1000 Genomes – http://browser.1000genomes.org/index.html
UCSC – https://genome.ucsc.edu/