Dissecting the Genetic Basis of Parkinson Disease, Dystonia and Chorea

A thesis submitted to the University College London for the degree of Doctor of Philosophy

June 2016

by

Dr Niccolò Emanuele Mencacci
Declaration of Authorship

I, Niccolò Emanuele Mencacci, confirm that the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated in the thesis.
Incontenibile andare, di monte in monte, inquieti
dietro un mistero che sempre ti seduce, da un’altra
valle
Abstract

In this thesis I used a range of genetic methodologies and strategies to unravel the genetic bases of Parkinson disease (PD), myoclonus-dystonia (M-D), and chorea.

First, I detail the work I performed in PD, including (1) the screening of GBA in a cohort of early-onset PD cases, which led to the identification of the allele E326K (p.Glu365Lys) as the single most frequent, clinically relevant, risk variant for PD; (2) a detailed genetic analysis in a large cohort of PD cases who underwent deep-brain stimulation treatment and a longitudinal comparison of the phenotypic features of carriers of mutations in different genes; (3) the observation that rare GCH1 coding variants, known to be responsible for the childhood-onset disorder DOPA-responsive dystonia, are a novel risk factor for PD.

Then, I describe the work I performed to identify novel causes of M-D, including (1) the discovery of the missense p.Arg145His mutation in KCTD17 as a novel cause of autosomal dominant M-D; (2) the identification of tyrosine hydroxylase deficiency as a novel treatable cause of recessive M-D; and (3) the conclusive disproof of the pathogenic role of the p.Arg1389His variant in CACNA1B as a cause of M-D.

Finally, I detail my work in the field of choreic syndromes, including (1) the genetic screening of NKX2-1 in the Queen Square cohort of benign hereditary chorea (BHC) cases; (2) the identification of ADCY5 mutations, the gene thought to be responsible for the condition familial dyskinesias with facial myokymia, as an important cause of BHC; and (3) the identification of de novo mutations in PDE10A as a novel genetic cause of chorea.

These findings are discussed in light of the recent literature. Following my analysis, I suggest future directions for the identification of novel genetic causes of movement disorders, in light of my recent findings and ongoing research.
# Table of Contents

Declaration of Authorship ................................................................. 2  
Abstract ................................................................................................. 3  
Table of Contents .................................................................................. 5  
List of Figures ......................................................................................... 9  
List of Tables .......................................................................................... 12  
Peer-Reviewed Publications ................................................................. 14  
Acknowledgements ................................................................................ 18  

## Chapter 1. General Introduction ......................................................... 20  
1.1 Thesis Aim ...................................................................................... 20  
1.2 Summary of Relevant Techniques in Genetics ................................. 21  
  1.2.1 Dideoxynucleotide Sequencing ............................................... 21  
  1.2.2 Next Generation Sequencing .................................................. 22  
  1.2.3 Linkage Analysis ...................................................................... 24  
1.3 Models and Strategies to Study the Genetic Architecture of Diseases .... 26  
1.4 Strategies for Gene Identification in Mendelian Diseases ................. 28  
1.5 Identifying Causal Alleles: How to Prioritise Variants? ..................... 31  
  1.5.1 Filtering Based on the Inheritance Mode and Pedigree Information .... 32  
  1.5.2 Filtering Based on Rarity .......................................................... 33  
  1.5.3 Filtering Based on Deleteriousness and Conservation .................. 34  
1.6 Limitations of Next-Generation Sequencing .................................... 35  

## Chapter 2. Overview on Genetics of Selected Movement Disorders ........ 37  
2.1 The Genetics of Parkinson disease ............................................... 37  
  2.1.1 Introduction .............................................................................. 37  
  2.1.2 Monogenic Causes of Parkinson Disease .................................... 40  
  2.1.3 Rare Variants Predisposing to PD ............................................ 50  
  2.1.4 Common Low-Risk Variants Predisposing to PD ....................... 56  
2.2 Overview on the Genetics of Dystonia ............................................. 59  
  2.2.1 Introduction .............................................................................. 59  
  2.2.2 Overview of Monogenic Forms of Dystonia .............................. 63  
2.3 Genetics of Choreas ........................................................................ 77
2.3.1 Introduction ................................................................. 77
2.3.2 Huntington’s disease ................................................... 81
2.3.3 Huntington’s disease-like syndromes ............................. 82
2.3.4 Benign Hereditary Chorea ............................................. 87

Chapter 3. Materials and Methods ........................................... 88
3.1 DNA Extraction ................................................................ 88
  3.1.1 Extraction from Whole Blood ...................................... 88
  3.1.2 Extraction from Saliva ............................................... 88
3.2 Polymerase-Chain Reaction .............................................. 89
3.3 Agarose Gel Electrophoresis ............................................. 89
3.4 PCR Purification ............................................................. 90
3.5 Sanger Sequencing ........................................................... 90
3.6 Multiplex Ligation-Dependent Probe Amplification ............. 90
3.7 Analysis of Microsatellites ............................................... 92
3.8 DNA Array SNP Analysis ................................................ 92
3.9 Genome-Wide Parametric Multipoint Linkage Analysis ....... 93
3.10 Whole-Exome Sequencing .............................................. 94
3.11 Bioinformatic Analysis of Whole-Exome Sequencing .......... 96
3.12 Regional Gene Expression Profiling in Brain Tissue .......... 98

Chapter 4. Exploring the Genetics of Parkinson Disease ............ 100
4.1 Outline Of The Chapter ................................................... 100
4.2 GBA Mutational Analysis In Early Onset Parkinson Disease .... 100
  4.2.1 Statement Of Contribution ........................................ 100
  4.2.2 Background ............................................................ 100
  4.2.3 Subjects, Materials and Methods ............................... 101
  4.2.4 Results ................................................................. 103
  4.2.5 Discussion .............................................................. 106
4.3 Genotype-Phenotype Correlation in A Cohort Of Parkinson Disease Patients Treated With Deep-Brain Stimulation ....... 109
  4.3.1 Statement of Contribution ......................................... 109
  4.3.2 Background ............................................................ 109
  4.3.3 Subjects, Materials and Methods ............................... 110
  4.3.4 Results ................................................................. 112
  4.3.5 Discussion .............................................................. 118
4.4 Mutational Analysis of the GTP cyclohydrolase-1 Gene in Parkinson Disease .................................................................................................................. 121
  4.4.1 Statement of Contribution .................................................................................. 121
  4.4.2 Background ....................................................................................................... 121
  4.4.3 Subjects, Materials and Methods ....................................................................... 122
  4.4.4 Results .............................................................................................................. 125
  4.4.5 Discussion ......................................................................................................... 144

Chapter 5. Exploring the Genetics Basis of Myoclonus-Dystonia ..................... 150
  5.1 Outline of the Chapter .......................................................................................... 150
  5.2 Overview of Myoclonus-Dystonia ........................................................................ 150
  5.3 Linkage Analysis and Exome Sequencing Identify A Novel Cause Of
    Familial Myoclonus-Dystonia .................................................................................. 153
      5.3.1 Statement of contribution .............................................................................. 153
      5.3.2 Introduction ................................................................................................... 153
      5.3.3 Subjects, Materials and Methods .................................................................. 153
      5.3.4 Results ......................................................................................................... 157
      5.3.5 Discussion ..................................................................................................... 176
  5.4 Tyrosine Hydroxylase Deficiency Causes Autosomal Recessive, DOPA-
    Responsive, Myoclonus-Dystonia ......................................................................... 182
      5.4.1 Statement of Contribution .............................................................................. 182
      5.4.2 Background ................................................................................................... 182
      5.4.3 Subjects, Materials and Methods .................................................................. 182
      5.4.4 Results ......................................................................................................... 183
      5.4.5 Discussion ..................................................................................................... 187
  5.5 Assessment of The Role Of The CACNA1B Arg1389His Variant In a large
    Myoclonus-Dystonia Cohort .................................................................................. 190
      5.5.1 Statement of Contribution .............................................................................. 190
      5.5.2 Background ................................................................................................... 190
      5.5.3 Subjects, Materials and methods ................................................................... 190
      5.5.4 Results ......................................................................................................... 191
      5.5.5 Discussion ..................................................................................................... 193

Chapter 6. Exploring the Genetics of Benign Hereditary Chorea ...................... 195
  6.1 Outline of chapter ............................................................................................... 195
  6.2 Benign Hereditary Chorea .................................................................................... 195
6.3  *N*<sup>1</sup>KX2-1 Mutational Analysis in the Queen Square Benign Hereditary Chorea Cohort
6.3.1 Statement of Contribution
6.3.2 Background
6.3.3 Subjects, Materials and Methods
6.3.4 Results
6.3.5 Discussion

6.4  *ADCY5* Mutational Analysis in a Cohort of Benign Hereditary Chorea Cases
6.4.1 Statement of Contribution
6.4.2 Introduction
6.4.3 Subjects, Materials and Methods
6.4.4 Results
6.4.5 Discussion

6.5  De novo Mutations in *PDE10A* Cause Childhood-Onset Chorea with Bilateral Striatal Lesions
6.5.1 Statement of Contribution
6.5.2 Background
6.5.3 Subjects, Materials and Methods
6.5.4 Results
6.5.5 Discussion

Chapter 7.  Conclusions
7.1 Genetics of Parkinson Disease
7.2 Genetics of Myoclonus-Dystonia
7.3 Genetics of Chorea

References
**List of Figures**

Figure 1-1 Schematic representation of the main technical steps of NGS..........................23

Figure 1-2 Scheme illustrating the necessary methods for gene discovery depending on the hypothesised underlying architecture of disease.........................................................26

Figure 1-3 NGS strategies for gene identification ..................................................................28

Figure 1-4 Prioritization of NGS candidate variants..........................................................31

Figure 2-1 Typical appearance of a patient with Parkinson disease. ..............................37

Figure 2-2 Pathological features of Parkinson disease.......................................................38

Figure 2-3. Schematic representation of the frequency and penetrance of the Parkinson disease genes ..................................................................................................................40

Figure 2-4 GBA gene structure and distribution of pathogenic mutations across the coding exons ..................................................................................................................51

Figure 2-5 Map showing the countries where the frequency of GBA mutations in Parkinson disease patients was assessed..............................................................54

Figure 2-6 Manhattan plot from the latest GWAS meta-analysis for Parkinson disease 58

Figure 2-7 Schematic comparison of the old and the new classification system of dystonia .........................................................................................................................61

Figure 2-8 The circuits of dystonia ......................................................................................62

Figure 2-9 Schematic illustration of the cellular pathway involved in dopamine synthesis in nigral neurons .................................................................................................72

Figure 3-1 Steps of the MLPA reaction..............................................................................91

Figure 3-2 Scheme of the library preparation using the Truseq Illumina kit ..................95

Figure 3-3 Cluster generation, bridge amplification and imaging .....................................96

Figure 3-4. Bioinformatic pipeline for WES analysis.......................................................97
Figure 4-1 Pedigrees and [123I]FP-CIT SPECT scan images of the four families with GCH1 mutations involved in this study......................................................... 126

Figure 4-2 RNA analysis performed in family-A ......................................................... 128

Figure 4-3 Phenylalanine-loading test performed in family B........................................ 130

Figure 4-4 Visual representation of coverage depth for GCH1 exon 1............................. 137

Figure 5-1 Pedigree of the index family where the KCTD17 c.434 G>A p.Arg145His mutation was identified.......................................................... 159

Figure 5-2 Linkage analysis results .............................................................................. 162

Figure 5-3 Sanger sequencing confirmation of the KCTD17 c.434G>A p.Arg145His mutation. ........................................................................................................... 165

Figure 5-4 Interspecies protein sequence alignment....................................................... 165

Figure 5-5 Pedigree of the German families with the KCTD17 c.434 G>A p.Arg145His mutation. ............................................................................................................. 167

Figure 5-6 KCTD17 brain regional mRNA expression data ........................................... 171

Figure 5-7 Immunocytochemistry in stably transfected SH-SY5 cells......................... 172

Figure 5-8 Network representation of the putamen KCTD17-containing gene module .......................................................... 173

Figure 5-9 Putamen KCTD17-containing module preservation across other brain regions .................................................................................................................. 174

Figure 5-10 Functional studies showing abnormalities of endoplasmic reticulum calcium signalling in KCTD17 p.Arg145His substitution bearing fibroblasts..... 176

Figure 5-11 Pedigree of a family with recessive L-DOPA responsive Myoclonus-Dystonia .................................................................................................................... 183

Figure 5-12 Genetic results .......................................................................................... 186

Figure 5-13 Chromatograms showing the CACNA1B c.4166G>A; p.Arg1389His variant
(A) and a control sequence (B). ................................................................. 192

Figure 6-1 Schematic representation of NKX2-1 point mutations reported to date. ..... 197

Figure 6-2 Clinical features described in cases with NKX2-1 mutations................. 198

Figure 6-3 Brain, thyroid and lung involvement in index patients and family members, with and without NKX2-1 mutations, in the largest cohort of Benign Hereditary Chorea cases screened to date........................................................................ 199

Figure 6-4. Array data of patient with BHC showing no deletions either encompassing or neighbouring the NKX2-1 locus................................................................. 204

Figure 6-5 Pedigree and genetic results of family 1 ................................................. 216

Figure 6-6. Pedigree of family 5 with the novel Arg418Gly mutation ...................... 218

Figure 6-7 Graphical Summary of ADCY5 Brain Expression Data.......................... 222

Figure 6-8 Graphical Summary of NKX2-1 Brain Expression Data........................ 223

Figure 6-9 Summary of human and mouse brain PDE10A mRNA expression data .... 235

Figure 6-10 Summary of human brain expression of PRMT7 and THDS1 ............ 236

Figure 6-11 Pedigree and haplotype analysis in the families of cases 1 and 2 .......... 237

Figure 6-12 Interspecies alignment of the identified pathogenic PDE10A mutations.. 238

Figure 6-13 MRI features associated with de novo dominant PDE10A mutations .... 241

Figure 6-14 Schematic representation of the PDE10A protein and in silico modelling of the 3D structure of the PDE10A GAF-B domain......................................................... 244

Figure 6-15 Enzyme Kinetics of Wild-Type and Mutant PDE10As....................... 246

Figure 6-16 Effects of PDE10A substitutions on the modulatory activity of the GAF-B domain............................................................................................................. 247
List of Tables

Table 2-1 Monogenic causes of Parkinson disease ..................................................41
Table 2-2 Traditional clinical classification of dystonia ...............................................60
Table 2-3 Monogenic causes of dystonia .................................................................65
Table 2-4 Monogenic causes of chorea .................................................................79
Table 4-1 Proportion of cases and controls with GBA mutations ..............................105
Table 4-2 Description of the genetic mutations in PARK2, GBA and LRRK2 in a cohort of PD patients with DBS surgery .................................................................114
Table 4-3 Pre-operative UPDRS scores and response to L-dopa according to genetic subgroup ............................................................................................................117
Table 4-4 DAT binding values of GCH1 mutation carriers evaluated in this study.....132
Table 4-5 Characteristics of PD cases with GCH1 pathogenic variants and abnormal dopaminergic imaging described in this study and present in the literature .........133
Table 4-6 Average exome sequencing coverage of the six GCH1 exons ...............136
Table 4-7 List of GCH1 variants identified by exome sequencing in patients with PD and controlsa ..............................................................................................................140
Table 4-8 Clinical features of PD cases with GCH1 variants identified in the exome-sequencing study ........................................................................................................142
Table 5-1 Clinical features of cases carrying the KCTD17 c.434 G>A p.Arg145His mutation ........................................................................................................................160
Table 5-2 Summary of position, size and genetic content of the regions with a LOD score > -2 ................................................................................................................161
Table 5-3 Summary of results from exome sequencing .............................................163
Table 5-4 Summary of novel variants detected by whole-exome sequencing and shared by individuals V-3 and IV-14 ................................................................................166
Table 5-5 Disease haplotype of the families with the *KCTD17* c.434 G>A p.Arg145His

Table 6-1 Clinical features of cases with pathogenic *ADCY5* mutations

Table 6-2 Summary of variants detected by whole-exome sequencing that survived the filtering strategy.

Table 6-3 Clinical and genetic features of cases with *PDE10A* mutations
Peer-Reviewed Publications

Publications Relating to the Work Contained Within This Thesis


Stamelou M, Mencacci NE, (joint first authors), Cordivari C, Batla A, Wood NW, Houlden H, Hardy J, Bhatia KP. Myoclonus-dystonia syndrome due to tyrosine

Other Significant Contribution During Thesis Period

Mencacci NE, Carecchio M. Recent advances in the genetics of chorea. Current Opinion in Neurology. In press.


Other Peer-Reviewed Publications During Thesis Period


McNeill A, Wu RM, Tzen KY, Aguiar PC, Arbelo JM, Barone P, Bhatia K, Barsottini O,


Acknowledgements

This thesis is dedicated to the loving memory of my godmother Franca. Her words and teaching shaped my mind and soul and will always guide me throughout my life.

I am deeply grateful to the many persons that made the years of my PhD one the most enjoyable and rewarding experiences of my life.

First, I would like to express my most sincere gratitude to my primary supervisor Professor Nicholas Wood for his support, mentorship, advice and encouragement throughout the course of my work at the Institute of Neurology. I am also grateful for the unlimited freedom he granted me, which allowed me to freely direct my research efforts.

I could not have made it through without the help of my secondary supervisor Dr Alan Pittman, who, since day one in the lab, patiently and carefully guided me into the fascinating world of genetics and taught me almost everything I know.

I am also immensely indebted to Professor John Hardy, who first gave me the opportunity to work in his lab as a visitor and then offered me to stay and pursue a PhD. I feel blessed to have had the chance to work close to a person who, thanks to his enlightening, enthusiastic and charismatic attitude, has made the history of genetics of neurological disorders and will always represents a source of inspiration.

Thanks also to Prof. Vincenzo Silani, who, with his open and longsighted views on education of young neurologists, gave me the opportunity to work in the laboratory of Prof John Hardy, while I was still completing my neurology training.

A special thank you goes to Professor Kailash Bhatia, for passionately teaching me the art of diagnosing movement disorders and for his great help in identifying most of the families and cases on which I have worked during my PhD. Thanks also to Prof. Tom Foltynie for the fantastic collaboration we had on the genetic characterisation of the DBS-PD cohort.

The list of people I have come across during my time at the Institute of Neurology, who helped and greatly contributed to the pleasantness of my time during the PhD, is long and I hope not to forget anyone. Big thanks certainly go to Arianna Tucci, Anna Sailer, Lucia Schottlander, Raquel Duran, Sarah Wiethoff, Sara Bandres-Ciga, Lea R’Bibo,
Heather Ging, Kin Mok, Steven Lubbe, Debbie Hughes, Mark Gaskin, Hallgeir Jonvik, and Lee Stanyer.

In particular, I would like to thank Dr Mina Ryten, Dr Daniah Trabzuni, Conceicao Bettencourt for their help with the gene network analysis and Dr Anselm Zdebik, Prof Andrey Abramov and Dr Marthe Ludtmann for assisting me with the functional work to characterise the \textit{KCDT17} mutation. Big thanks also to Dr James Polke for his help for the characterisation of the \textit{GCH1} splice-site mutation.

Thanks also to all the international collaborators who contributed to many of the projects contained in this thesis, including the Tubingen group (Prof Tom Gasser, Dr Ebba Lohmann, Dr Fritz Asmus), Wurzburg (Prof Jens Volkmann, Dr Ioannis Isaias), Milan (Dr Miryam Carecchio, Dr Barbara Garavaglia and Dr Nardo Nardocci), Takeda Pharmaceutical Company (Dr Haruhide Kimura, Dr Kosuke Nakashima, Dr Kazunori Suzuki, Ceri H. Davies) and all the members of the International Parkinson disease Genomic Consortium.

A special mention goes to Dr Marc Soutar, Dr Suran Nethisinge, Dr Joshua Hersheson, Dr Alastair Noyce, not just colleagues but fantastic friends. With your fresh, down-to-earth, outgoing, inclusive, and generous attitude you taught me a lot about science and how to be a doctor in the UK and, most importantly, you greatly helped to make me feel at home in London.

On a personal note, I am grateful to my family, my father Claudio, my mother Giacinta and my brother Jacopo, for always supporting my decision of having moved to London and pursue this PhD.

Finally, thanks to my \textit{enaim} Gila for patiently standing by my side with constant love and support since the very beginning of this journey. We have gone through a lot together. You felt, understood, and always supported without complaining my passion and dedication for this research, which certainly drained a lot of my energies. Thanks also for carefully proofreading this thesis for me, certainly another great act of love. I am really grateful and I will always be.
Chapter 1. General Introduction

1.1 Thesis Aim

Movement disorders comprise a large group of clinically and genetically heterogeneous conditions, which may be subdivided into various clinical entities, including Parkinson disease (PD), tremor, dystonia, myoclonus and chorea. Mendelian disorders are disorders in which inheritance of one (dominant) or two (recessive) mutations in a single gene leads to the development of a disease phenotype.

Many movement disorders, such as Huntington’s disease (HD) or spinocerebellar ataxias (SCAs), are transmitted as Mendelian diseases. Furthermore, a small but significant percentage of patients with diseases traditionally considered sporadic (e.g. PD or dystonia) are inherited as Mendelian traits. Despite their rarity, Mendelian diseases provide fertile ground for understanding a large number of essential biological processes that are critically involved in the pathogenesis of these complex disorders and represent fitting targets for the development of future treatments.

Although, overall, monogenic causes are rare, more than 200 genes are known to cause either an isolated movement disorder or a syndromic form of movement disorder when mutated (Klein, 2014, Spatola and Wider, 2014, Gardiner et al., 2015). However, in total, mutations in these genes only explain a small proportion of cases, suggesting that mutations in more genes await discovery.

The introduction and diffusion of ‘next generation’ sequencing (NGS) technologies, in particular whole-exome sequencing (WES) and micro-array based genotyping, has made the time and cost-effective generation of vast amounts of sequencing data increasingly easy, allowing an exponential increase in the understanding of the genetic bases of human diseases.

In the work outlined in this thesis, I used a combination of various genetic techniques and strategies to try and identify novel disease-causing genes or risk factors for familial and sporadic forms of PD, dystonia and chorea. The results of this work represent an example of how genetics can both guide the clinical practice and facilitate the understanding of the cellular pathways implicated in the pathogenesis of these diseases, paving the way for the development of future treatments.
1.2 Summary of Relevant Techniques in Genetics

What follows is a summary of the genetic techniques that I used for the work detailed in this thesis.

1.2.1 Dideoxynucleotide Sequencing

Dideoxynucleotide sequencing, commonly referred to as Sanger sequencing, named after Frederick Sanger who pioneered this technique in 1977, has been the most widely used sequencing method for the last three decades (Sanger et al., 1977).

Sanger sequencing still represents the most reliable and accurate method for variant detection and relevant NGS results are regularly validated by Sanger sequencing for the exclusion of artefacts (Pittman and Hardy, 2013).

The method relies on the principle of random inhibition of the DNA polymerase, creating newly synthesised DNA strands of different lengths that can be subsequently separated by capillary electrophoresis depending on the size. The DNA chain elongation is randomly inhibited by the inclusion in the reaction mix of a small quantity of dideoxynucleotides (ddNTPs). These differ from the normal deoxyribonucleotide (dNTP) as they lack the hydroxyl group at the 2’ and 3’ carbon positions. This feature prevents the formation of a phosphodiester bond with the subsequent dNTP, stopping DNA from extending the chain elongation any further.

The input DNA sample for Sanger sequencing is a population of identical molecules of DNA obtained by polymerase chain reaction (PCR) amplification. PCR primers are designed to target and amplify individually the portions of the genome of interest, such as the coding exons and the exon-intron boundaries.

Each of the fragments produced during the sequencing reaction will have a common 5’ end (determined by the sequencing primer) but a different 3’ end (depending on the insertion of the appropriate ddNTP into any one of the different positions that will match that base). As the different ddNTPs are combined with fluorescent dyes with different emission wavelengths, automated DNA sequencing machines, which use laser excitation during electrophoresis, are able to construct the DNA sequence from the intensity profiles of the four different fluorophores.
1.2.2 Next Generation Sequencing

In the last 5-10 years NGS technologies have completely transformed the hunt for Mendelian disease genes due to a dramatic reduction in cost and commensurate increase in the speed of DNA sequencing (Bras et al., 2012).

NGS is applied to three main techniques: whole-genome sequencing (WGS; for sequencing all coding and non-coding portions of the genome), WES (for sequencing only the coding portions of the genome) and targeted gene sequencing (for capturing and sequencing of a more limited set of genes or specific genomic regions).

Although WGS certainly represents the most comprehensive approach, so far its use has been limited by some important factors, including the cost and the capacity required to store the data. Furthermore, interpreting variants in the non-protein-coding portion of the genome is extremely challenging.

Protein-coding genes, which constitute approximately 1-2% of the human genome (the ‘exome’), harbor the majority of disease-causing mutations. For this reason WES has become the most widely used NGS technique for the identification of genes underlying Mendelian disorders.

Differently from Sanger sequencing, which is based on conventional capillary-based sequencing, NGS relies on massively parallel sequencing of DNA molecules, which are spatially separated in a flow cell. This enables the automated processing of millions of sequence reads and hundreds of gigabases of nucleotide-sequence output in parallel in a time and cost-effective manner. Different NGS chemistries, technologies and platforms- differing in specific technical details with respect to cost, accuracy, read length and multiplexing capability- are currently available (Metzker, 2010), but the details of this are beyond the scope of this chapter.

The main technical steps of the NGS process are summarised in Figure 1-1. In brief, genomic DNA is first broken up into small fragments. Then, sequence adaptors, which are platform-specific and allow each fragment to be hybridized to the flowcell where the sequencing occurs, are added to both ends of each fragment. WES protocols proceed with the hybridization of the fragments to probes that are complimentary to all the known exons in the genome, which are then captured while the remaining DNA is washed away, leaving a pool of fragments containing exons. WGS requires no extra steps following the addition of adaptors and the library is ready to be sequenced at that
point. The raw sequencing data is subsequently analysed through a bioinformatic pipeline, composed of a series of software tools, which allow the production of the sequence reads, quality score assessment of the reads, alignment of the reads to the reference genome, variant calling and annotation of the called variants.

Figure 1-1 Schematic representation of the main technical steps of NGS
Reproduced from (Bras et al., 2012)
One of the first steps and most important aspects for understanding and interpreting the results of WES is the definition of the ‘exome’. Indeed, considerable uncertainty still exists regarding which portions of the human genome truly encode for proteins. Initially, exome-capture kits used the CCDS (Consensus Coding Sequence Project) definition (Pruitt et al., 2009), which is subset of genes determined to be coding with high confidence. Currently, most of the available commercial kits target all of the genes contained in Refseq (Pruitt et al., 2014) and an increasingly large number of intronic promoter regions, hypothetical proteins and non-coding RNAs. This aspect is particularly important as mutations may be missed if they are located in portions of the genome not targeted by the used exome-capture kit.

1.2.3 Linkage Analysis

As I will discuss in the next sections of this chapter, even in the age of NGS, linkage analysis, which enables the identification of areas of the genome where the causal variants are more likely to lie, remains of invaluable importance to narrow down the number of possible candidate variants under consideration.

Briefly, genetic linkage analysis is a statistical method that is used to associate a phenotype and the chromosomal location of the causative mutation (Dawn Teare and Barrett, 2005). Linkage analysis relies on the principle that genes or genetic markers, which are located proximal to each other on a chromosome, tend to be inherited together during meiosis. The greater the frequency of recombination between two genetic markers, the further apart they are likely to be. Conversely, the lower the frequency of recombination of two markers, the closer they are. Therefore the position of a disease gene can be inferred by finding markers that tend not to recombine in all affected individuals.

The markers generally assessed in linkage analysis are, on the one hand, the detectable phenotypes of interest, and, on the other, DNA markers evenly spread throughout the genome.

In the past microsatellites, tracts of repetitive DNA in which certain DNA motifs (ranging in length from 2–5 base pairs) are repeated, used to be the most widely adopted genetic markers for linkage analysis. However, sizing of hundreds of microsatellites is a time-consuming and laborious process. Consequently, the use genome-wide arrays, which enable the simultaneous detection of thousand of single-nucleotide
polymorphisms in an automated way, has become the most popular way to generate genetic markers for linkage analysis.

Linkage analysis may be either parametric (if the pattern of transmission of the trait of interest is known) or non-parametric. Parametric linkage analysis is the traditional approach in the study of pedigree with Mendelian diseases.

The proportion of offspring in which two parental alleles are separated by recombination is the recombination fraction ($\theta$). The recombination fraction varies from 0 (for adjacent loci) to 0.5 (for distant loci) and may serve as a measure of the distance between the loci. The probability that a causative gene is linked to a genetic marker is assessed through the LOD score (logarithm -base 10- of odds), which is a function of $\theta$. The LOD score is a value that results from comparing the likelihood of obtaining the data if the two loci are indeed linked, to the likelihood of observing the same data by chance. In order to calculate the LOD score, a model for disease must be specified, which includes the frequency of the disease allele and mode of inheritance (dominant versus recessive), marker allele frequencies, and a full marker map for each chromosome.

Large positive scores are evidence for linkage, and negative scores are evidence against it. By convention, a LOD score greater than 3.0 (equivalent to $p=0.001$) is considered significant evidence for linkage. On the other hand, a LOD score $<-2.0$, indicating a probability against linkage of 100 to 1, is considered enough evidence to exclude linkage.

A monogenic mode of inheritance established by segregation analysis and a correct phenotypic designation of affected and unaffected status are the two essential requirements for successful genetic mapping using linkage analysis.
1.3 Models and Strategies to Study the Genetic Architecture of Diseases

Three main models are thought to describe the genetic basis of human diseases and, depending on the postulated underlying model, different sets of genetic tools and strategies are necessary to unravel the architecture of a disease (Singleton et al., 2010).

![Diagram](Image reproduced from Singleton et al., 2010)

**Figure 1-2** Scheme illustrating the necessary methods for gene discovery depending on the hypothesised underlying architecture of disease.

Mendelian Disease Model

Mendelian diseases are caused by very rare and highly penetrant variants (minor allele frequency [MAF] far below 1%) that often, but not always (i.e. sporadic cases with recessive or *de novo* mutations) affect multiple individuals in a family. The most effective way of identifying these types of mutations is based on mapping strategies (e.g. linkage analysis or homozygosity mapping). However, as I will discuss more in the next section, WES has revolutionised this field, introducing a series of potential new approaches.

Common Disease-Common Variant Model

Genetically complex diseases may be influenced by single variants or by a combination
of variants that are common in the general population (MAF > than 5%), but exert only a small role in increasing the disease risk. Genome-wide association studies (GWAS) have been shown to be an excellent resource for identifying these types of variants.

Common Disease-Rare Variant Model

Genetically complex diseases may also be caused or influenced by rare variants that exert a moderate effect on disease risk.

Detection of these types of variants is notoriously difficult because they do not have a strong enough effect to be transmitted as Mendelian traits and therefore elude mapping strategies, such as linkage analysis. Furthermore, multiple variants at the same locus may contribute to disease risk in a cumulative way and individual variants tend not be frequent enough to be detected through GWAS and association strategies. For these reasons, rare variants with intermediate effect represent a major challenge for genetic studies currently and in the near future.

Currently, there are two main approaches for the detection of these types of variants:

1) A candidate gene approach, whereby researchers assess the role of variants in genes that are plausibly involved in the biology of a disease. Interestingly, most of the genes identified to date through this approach had initially been suspected thanks to astute clinical observations (e.g. GBA mutations in PD or TREM2 variants in Alzheimer disease) (Sidransky and Lopez, 2012, Guerreiro et al., 2013).

2) A hypothesis-free exome (or genome)-wide approach, based on the comparison of NGS data obtained from large cohorts of cases and controls. This approach requires the assembly of very large cohorts of cases and controls and, to date, no novel risk genes for movement disorders have been identified. However, this approach has recently led to the identification of TUBA4A and TBK1 mutations in amyotrophic lateral sclerosis/fronto-temporal dementia (Smith et al., 2014, Cirulli et al., 2015).

The architecture of complex diseases, such as PD, may be composed of a combination of various types of causative variants, and some genes may contain multiple types of variants contributing to the disease risk (pleomorphic gene loci).

For the purposes of this thesis, I have elected to focus only on the discovery of rare Mendelian variants and rare variants with intermediate effect.
1.4 Strategies for Gene Identification in Mendelian Diseases

When attempting to identify pathogenic variants, researchers should select the correct technological and analytical approach based on the most likely genetic architecture of the disease of interest, the suspected type of variants, the frequency of the disease of interest, the inheritance pattern, and the penetrance of the mutations. The strategies for gene identification in the NGS era are schematically summarised in Figure 1-3.

The identification of mutations causing Mendelian diseases has traditionally relied on the use of mapping strategies (e.g. linkage analysis or homozygosity mapping), followed by gene identification through positional cloning or candidate gene approach (i.e. prioritising genes for Sanger sequencing within the mapped genomic region, based
on the biological plausibility or possible relevance to the disease of interest).

While in theory the power of NGS would not be absolutely necessary for this approach, NGS has made gene discovery tremendously easier when used in combination with traditional mapping strategies. When informative families with several affected individuals from the same family are available, the disease locus can be narrowed down significantly and then NGS can be used to interrogate all the coding exons and splicing regions of the genes contained within the locus for the presence of the causative mutations.

For dominant traits, linkage analysis (discussed in detail in section 1.2.3) represents the gold standard for genetic mapping (Figure 1-3A). However, families that are genetically informative enough for linkage analysis are exceedingly rare. Thus, application of this approach for gene discovery is becoming harder by the day, as there are not many families left which may be suitable for this analysis.

For the study of recessive disorders where a homozygous variant is suspected (e.g. study of consanguineous families), NGS can be complemented by the use of homozygosity mapping (Figure 1-3B), which relies on the identification of regions that are identical by descent in affected individuals only (Alkuraya, 2013).

As discussed previously, NGS can be also used to rapidly screen candidate disease-causing or predisposing genes, selected based on their biological plausibility or on grounds of clinical observations hinting towards a possible genetic association (Figure 1-3C).

As discussed in more detail in the next chapter, applications of these traditional strategies in combination with NGS has been already successfully used in multiple occasions for the discovery of novel genes causing movement disorders.

Besides facilitating traditional mapping strategies, NGS has also offered the opportunity to identify novel disease-causing gene mutations in ways that simply were not possible before the advent of NGS (Olgiati et al., 2016).

For the study of dominant families that are too small for the application of linkage analysis, one novel approach is sequencing the exomes (or genomes) of several distant affected individuals belonging to the same family (Figure 1-3D) and look for shared
variants after applying filtering criteria to prioritise the likely pathogenic variants (see next section for details of prioritisation strategies). Depending on the number of family members available for sequencing and their distance in the pedigree, this approach can result in the identification of relatively small lists of candidate variants that can be then used for replication studies in additional families or cohorts of similarly affected subjects.

For the study of recessive families where homozygosity mapping is not applicable (e.g. non-consanguineous pedigrees), the double hit strategy (Figure 1-3E) may be applied to look for compound heterozygous mutations. Here, the exomes of siblings affected by the same condition are interrogated for the presence of at least two separate, likely pathogenic, mutations in the same gene. With this strategy, when the number of affected individuals is small (i.e. only two affected siblings), it is useful to sequence also the DNA of the healthy parents, as this makes it possible to exclude genes carrying multiple mutations inherited from the same parent (and therefore located on the same allele), a mode of inheritance not compatible with a recessive disease.

Another important novel approach is the study of family trios by NGS, looking for the presence of de novo mutations (Figure 1-3F). These are mutations present in the DNA of the affected subject, but not in the DNA of the parents. This type of approach is generally used for individuals without a family history with severe neurological phenotypes that reduce the fitness to procreate. To date, this approach has been incredibly successful for the study of autism-spectrum disorders, schizophrenia and intellectual disability (Rauch et al., 2012, Xu et al., 2012, Gulsuner et al., 2013, Genome of the Netherlands Consortium, 2014, Iossifov et al., 2014).

Finally, NGS can also be used to study a cohort of unrelated patients, looking for the presence of causative mutations in the same gene (Figure 1-3G). This strategy is more likely to be effective for identifying the disease-causing gene when the disease of interest is rare and highly specific (e.g. presence of radiological hallmarks), as this would significantly reduce the possibility of genetic heterogeneity (i.e. multiple genes responsible for the same phenotype). This approach has been proven to be successful even with relatively small cohort of carefully characterised patients (Haack et al., 2012, Simons et al., 2013).
1.5 Identifying Causal Alleles: How to Prioritise Variants?

The key challenge in NGS is the interpretation of the large number of variants that are routinely detected. An average ‘exome’ would reveal the presence of ~24,000 coding and splice-site variants in individuals of African origin and ~20,000 variants in individuals of European origin (Bamshad et al., 2011). Of these, on average, ~98% are known benign polymorphisms, and about ~50-60% are synonymous variants, meaning that they not lead to an amino acid change. However, hundreds of novel non-synonymous variants are found in the exome of each individual. Among these, only one (in case of dominant disease) or two (in recessive diseases) variants will be causative in most families with a Mendelian disease. Therefore, the crucial and more challenging step in the search of THE disease-causing mutation is recognising and differentiating the true disease-related alleles among the myriad of benign common and rare/novel variants and sequencing artefacts.

I will describe in the next section the strategies that are commonly used to filter and reduce the number of possible pathogenic variants, schematically summarised in Figure 1-4.

![Figure 1-4 Prioritization of NGS candidate variants.](image)

Image reproduced from (Olgiati et al., 2016)
1.5.1 Filtering Based on the Inheritance Mode and Pedigree Information

In the case of Mendelian disorders, a correct understanding of the mode of inheritance can effectively reduce the search space for candidate pathogenic changes. As previously explained, the pattern of inheritance of a monogenic disorder will greatly influence both the experimental design (e.g. the number of cases to sequence and the most informative cases to select for sequencing) and the analytical approach.

Generally, family studies will examine variants that are shared between multiple affected relatives and consequently, it is preferable to select for sequencing subjects that are as distantly related as possible, as they will share a smaller number of genes. Where a dominant mode of inheritance is suspected, only shared heterozygous variants are considered, unless there is the suspicion that recessive mutations are transmitted in a pseudo-dominant way due to multiple loops of consanguinity in the pedigree. When a recessive inheritance is suspected, homozygous or compound heterozygous variants should be considered. Recessive inheritance or de novo dominant mutations should be suspected in sporadic cases. However, an increasing number of reports indicate that de novo mutations can also be inherited by multiple individuals in the same family due to germinal mosaicism in one of the unaffected parents, leading to pseudo-recessive presentations (Kancheva et al., 2015).

Segregation of candidate variants, usually by Sanger sequencing, in other affected family members can further refine the list of variants. Unaffected individuals can be also of great help in reducing the number of candidate variants, as variants carried by these subjects are less likely to be responsible for the phenotype. However, two caveats should be kept in mind when using unaffected subjects to filter out variants: (i) pathogenic variants may be associated with incomplete penetrance (meaning that carriers of the variants may not manifest the disease); (ii) carriers of pathogenic variants may be too young at the time of the genetic study to firmly exclude that they will not develop the disease.

Importantly, segregation analysis alone in a single family cannot definitively implicate a variant as disease-causing, as there is the possibility that a separate, unobserved pathogenic mutation may lie on the same haplotype as the candidate variant.

Finally, it is important to mention that phenocopies, cases with the familial phenotype or
closely mimicking it but not carrying the causative variant, may be present in pedigree, greatly complicating the correct interpretation of the pedigree structure.

In summary, a substantial number of benign variants can be excluded from the list of disease-causing candidates using as filters the pedigree structure and the assumed mode of inheritance of the disease.

**1.5.2 Filtering Based on Rarity**

Filtering the list of variants against a variety of publicly available datasets, which contain NGS data from thousands of individuals originating from different human populations, is essential to removing common benign variations and narrowing the fraction of candidate variants down to a workable number.

The list of the most widely used datasets of genetic variations includes the Exome Aggregation Consortium (ExAC; http://exac.broadinstitute.org/), dbSNP version 129 and 137 (www.ncbi.nlm.nih.gov/projects/SNP), 1000 Genomes project (1kGP; www.1000genomes.org), NHLBI Exome Variant Server (EVS; evs.gs.washington.edu) and Complete Genomics 69 (www.completegenomics.com/public-data/69-Genomes).

Furthermore, variants should also be filtered against data from large cohorts of individuals sequenced in-house. This step is particularly important for excluding sequencing artifacts and false positive calls, which tend to be specific for the sequencing chemistry and the pipeline used in the analysis of the NGS data.

The population frequency of a variant is particularly important as it would be expected that Mendelian disease causing alleles are extremely rare, if not absent, in the general population. For dominant disorders, it is sensible to consider only novel or extremely rare variants. Alleles causative for recessive diseases are usually present in the population at MAF of < 0.01. However, it is essential to remember that these are largely arbitrary cut-offs. The frequency filter should be set depending on the frequency of the disease of interest in the general population, the clinical presentation (i.e. early onset/aggressive versus late onset/indolent phenotype), and the expected penetrance of the putatively disease-causing variants.

There are other two important issues that should be considered during this crucial step. First, the frequency of both benign and pathogenic variants can vary substantially among different ethnicities and certain variants are found exclusively in certain specific
populations. This means that variants that are completely absent in one population could be much more frequent in another. Therefore, it is essential to filter the WES results against data derived from ethnicity-matched populations.

Secondly, it is important to consider that well-established pathogenic mutations (e.g. the \textit{LRRK2} mutations p.Gly2019Ser causing autosomal dominant PD) may be present in population variant databases, although at low frequencies. Being too strict with filtering based on MAF may therefore sometimes result in the exclusion of pathogenic alleles from the analysis. This is particularly true for diseases such as PD, which tend to manifest later in life and thus non-manifesting carriers of pathogenic alleles may be present in large control populations.

1.5.3 Filtering Based on Deleteriousness and Conservation

Further stratification of variants can be undertaken based on predictions of their deleteriousness. Synonymous variants are generally excluded from the analysis, as it is expected that causative alleles for Mendelian diseases will be missense, nonsense, frameshift or splice site variants. Importantly, although in rare instances, synonymous variants can potentially affect splice sites and therefore result in loss-of-function alleles (Korvatska \textit{et al.}, 2013). In addition, variants located in deep intronic and in other non-coding genomic regions are generally removed from the list of candidates under the assumption that disease causing-mutations would affect the coding region of the genome. However, one should be aware that this assumption is based on the fact that, as a matter of feasibility, in the Sanger sequencing era only the coding portion of the genes were screened. It is hence plausible that a significant proportion of the missing variability responsible for Mendelian diseases is located in the intronic region.

Greater importance may be given to loss-of-function variants (i.e. frameshift, stop codons and variants disrupting canonical splice sites), than to missense variants.

Additional strategies can be used to further prioritize the remaining candidate variants, especially if they are missense. A powerful approach uses the quantitative estimates of mammalian evolution at the nucleotide level. This approach relies on the principle that deleterious mutations generally present high sequence conservation as a result of purifying selection. Therefore, highly conserved nucleotides and amino acids are likely to be very relevant for the function of the gene and the encoded protein. Examples of tools used in this thesis for assessment and quantification of evolutionary conservation
are PhiloP (Pollard et al., 2010) and GERP++ (Davydov et al., 2010).

The predicted deleteriousness of amino acid substitutions can also be assessed using *in silico* tools such as SIFT (http://sift.jcvi.org/) (Kumar et al., 2009), PolyPhen2 (http://genetics.bwh.harvard.edu/pph2/) (Adzhubei et al., 2010), Provean (http://provean.jcvi.org/) (Choi et al., 2012), MutationTaster (http://www.mutationtaster.org/) (Schwarz et al., 2014) and CADD (http://cadd.gs.washington.edu/home) (Kircher et al., 2014).

Some of these tools (e.g. SIFT or Provean) rely on specific biological aspects of deleterious mutations (often conservation) to classify variants in discrete classes of pathogenicity (first-principles approaches). Another group (e.g. Polyphen-2 or MutationTaster) consists of classifiers that are trained to distinguish between a set of known deleterious mutations against a set of benign polymorphisms (trained classifiers).

Importantly, these methods are particularly useful when assessing the pathogenicity of putatively loss-of-function variants, either causing recessive diseases or causing dominant diseases through haplo-insufficiency or dominant negative effects. However, the *in silico* tools are less reliable for the assessment of dominant variants that are pathogenic through a gain-of-function mechanism.

### 1.6 Limitations of Next-Generation Sequencing

Despite many successful reports of disease-associated novel genes identified by WES, it is not precisely known how often this approach fails, as negative results are not reported most of the times. The experience in our department teaches that failure in the analysis of kindreds with suspected Mendelian diseases is unfortunately not uncommon.

NGS bears a number of important limitations, which have to be kept in mind when the causative variant is not identified.

1. Current NGS technologies easily detect single nucleotide variants or small deletions are, but there remain significant issues with the identification of large deletions and copy number variants (CNVs).

2. The currently used short read libraries (~250bp) lead to particular problems with the detection of repeat expansions, which not uncommonly are the responsible cause for neurodegenerative diseases, such as the dominant SCAs or HD.
3. Variable coverage of the targeted regions, as mutations may be missed if they are located in poorly covered areas.

4. The causative gene is not in the exome capture kit target definition (for instance, it is not a known gene, or there is a failure in the design)

Finally, even once a single or few promising candidate variant are found in a given pedigree, the significant genetic heterogeneity that exists for many inherited neurological conditions, can hamper the confirmation of the finding, which generally relies on the identification of other mutations in the same gene in unrelated pedigrees.
Chapter 2. Overview on Genetics of Selected Movement Disorders

2.1 The Genetics of Parkinson disease

2.1.1 Introduction
PD is the second most common neurodegenerative disease after Alzheimer disease, affecting >2% of those over 75 years (Mayeux et al., 1995).

The main clinical features of PD and clinical diagnostic criteria include bradykinesia and at least one other sign amongst muscular rigidity, rest tremor, or postural instability (Figure 2-1). A progressive course, unilateral presentation with asymmetrical signs, a classic pill rolling tremor, good sustained response to levodopa (L-DOPA) and a number of non-motor symptoms accompanying or preceding the motor onset (eg. hyposmia, constipation, genito-urinary symptoms, REM-behavioural sleep disorders) generally support the diagnosis (Lees et al., 2009, Noyce et al., 2016).

Figure 2-1 Typical appearance of a patient with Parkinson disease.
Reproduced from (Lees et al., 2009)
The pathological signature of PD is a selective loss of pigmented dopaminergic neurons from the pars compacta of the substantia nigra (Figure 2-2A). Nigral neuron degeneration and consequent decrease in dopaminergic striatal innervation result in classic PD motor symptoms. This nerve cell loss is accompanied by three distinctive intraneuronal inclusions: Lewy bodies (LBs; Figure 2-2B), the pale body, and the Lewy neurite.

α-synuclein (α-syn) is the main component of LBs (Spillantini et al., 1997). According to the model proposed by Braak and colleagues, α-syn pathology may spread ascending from the dorsal motor nucleus of the vagus and involving progressively, the lower brainstem, the olfactory bulb, the midbrain and finally the cortex (Braak et al., 2003).

Figure 2-2 Pathological features of Parkinson disease
(A) Cross section of the midbrain showing the pigmented substantia nigra in a normal brain (bottom) and depigmented nigra in a brain with PD (upper). (B) Microscopic section of a substantia nigra pigmented neuron containing neuromelanin (white arrow) and a Lewy body (black arrow) within the cytoplasm of the neuron. The Lewy body has a dense core and a lighter halo.
The aetiology of PD is poorly understood and, similarly to other neurodegenerative conditions, age is the major risk factor. However, 10% of people with the disease are younger than 45 years of age (early-onset PD).

PD was long thought to be a sporadic disorder without genetic causation with most of the research interest focused on the identification of environmental risk factors (reviewed in Noyce et al., 2012).

Several studies estimating the genetic contribution to PD pathogenesis using twins have shown low concordance rates in monozygotic and dizygotic twins (Duvoisin et al., 1981, Ward et al., 1983, Marttila et al., 1988). However, in studies based on PET scan data, rather than simple cross-sectional observance of the phenotype, the concordance rate was significantly higher for monozygotic twins, than for dizygotic twins (55% versus 18%), suggesting a substantial genetic contribution to the PD pathogenesis (Piccini et al., 1999). It is estimated that approximately 5-10% of people with PD have a positive family history in a first degree-relative.

Over the last 25 years, since the identification of the a-synuclein (SNCA) locus, the views on the aetiology of PD have profoundly changed. Mutations in a growing list of genes have been incontrovertibly linked to monogenic forms of PD (recently reviewed in Hernandez et al., 2016). Furthermore, genome-wide-association studies (GWAS) have succeeded in identifying many common, low risk variants that are likely to play a direct role in the aetiology of the common sporadic disease as well (Nalls et al., 2014). In addition, rare variants with intermediate-effect have been identified as important risk factors (e.g. GBA). A schematic overview of the frequency and penetrance of the genes linked to PD is shown in Figure 2-3.
Figure 2-3. Schematic representation of the frequency and penetrance of the Parkinson disease genes
*Pleiotropic loci. Figure adapted from (Verstraeten et al., 2015)

2.1.2 Monogenic Causes of Parkinson Disease

The genes associated with monogenic forms of PD are listed in Table 2-1.

Mendelian forms of PD are overall rare and generally have an earlier disease manifestation than sporadic PD. Mutations in genes identified as causing Mendelian forms of PD explain 30% of familial PD and only 3%-5% of patients with sporadic disease occurrences (Kumar et al., 2011). In the next section I will describe the main clinical and biological features of the most relevant Mendelian genes that have been shown to cause PD.
<table>
<thead>
<tr>
<th>Locus</th>
<th>Gene</th>
<th>Features</th>
<th>Inheritance</th>
<th>Mutational spectrum</th>
<th>Pathology</th>
</tr>
</thead>
<tbody>
<tr>
<td>PARK1/PARK4</td>
<td>SNCA</td>
<td>Early and late-onset, fast progression, severe dementia,</td>
<td>AD</td>
<td>Missense and gene dosage</td>
<td>Synucleinopathy</td>
</tr>
<tr>
<td>PARK2</td>
<td>Parkin</td>
<td>Juvenile- or young-onset parkinsonism, slow progression</td>
<td>AR</td>
<td>Missense, truncating and gene dosage</td>
<td>Synucleinopathy (occasionally)</td>
</tr>
<tr>
<td>PARK6</td>
<td>PINK1</td>
<td>Early-onset parkinsonism, slow progression</td>
<td>AR</td>
<td>Missense, truncating and gene dosage</td>
<td>Synucleinopathy (occasionally)</td>
</tr>
<tr>
<td>PARK7</td>
<td>DJ-1</td>
<td>Early-onset parkinsonism, psychiatric features, slow progression</td>
<td>AR</td>
<td>Missense, truncating and gene dosage</td>
<td>Synucleinopathy</td>
</tr>
<tr>
<td>PARK8</td>
<td>LRRK2</td>
<td>Late onset parkinsonism, slow progression</td>
<td>AD</td>
<td>Missense</td>
<td>Synucleinopathy, tauopathy</td>
</tr>
<tr>
<td>PARK9</td>
<td>ATP13A2</td>
<td>Juvenile and early onset, atypical, fast progression</td>
<td>AR</td>
<td>Missense, truncating</td>
<td>Not known</td>
</tr>
<tr>
<td>PARK14</td>
<td>PLA2G6</td>
<td>Juvenile and early onset, atypical, fast progression</td>
<td>AR</td>
<td>Missense, truncating</td>
<td>Synucleinopathy</td>
</tr>
<tr>
<td>PARK15</td>
<td>FBX07</td>
<td>Juvenile and early onset, atypical, fast progression</td>
<td>AR</td>
<td>Missense, truncating</td>
<td>Not known</td>
</tr>
<tr>
<td>PARK17</td>
<td>VPS35</td>
<td>Late onset parkinsonism, slow progression</td>
<td>AD</td>
<td>Missense</td>
<td>Not known</td>
</tr>
<tr>
<td>PARK17</td>
<td>DNAJC6</td>
<td>Juvenile and early onset, atypical, fast progression</td>
<td>AR</td>
<td>Missense, truncating</td>
<td>Not known</td>
</tr>
<tr>
<td>PARK20</td>
<td>SYNJ1</td>
<td>Juvenile and early onset, atypical, fast progression</td>
<td>AR</td>
<td>Missense, truncating</td>
<td>Tauopathy</td>
</tr>
<tr>
<td>NA</td>
<td>DNAJC13</td>
<td>Late-onset, typical, slow progression</td>
<td>AD</td>
<td>Missense</td>
<td>Synucleinopathy</td>
</tr>
<tr>
<td>NA</td>
<td>RAB39B</td>
<td>Early onset, intellectual disability</td>
<td>X-linked</td>
<td>Missense, truncating and gene dosage</td>
<td>Synucleinopathy</td>
</tr>
<tr>
<td>NA</td>
<td>CHCHD2</td>
<td>Late onset parkinsonism, slow progression</td>
<td>AD</td>
<td>Missense</td>
<td>Synucleinopathy?</td>
</tr>
<tr>
<td>NA</td>
<td>VPS13C</td>
<td>Early onset, fast progression</td>
<td>AR</td>
<td>Truncating</td>
<td>Synucleinopathy</td>
</tr>
</tbody>
</table>
**Autosomal Dominant Parkinson Disease**

**SNCA (PARK1, PARK4)**

The identification in 1997 of the p.Ala53Thr missense mutation in *SNCA* as a cause of autosomal dominant PD in families of Greek and Southern Italian origin was the first proof for a genetic aetiology in a subset of PD cases (Polymeropoulos *et al.*, 1997). Furthermore soon after the identification of SNCA mutations in familial PD cases it was discovered that α-Syn is the major component of LBs, the pathological hallmark of PD, therefore linking familial to sporadic PD (Spillantini *et al.*, 1997).

Subsequently, whole-locus duplication and triplication of *SNCA* were identified, proving that not only structurally altered α-synuclein, but also simple overexpression of the wild-type protein is pathogenic (Singleton *et al.*, 2003, Kara *et al.*, 2014). The p.Ala30Pro and p.Glu46Lys mutations have been detected in a single family of German and Spanish ancestry, respectively (Kruger *et al.*, 1998, Zarranz *et al.*, 2004). Only recently, three additional *SNCA* point mutations (p.His50Gln, p.Gly51Asp and Ala53Glu) have been discovered (Kiely *et al.*, 2013, Proukakis *et al.*, 2013, Pasanen *et al.*, 2014).

The missense mutations are pathogenic through a toxic gain of function, whereby they increase the tendency of α-Syn to form β-pleated sheets, which, in turn, are thought to predispose α-Syn to form the aggregates that are localised in LBs (Devine *et al.*, 2011).

Patients with *SNCA* missense mutations usually have EOPD with initial good response to L-DOPA. The disease tends to be aggressive, with dementia often complicating early on the clinical picture.

A dose relationship between α-Syn levels and severity of the clinical phenotype is observed. In *SNCA* multiplications families, duplications (three copies of *SNCA*) are not fully penetrant and lead to late-onset dopa-responsive PD, as seen in typical sporadic PD (Ibanez *et al.*, 2009); triplications (four copies of *SNCA*) are conversely fully-penetrant and cause a much more aggressive phenotype with earlier onset and severe rapidly progression, almost inevitably complicated by early dementia and additional features such as dysautonomia (Farrer *et al.*, 2004, Fuchs *et al.*, 2007).

Brain pathology in patients with *SNCA* mutations is characterized by abundant α-Syn-
positive neuronal inclusions.

**LRRK2 Mutations (PARK8)**

The PARK8 locus was first mapped on chromosome 12p11.23-q13.11 in a large Japanese family with autosomal dominant PD (Funayama et al., 2002). A few years later two independent groups identified the causative mutations underlying the PARK8 locus in the **LRRK2** gene (Paisan-Ruiz et al., 2004, Zimprich et al., 2004).

Mutations in **LRRK2** have been since then recognized as the most common Mendelian cause of PD. A large number of variants have then been described in PD cases. However, based on segregation with disease in families and frequency in population controls, only few of these variants are firmly considered pathogenic: p.Arg1441Gly, p.Arg1441Cys, p.Arg1441His, p.Tyr1699Cys, p.Gly2019Ser and p.Ile2020Thr (Di Fonzo et al., 2005, Gilks et al., 2005, Zabetian et al., 2005).

The most frequent mutation is by far the c.6055G>A; p.Gly2019Ser which is found in up to 40% of cases of PD in North African populations (Lesage et al., 2006), in about 20% of Ashkenazi Jewish patients (Ozelius et al., 2006) and 1-7% of PD patients of European origin (Di Fonzo et al., 2005, Clark et al., 2006, Zabetian et al., 2006).

The p.Arg1441Cys is the second most common mutation and has been identified worldwide (Nuytemans et al., 2008). Conversely the p.Arg1441Gly and the p.Ile2020Thr are very rare and population-specific, being the first mainly found in the Basques (Simon-Sanchez et al., 2006) and the second in the Japanese population (Tomiyama et al., 2006).

The penetrance of the **LRRK2** p.Gly2019Ser mutation has been established to be ~28% in subjects younger than 60 years of age, but increases up to 74% at 79 years of age (Healy et al., 2008). Of note, there has been a recent report of two monozygotic twins in their 70s with the p.Gly2019Ser but discordant for the presence of PD, suggesting that the variability in onset and penetrance between **LRRK2**, maybe largely non-genetic in origin (e.g. stochastic; environmental) (Xiromerisiou et al., 2012). Furthermore, a recent report suggested that the penetrance in the Ashkenazi Jewish population could be as low as 25% by the age of 80, a figure much lower than in other ethnic background, suggesting the existence of genetic or environmental factors which affect the penetrance of the p.Gly2019Ser mutation in this population (Marder et al., 2015).
The majority of patients with LRRK2 mutations presents with late-onset asymmetric L-DOPA-responsive tremor-predominant PD, clinically indistinguishable from PD (Healy et al., 2008). Neuropathological findings showed pleomorphic pathologies overlapping with other neurodegenerative diseases, including LBs, tau- and ubiquitin-containing inclusions and pure nigral degeneration without LBs (Zimprich et al., 2004, Giasson et al., 2006).

LRRK2 is a large gene that consists of 51 exons and encodes the 2527-amino acid protein leucine-rich repeat kinase 2. The physiological and pathological functions of LRRK2 have not yet been fully characterized. Several pathogenic LRRK2 mutations cluster in functionally important domains and impair their function: p.Arg1441Cys, p.Arg1441Gly, and p.Arg1441His reduce protein GTPase activity, whereas the p.Gly2019Ser and p.Ile2020Thr mutations in the kinase domain have been shown to increase its activity (Simon-Sanchez et al., 2006).

VPS35 (PARK17)

A single missense mutation in VPS35, p.Asp620Asn, has been identified by two groups through WES as an important cause of autosomal-dominant PD (Vilarino-Guell et al., 2011, Zimprich et al., 2011). Several other studies in which the sequencing of the entire VPS35 coding region was performed have disclosed a handful of other missense mutations (p.Gly51Ser, p.Met57Ile, p.Thr82Arg, p.Ile241Met, p.Pro316Ser, p.Arg524Trp, and p.Leu774Met). However, none of these additional mutations have been proven to be pathogenically relevant (Lesage et al., 2012, Sharma et al., 2012, Nuytemans et al., 2013). Overall, VPS35 the p.Asp620Asn mutation is a rare cause of PD accounting for only about 1% of familial parkinsonism and 0.2% of sporadic PD. VPS35 encodes a highly-conserved component of the retromer, a complex that mediates retrograde transport of transmembrane cargo from endosomes back to the trans-Golgi network (Seaman, 2005, Bonifacino and Rojas, 2006). Clinically VPS35-related PD appears to be similar to idiopathic PD (Sheerin et al., 2012).

DNAJC13

WES and linkage analysis identified the missense mutation p.Asns55Ser in DNAJC13 in a Saskatchewan autosomal-dominant family affected by multi-generational PD with evidence LB pathology. Subsequent genotyping in 2928 patients and 2676 control subjects from Canada, Norway, Taiwan, Tunisia, and the USA identified four additional
patients carrying the same mutation, of which two had familial parkinsonism, but none in controls (Vilarino-Guell et al., 2014). DNAJC13 regulates the dynamics of clathrin coats on early endosomes. Cellular analysis shows that the mutation confers a toxic gain-of-function and impairs endosomal transport. Subsequent studies have shown that the DNAJC13 p.Asn855Ser substitution is not a common cause of PD among Caucasian populations (Gustavsson et al., 2015, Lorenzo-Betancor et al., 2015).

**CHCHD2**

Funayama and colleagues used linkage analysis and WES to successfully identify in a Japanese dominant PD family a missense variant in the CHCHD2 gene, encoding the mitochondrial protein coiled-coil-helix-coiled-coil-helix domain-containing 2 (Funayama et al., 2015). Screening of a further 340 index patients with autosomal dominant PD identified three additional pathogenic missense and splice-site mutations, p.Thr61Ile, p.Arg145Gln, and c.300+5G>A. Subsequently, other independent groups have confirmed that coding mutations in CHCHD2 represent a rare genetic cause of autosomal dominant PD and may also represent a risk factor for sporadic PD and other synucleinopathies (Jansen et al., 2015, Ogaki et al., 2015, Koschmidder et al., 2016, Shi et al., 2016, Zhang et al., 2016).

**Autosomal Recessive Parkinson Disease**

**PARK2**

Mutations in PARK2, firstly identified by positional cloning in Japanese families (Matsumine et al., 1997, Kitada et al., 1998), are the most common cause of autosomal recessive, EOPD. PARK2 mutations are detected in up to ~50% of cases with recessive familial EOPD (Lucking et al., 2000), as well as ~10% of sporadic cases with onset before the age of 45 (Klein and Lohmann-Hedrich, 2007). A range of loss-of-function variants are detected, including missense, truncating mutations and whole-exon deletions or duplications, most commonly in the compound heterozygous state.

Whether heterozygous mutations in PARK2 represent a risk factor for PD is still a matter of debate (Klein et al., 2007) and circumstantial evidence suggests that copy-number variants may increase the disease risk (Huttenlocher et al., 2015), but not point mutations (Kay et al., 2007, Kay et al., 2010).

PARK2-related disease usually shows a slowly progressive course with a dramatic
response to dopaminergic treatment. No specific features differentiate PARK2 mutation carriers from other EOPD forms; however, symmetrical onset, dystonia at onset, hyperreflexia, slower progression of the disease, a tendency toward a better response to L-DOPA, and preservation of olfaction might be more frequent among patients with PARK2 mutations (Khan et al., 2003, Alcalay et al., 2011).

Post-mortem examination of PARK2-mutated cases shows a more restricted distribution of the pathology than in classic sporadic PD, with severe but selective ventral nigral degeneration and absent or scarce LB pathology (Doherty et al., 2013).

PARK2 encodes parkin, a protein that functions as an E3 ubiquitin ligase in the process of ubiquitination, a form of post-translational modification that conjugates ubiquitin proteins to lysine residues of target proteins, in order to address them towards proteosomal degradation. Most PARK2 mutations lead to loss of its E3 ligase function. Disturbed elimination of damaged mitochondria, which is an important cellular process called mitophagy, appears to be one of the most important mechanisms for neuronal death in parkin deficiency (Deas et al., 2011).

PINK1 (PARK6)

Mutations in phosphatase and tensin homolog-induced putative kinase 1 (PINK1) are the second-most common cause of autosomal recessive EOPD, accounting for ~1% to 8% of recessive familial EOPD and ~1% of sporadic EOPD cases (Healy et al., 2004, Rogaeva et al., 2004, Valente et al., 2004, Bonifati et al., 2005). Mutations are found across different populations and one mutation with a founder effect is particularly frequent in the Philippines (Bonifati et al., 2005).

PINK1 is a 581 amino acid ubiquitously expressed protein kinase. It consists of an amino-terminal mitochondrial targeting motif, a conserved serine-theonine kinase domain, and a carboxy-terminal autoregulatory domain. Two-thirds of the reported pathogenic mutations determine a loss-of-function of the PINK1 kinase domain, supporting the relevance of PINK1 enzymatic activity in the pathogenesis of PD. More than 90% of the reported mutations in PINK1 are missense mutations, whereas the remaining are copy number mutations or truncating mutations. Similarly to PARKIN, the role of heterozygous PINK1 mutations is debated (Abou-Sleiman et al., 2006).

Clinically, PINK1 mutation carriers present with early-onset, slowly progressive L-
DOPA-responsive parkinsonism, initial gait impairment and often psychiatric symptoms, such as depression and anxiety. Cognitive involvement is rare. Differently from PARK2 mutated cases, olfactory dysfunction is common in PINK1 (Ferraris et al., 2009).

The first post-mortem analysis in a case with PINK1-linked PD showed neuronal loss in the substantia nigra and the presence of LBs (Samaranch et al., 2010). However, a more recent report showed absence of LB pathology (Takanashi et al., 2016).

PINK1 and Parkin function in a common pathway for detecting mitochondrial damage and selectively eliminating corrupted mitochondria. PINK1 is stabilized on mitochondria with lower membrane potential and recruits Parkin from the cytosol. Once recruited to mitochondria, Parkin becomes enzymatically active and initiates the autophagic clearance of mitochondria by lysosomes (i.e., mitophagy) (Youle and Narendra, 2011).

**DJ-1 (PARK7)**

Bonifati and colleagues identified mutations in DJ-1 as an additional cause of EOPD (Bonifati et al., 2003). DJ-1 mutations are more rare than PARK2 and PINK1, being found in only ~1% of EOPD cases (Abou-Sleiman et al., 2003, Pankratz et al., 2006). The clinical phenotype is similar to that of PARK2- and PINK1-related EOPD (Abou-Sleiman et al., 2003). Dementia and features of motorneuron disease have been described in individual cases with bi-allelic DJ-1 mutations (Annesi et al., 2005, Hanagasi et al., 2016).

**RAB39B**

Genetic analysis of an Australian and of an American family with a combination of EOPD and intellectual disability with an X-linked pattern of inheritance identified a large deletion and a missense mutation (p.Thr168Lys) in the RAB39B gene (Wilson et al., 2014). All affected cases initially presented in childhood with various degrees of intellectual disability, including developmental delay, cognitive impairment,
macroencephaly, and, seizures. EOPD appeared later on, with tremor as the first symptom. Post-mortem studies show that \textit{RAB39B} mutations causes extensive nigral dopaminergic neuronal loss with widespread LB pathology.

Other groups have subsequently replicated this finding and found \textit{RAB39B} mutations in individuals affected by classic EOPD, in absence of the atypical signs observed in the families where the gene was first identified (Lesage \textit{et al.}, 2015, Mata \textit{et al.}, 2015).

Interestingly, previous work had identified \textit{RAB39B} as a cause of X-linked intellectual disability in absence of parkinsonian signs (Giannandrea \textit{et al.}, 2010), indicating that \textit{RAB39B} mutations cause a spectrum of overlapping syndromes. Furthermore, this suggests that shared pathogenic mechanisms may exist linking PD to developmental disorders. Rab proteins are small molecular weight guanosine triphosphatases involved in the regulation of vesicular trafficking (Cheng \textit{et al.}, 2002).

\textbf{VPS13C}

Very recently, \textit{VPS13C} loss of function variants have been recognised as a novel cause of autosomal-recessive EOPD (Lesage \textit{et al.}, 2016). Homozygous or compound heterozygous truncating mutations were detected in three isolated PD subjects. The phenotype associated with \textit{VPS13C} mutations is characterised by rapid and severe disease progression and early cognitive decline. Neuropathological assessment showed evidence of diffuse LB pathology. Functional work demonstrated that VPS13C knock-down leads to a severe mitochondrial pathology and exacerbation of PINK1/Parkin-dependent mitophagy.

\textbf{Complex Forms of Autosomal Recessive Parkinsonism}

Mutations in the genes described in this section are generally associated with more rare and aggressive forms of recessive parkinsonism, characterised by very early-onset (<30 years) and usually additional, atypical features (pyramidal, dystonic, ocular movement, and cognitive disturbances). However, in spite of the different phenotypic presentation, several biological evidences link the pathogenic mechanisms of these conditions to that of PD. Furthermore, several recent observations indicate that mutations in most of these genes can lead to more typical forms of autosomal recessive EOPD.

\textbf{ATP13A2 (PARK9)}

Recessive loss-of-function of \textit{ATP13A2} cause Kufor-Rakeb syndrome, a rare, juvenile-
onset disorder characterized by severe parkinsonism, dementia, pyramidal signs and supranuclear gaze palsy (Ramirez et al., 2006). Subsequently, ATP13A2 recessive mutations have been detected in cases with a broader phenotype, including isolated juvenile-parkinsonism (Di Fonzo et al., 2007, Martino et al., 2015), parkinsonism with brain iron accumulation (Schneider et al., 2010) and, more recently, also neuronal ceroidal lipofuscinosis (a severe infantile neurodegenerative disorder) (Bras et al., 2012).

ATP13A2 is a large gene encoding for a protein normally located in the lysosomal membrane (Ramirez et al., 2006). Loss of ATP13A2 leads to lysosomal dysfunction and subsequent α-Syn accumulation (Tsunemi and Krainc, 2014).

**PLA2G6 Mutations (PARK14)**

A broad range of truncating recessive mutations in the phospholipase A2 group VI (PLA2G6) gene, encoding for a phospholipase involved in metabolism of phospholipids to free fatty acids, were first described as the cause of infantile neuroaxonal dystrophy and neurodegeneration associated with brain iron accumulation (Kurian et al., 2008).

PLA2G6 mutations were subsequently detected also in subjects with a different presentation, featuring L-DOPA-responsive dystonia-parkinsonism, pyramidal signs and cognitive/psychiatric features, with onset in early adulthood, without evidence of iron accumulation (Paisan-Ruiz et al., 2009). More recently, bi-allelic missense mutations, more likely to have a milder effect on protein function, have also been identified in cases with uncomplicated EOPD (Xie et al., 2015, Giri et al., 2016).

LBs have been described in association with PLA2G6 mutations (Paisan-Ruiz et al., 2012). Functional work performed in a Drosophila model of PLA2G6 deficiency shows that loss of normal gene activity leads to lipid peroxidation and mitochondrial dysfunction (Kinghorn et al., 2015).

**FBOX7 (PARK15)**

FBOX7 mutations were first described in a Persian family and then in an Italian family with severe juvenile-onset parkinsonism associated with spasticity and prominent pyramidal signs (Najim al-Din et al., 1994, Di Fonzo et al., 2009). More recently a homozygous FBOX7 missense mutation was described also in a family with more classic EOPD (Lohmann et al., 2015). FBOX7, like Parkin, is part of an E3 ubiquitin
ligase complex necessary for the correct initiation of the mitophagy process (Laman, 2006, Burchell et al., 2013).

**SYNJ1**

The same homozygous mutation in **SYNJ1** (p.Arg258Gln) was identified by WES, in two unrelated consanguineous kindred’s, as a cause of a complex autosomal recessive parkinsonian syndrome (Krebs et al., 2013, Quadri et al., 2013). The phenotype includes aggressive parkinsonism with severe levodopa-induced dystonia. Additional features are generalized seizures, cognitive decline, dysarthria, eyelid apraxia and supranuclear gaze palsy.

**SYNJ1** encodes synaptojanin 1, a phosphoinositide phosphatase protein involved in clathrin-mediated endocytosis in the adult brain (Drouet and Lesage, 2014). A recent report showed that bi-allelic loss of synaptojanin 1 cause severe infantile epileptic encephalopathy with brain accumulation of tau protein (Dyment et al., 2015, Dyment et al., 2015).

**DNAJC6**

Recessive truncating mutations in **DNAJC6** cause juvenile onset parkinsonism with rapid progression. Associated features include dysarthria, generalised seizures and brain atrophy on MRI (Edvardson et al., 2012, Koroglu et al., 2013). The group of Bonifati recently described **DNAJC6** mutations in cases with a milder presentation, consistent with classic EOPD (Olgiati et al., 2016). **DNAJC6** encodes the protein auxilin, which is a clathrin-associated protein enriched in nerve terminals and may play a role in synaptic vesicle recycling (Ahle and Ungewickell, 1990).

2.1.3 Rare Variants Predisposing to PD

**GBA**

Homozygous or compound heterozygous loss-of-function mutations in **GBA** are the genetic abnormality underlying Gaucher’s disease (GD), the most common lysosomal storage disorder. GD has a prevalence of 1/40,000 to 1/50,000 in all ethnicities and is particularly frequent in the Ashkenazi Jewish population, where it affects 1/800 to 1/1,000 live births (Cox, 2001, Grabowski, 2008).

GD is an autosomal recessive disorder and more than 300 different causative **GBA**
mutations spanning the entire coding region have been described (Figure 2-4), including insertions, deletions, point mutations and complex alleles resulting from recombination events with the neighbouring pseudogene GBAP (Hruska et al., 2008).

![Figure 2-4 GBA gene structure and distribution of pathogenic mutations across the coding exons](image)

Image reproduced from (Hruska et al., 2008)

GD is caused by an insufficient activity of glucocerebrosidase (GCase), the lysosomal enzyme coded by GBA. Pathogenic mutations affect the enzyme’s catalytic function, cellular stability and/or subcellular trafficking, resulting in intracellular build-up of GCase’s main substrate, glucosylceramide, mainly in the cells of the reticuloendothelial system of the spleen, liver and bone marrow, and often resulting in organomegaly (Grabowski, 2008).

Clinically, GD presents with a spectrum of manifestations, the severity of which partially depend on the deleteriousness of the mutation on GCase activity: type 1 GD, the non-neuronopathic, most common and mildest form of GD; type 2 (GD2) the acute neuronopathic, most severe form associated with a complete deficiency in GCase activity and short life-expectancy; and type 3 (GD3) the chronically neuronopathic form which has a later onset and longer life-expectancy than GD2.
The main clinical features of GD are summarised in the panel below taken from (Sidransky and Lopez, 2012).

<table>
<thead>
<tr>
<th>The three types of Gaucher's disease</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Type 1: non-neuronopathic</strong></td>
</tr>
<tr>
<td>- Panethnic disorder, although more common in Ashkenazi Jews</td>
</tr>
<tr>
<td>- Associated with clinical heterogeneity</td>
</tr>
<tr>
<td>- Wide range of symptom severity</td>
</tr>
<tr>
<td>- Hepatosplenomegaly, anaemia, and thrombocytopenia are common</td>
</tr>
<tr>
<td>- Bone disease is a frequent cause of morbidity</td>
</tr>
<tr>
<td>- Treated with enzyme replacement therapy</td>
</tr>
<tr>
<td><strong>Type 2: acute neuronopathic</strong></td>
</tr>
<tr>
<td>- Rare, panethnic disorder</td>
</tr>
<tr>
<td>- Can present prenatally, at birth, or in the first year of life</td>
</tr>
<tr>
<td>- Rapidly progressive neurological deterioration</td>
</tr>
<tr>
<td>- Enzyme replacement therapy does not reverse or halt neurological progression</td>
</tr>
<tr>
<td>- Early death within days to years</td>
</tr>
<tr>
<td><strong>Type 3: chronic neuronopathic</strong></td>
</tr>
<tr>
<td>- Includes several different phenotypes with variable longevity</td>
</tr>
<tr>
<td>- Accompanied by a specific disorder of horizontal saccadic eye movements</td>
</tr>
<tr>
<td>- Some patients develop myoclonic epilepsy</td>
</tr>
<tr>
<td>- A subgroup of patients develop cardiac calcification, hydrocephalus, and other abnormalities</td>
</tr>
<tr>
<td>- Associated with distinct learning disabilities in some patients</td>
</tr>
</tbody>
</table>
In the past decade a previously unexpected link between GD and PD has been recognised. The first findings were based on isolated clinic observations of GD cases presenting a parkinsonian picture with clinical features very similar to those classically seen in patients with sporadic PD (Neudorfer et al., 1996, Tayebi et al., 2001, Tayebi et al., 2003). These observations were subsequently extended to the GD first-degree relatives, obligate carriers of GBA heterozygous mutations (Goker-Alpan et al., 2004).

Then, many independent case-control studies, performed in PD cases from different ethnical backgrounds (reviewed in (Sidransky and Lopez, 2012), have then consistently observed a higher frequency of GBA mutations in PD patients (Figure 2-5), indicating a causal relationship between GD and PD.

GBA mutations are now considered the most relevant genetic risk factors for PD (Neumann et al., 2009, Sidransky et al., 2009). Among Ashkenazi Jewish subjects, GBA mutation are found in ~15% of patients and among non-Ashkenazi Jewish subjects mutations are identified in ~7%, with an odds ratio for any GBA mutation in patients versus controls of ~5.

Neuropathological analysis of brains from parkinsonian patients with GD or carriers of heterozygous GBA mutations revealed the presence of abundant α-syn positive LBs (Bembi et al., 2003, Neumann et al., 2009).

Clinically, patients with PD and GBA mutations (both homozygous and heterozygous carriers) present with a classic clinical picture, indistinguishable from that of sporadic PD. Patients are characterized by asymmetric resting tremor, rigidity and bradykinesia and a good response to L-DOPA treatment. Most series of patients with GBA mutations report a lower mean age at onset and a higher incidence of cognitive and behavioural symptoms, such as dementia and visual hallucinations, compared to wild type cases (Alcalay et al., 2012, Winder-Rhodes et al., 2013).

Importantly, the replacement therapy used to treat GD cases does not modify the risk of PD. It is not clear though whether the drug simply does not cross the blood-brain barrier or whether enzyme replacement is not targeting the pathogenic mechanisms critical for PD (Rosenbloom et al., 2011).
Figure 2-5 Map showing the countries where the frequency of GBA mutations in Parkinson disease patients was assessed
Image taken from (Sidransky and Lopez, 2012)
Despite the numerous investigations that followed from the discovery of this association, the molecular nature of the link between GBA mutations, PD and α-syn aggregation remains largely unexplained. Furthermore there exists a deepening debate as to whether the mutant GBA protein acts by a loss- or toxic-gain-of-function mechanism (Goldin, 2010).

As previously mentioned, GD is caused by a drastic reduction in the activity of the encoded lysosomal enzyme β-glucocerebrosidase (GCase) and a consequent accumulation of substrates, mainly glucosylceramide and glucosyphingosine, in the tissues where GD is associated with lysosomal storage dysfunction.

However, while it would seem likely that GCase loss of enzymatic activity contribute to the risk of PD, many elements argue against this view, suggesting the mutant GBA could instead increase the risk of PD through other mechanisms.

Most patients with GBA mutations and PD are heterozygous carriers, which implies that at least 50% of residual enzymatic activity is conserved. This is largely sufficient to avoid the pathogenic build up of substrates, as enzymatic activity needs to be reduced by at least 70-90% before substrate accumulation occurs. However, it is possible that GBA has other distinct unidentified regulatory functions of lysosomal activity the loss of which could favour α-syn aggregation and neuronal cell death. Furthermore, studies assessing age-specific penetrance and estimated life cumulative risk for PD in GD patients and heterozygous carriers have demonstrated that the risk is similar in the two groups (from 9 to 12% by the age of 80), thus arguing against the presence of a gene dosage effect, which is a typical of pure loss-of-function mutations (Rosenbloom et al., 2011, McNeill et al., 2012, Rana et al., 2013, Alcalay et al., 2014).

On the other hand, there is growing evidence of a direct relationship between the severity of the mutation, in terms of GCase activity reduction, and PD risk. For instance, Gan-Or et al. found that carriers of severe mutations had a considerably higher disease risk and reduced age at onset compared to carriers of mild mutations (Gan-Or et al., 2008, Gan-Or et al., 2015). Moreover, Lesage et al. found that the PD onset tended to occur earlier in carriers of severe GBA mutations (Lesage et al., 2011). Finally, recent work by Barrett and colleagues determined in a cohort of relatives of GD1 patients that non-N370S mutations (severe or null mutations) conferred a higher risk for PD compared to the mild N370S mutation (Barrett et al., 2013).
Different investigations carried out in cell and animal models of GD and synucleinopathies support the idea that both mechanisms, toxic gain of function due to misfolded mutant proteins and loss of enzymatic activity, are likely to equally contribute to the complex lysosomal dysfunction that leads to aberrant α-syn aggregation and ultimately to nigral neuronal death (Cullen et al., 2011, Mazzulli et al., 2011, Sardi et al., 2011).

**SPMD1**

*SPMD1* encodes sphingomyelin phosphodiesterase 1, a lysosomal enzyme that cleaves the phosphocholine head group of sphingomyelin to generate ceramide. Bi-allelic mutations in the gene are the cause of Niemann-Pick type A, another multi-systemic lysosomal disorder, particularly frequent in the Ashkenazi Jewish population (Schuchman, 2007).

Gan-Or and colleagues recently suggested that a heterozygous missense mutation in this gene, p.Leu230Pro, could be a risk factor for the development of PD in the Ashkenazi Jewish population, with an odds ratio of ~9 (Gan-Or et al., 2013). Importantly, other groups have subsequently replicated this finding in populations of different ethnical background (Foo et al., 2013, Wu et al., 2014), strongly supporting the view that dysfunctional lysosomal activity is critical to α-syn accumulation and PD pathogenesis.

### 2.1.4 Common Low-Risk Variants Predisposing to PD

Candidate gene association studies have been used in the past to try and identify common loci that increase the risk for PD. Through this approach common variability at the *SNCA* locus (Kruger et al., 1999) and at the *MAPT* locus were shown to contribute to the risk of disease (Golbe et al., 2001).

However, it is mainly thanks to unbiased GWAS that in recent years a growing number of susceptibility genes and risk loci have been undoubtedly associated with increased risk of PD. These studies rely on the common-disease-common variant hypothesis, according to which common genetic variants of small effect size can increase the risk of diseases, such as PD, that are frequent in the general population. GWAS have been made possible by the development of platforms that allow the assessment of genetic variability (usually single-nucleotide polymorphisms; SNPs) across the genome with the disease on the order of 500,000 to 1 million variants (Hardy and Singleton, 2009).
While the first GWAS in PD were underpowered and therefore yielded negative results (Fung et al., 2006), larger GWAS identified that common variability in several loci, including \textit{SNCA}, \textit{MAPT}, \textit{LRRK2}, \textit{PARK16}, \textit{BST1}, \textit{HLA-DRB5}, and \textit{GAK}, increase PD risk in independent populations (Pankratz et al., 2009, Satake et al., 2009, Simon-Sanchez et al., 2009, Edwards et al., 2010, Saad et al., 2011, Spencer et al., 2011, Sharma et al., 2012)

Subsequently, thanks to imputation of variants and a meta-analysis of different GWAS datasets, additional genes were identified, including \textit{ACMSD}, \textit{STK39}, \textit{MCCC1/LAMP3}, \textit{SYT11}, and \textit{CCDC62/HIP1R} (Nalls et al., 2011).

The largest and most recent meta-analysis of all existing GWAS, which includes a total of \~13,000 PD cases and \~95,000 controls, showed that 24 risk loci containing 28 independent risk variants are associated with an increased risk of PD (Nalls et al., 2014). The Manhattan plot showing the results of this meta-analysis is represented in Figure 2-6.

Importantly, most of the SNPs associated with increased PD risk are non-coding and the functional role of the vast majority of them is far from being understood. Furthermore, the effect size of the individual risk variants is low, although initial evidence suggests that the risk conferred by the combination of multiple risk variants may be more substantial.

It is of interest to note that some of the most robustly associated SNPs are located in close proximity of genes that also contain high penetrant mutations causing Mendelian forms of familial PD, such as \textit{SNCA}, \textit{LRRK2} and \textit{VPS13C}. Importantly, this suggests that there could be a substantial overlap in the pathogenic mechanisms underlying rare Mendelian forms of PD and the more common sporadic forms, reinforcing the crucial importance of dissecting the bases of PD with Mendelian transmission.
Figure 2-6 Manhattan plot from the latest GWAS meta-analysis for Parkinson disease
Image reproduced from (Nalls et al., 2014)
2.2 Overview on the Genetics of Dystonia

2.2.1 Introduction

The dystonias are a clinically and genetically heterogeneous group of movement disorders. In 2013, an international panel of experts provided a consensus update on definition, phenomenology and classification of dystonia (Albanese et al., 2013). Dystonia was defined as a “movement disorder characterized by sustained or intermittent muscle contractions causing abnormal, often repetitive, movements, postures, or both. Dystonic movements are typically patterned, twisting, and may be tremulous. Dystonia is often initiated or worsened by voluntary action and associated with overflow muscle activation”.

Dystonia is the third most common movement disorder worldwide and approximately 70,000 people are affected by dystonia in the UK alone (Paudel et al., 2012).

Dystonia has been traditionally classified based on the four main features (Table 2-2): (1) age of onset (early onset vs. adult onset), (2) distribution of affected body parts (focal, segmental or generalised), (3) aetiology (primary vs secondary dystonia) (4) special clinical features (paroxysmal, exercise-induced, task-specific or dopa-responsive)(Albanese et al., 2011). Given the existence of several problems and confusion with this system, a new classification has been recently proposed, which uses two axes to define dystonia. (Albanese et al., 2013).

Axis I describes the clinical features of dystonia and includes age at onset, body distribution, temporal pattern, coexistence of other movement disorders and other neurological manifestations.

Axis II focuses on the presumed aetiology; the two classifying criteria are the presence of nervous system pathology (evidence of degeneration vs structural lesions vs absence of any pathology) and a differentiation among inherited dystonia (when a precise genetic cause is recognized) and acquired dystonia (e.g. perinatal brain injury, infections, drugs among others) or idiopathic dystonia. The old and the new classification are shown and compared in the figure Figure 2-7.
Table 2-2 Traditional clinical classification of dystonia

<table>
<thead>
<tr>
<th>A) Aetiology</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Primary dystonia</td>
<td>No underlying structural, pharmacological or metabolic cause can be identified</td>
</tr>
<tr>
<td>1.1) Primary isolated dystonia</td>
<td>Dystonia is the only clinical sign (apart from tremor) and there is no identifiable exogenous cause or other inherited or degenerative disease</td>
</tr>
<tr>
<td>1.2) Primary plus dystonia</td>
<td>Dystonia is a prominent sign but is associated with another movement disorder, e.g. myoclonus or parkinsonism.</td>
</tr>
<tr>
<td>1.3) Primary paroxysmal dystonia</td>
<td>Torsion dystonia occurs in brief episodes with normalcy in between.</td>
</tr>
<tr>
<td>2. Heredodegenerative dystonia</td>
<td>Dystonia is a feature, among other neurological signs, of a genetically determined neurodegenerative disorder</td>
</tr>
<tr>
<td>3. Secondary dystonia</td>
<td>Dystonia is a symptom of an identified neurological condition, e.g. a focal brain lesion, exposure to drugs or chemicals</td>
</tr>
</tbody>
</table>

B) By age at onset

| 1. Early onset (≤30 years of age) | Usually starts in a leg or arm and frequently progresses to involve other limbs and the trunk |
| 2. Late onset | Usually starts in the neck (including the larynx), the cranial muscles or one arm. Tends to remain localized with restricted progression to adjacent muscles |

C) By distribution of affected body parts

| 1. Focal | Single body region (e.g. writer’s cramp, blepharospasm) |
| 2. Segmental | Contiguous body regions (e.g. cranial and cervical, cervical and upper limb) |
| 3. Multifocal | Non-contiguous body regions (e.g. upper and lower limb, cranial and upper limb) |
| 4. Generalised | Both legs and at least one other body region (usually one or both arms) |
| 5. Hemidystonia | Half of the body (usually secondary to a lesion in the contralateral basal ganglia) |
Figure 2-7 Schematic comparison of the old and the new classification system of dystonia
Table reproduced form (Balint and Bhatia, 2014).
The pathophysiology basis of dystonia is poorly understood. It has been proposed that dystonia may derive from a functional impairment of the neural circuits underlying the control and coordination of voluntary movements (Figure 2-8). These include the basal ganglia, the cerebellum, the sensorimotor cortex, and the interactions between these regions of the brain (Neychev et al., 2011). The impairment of these circuits may occur as a consequence of structural lesions of these brain areas (either focal lesions or degeneration), as a consequence of a pharmacological treatment altering the activity of the neurons constituting these pathways (e.g. L-DOPA treatment in PD patients or antipsychotic treatments blocking the dopamine receptors), or, most frequently, in absence of any overt brain pathology, which defines the primary dystonias. This last group of conditions is thought to derive from anatomically preserved, but functionally abnormal, basal ganglia and cerebellar circuits.

![Figure 2-8 The circuits of dystonia](image.png)

Reproduced from (Berman and Jinnah, 2015)
2.2.2 Overview of Monogenic Forms of Dystonia

Genetic factors largely contribute to the pathogenesis of dystonia. A positive family history in first-degree family members is reported in a very high proportion of patients with primary dystonia, ranging from 18 to 50% of cases (Rubio-Agusti et al., 2013). Moreover, in a small but defined subset of families, dystonia can be transmitted across generations as a clear Mendelian trait.

Mutations in a growing number of genes are responsible for the Mendelian forms of dystonia. Importantly, the identification of these genes has allowed the initial recognition of different cellular pathways that could be involved in the molecular pathogenesis of primary dystonia, including perturbed synaptic transmission and plasticity (in particular of the dopamine system), abnormal transcription and cell-cycle regulation and endoplasmic reticulum (ER) dysfunction (Ledoux et al., 2013).

The genes that have been associated with dystonia are summarised in Table 2-3.

This overview will only focus on genes that cause primary dystonia.

The Primary Isolated Dystonias

This term defines a group of syndromes in which dystonia is the only clinical feature (except for tremor of the arms or head and neck), without any evidence of neurodegeneration or any obvious secondary cause.

TOR1A Mutations (DYT1-related Dystonia)

In 1997 an in-frame 3 base pairs (CAG) deletion in exon 5 of the TOR1A gene, was identified as the genetic abnormality underlying the DYT1 locus (Ozelius et al., 1997). This mutation is responsible for approximately 50% of cases with early onset generalized dystonia across different ethnicities (Kramer et al., 1994). Due to a founder mutation, the DYT1 mutation is particularly frequent among Ashkenazi Jewish population, where it accounts for 80% of early onset dystonia.

Clinically, TOR1A-related dystonia would present in childhood with dystonic posturing of lower limbs and subsequent generalisation. Dystonia can nevertheless begin in any part of the body, with usual sparing of the neck and face. Late onset forms and atypical presentation of the disorder are not infrequently recognised (Jamora et al., 2006, Stamelou et al., 2013). Most cases would report a family history consistent with
autosomal dominant inheritance. However, the penetrance of the mutation is markedly reduced (~30%), which explains the frequent observation of pseudo-sporadic cases (Hjermind et al., 2002). A small number of other variants have been described in sporadic cases with various forms of dystonia, but none of them has been convincingly shown to be pathogenic due to lack of clear segregation in other family members (Leung et al., 2001, Kabakci et al., 2004, Vulinovic et al., 2014).

TorsinA is a member of the AAA+ superfamily (ATPases associated with different cellular activities) with multiple functions in the cell. TOR1A is exclusively expressed in neurons and the protein localizes in the endoplasmic reticulum (ER). Cellular models have demonstrated that mutant torsinA is displaced from the ER to the nuclear envelope (Goodchild and Dauer, 2004). The abnormal localisation may result in dysfunction of the secretory pathway and synaptic vesicle machinery (e.g. transportation of dopamine vesicles and/or dopamine receptors to the pre- and post-synaptic membrane, respectively) (Warner et al., 2010).

Recently the first mouse model of DYT1-dystonia, exhibiting a motor phenotype closely resembling the human disease, has been described (Liang et al., 2014).
<table>
<thead>
<tr>
<th>Locus</th>
<th>Gene locus</th>
<th>Features</th>
<th>Inheritance</th>
<th>Mutational spectrum</th>
</tr>
</thead>
<tbody>
<tr>
<td>DYT1</td>
<td>TOR1A</td>
<td>Early-onset primary torsion dystonia</td>
<td>AD with reduced penetrance</td>
<td>Only an in-frame 3 base pairs (CAG) deletion in exon 5 is pathogenic</td>
</tr>
<tr>
<td>DYT2</td>
<td>HPCA</td>
<td>Early-onset primary dystonia with prominent cranio-cervical involvement</td>
<td>AR</td>
<td>Missense</td>
</tr>
<tr>
<td>DYT4</td>
<td>TUBB4A</td>
<td>Whispering dystonia (adult onset spasmodic dysphonia) with generalisation and ‘hobby horse’ gait. Hypomyelination of basal ganglia and cerebellum (H-ABC). Spastic paraparesis</td>
<td>AD/De novo</td>
<td>Missense</td>
</tr>
<tr>
<td>DYT5a</td>
<td>GCH1</td>
<td>Dopa-responsive dystonia</td>
<td>AD/AR</td>
<td>Missense, truncating and large deletions or duplications</td>
</tr>
<tr>
<td>DYT5b</td>
<td>TH</td>
<td>Dopa-responsive dystonia</td>
<td>AR</td>
<td>Missense and truncating</td>
</tr>
<tr>
<td>NA</td>
<td>SPR</td>
<td>Dopa-responsive dystonia</td>
<td>AR</td>
<td>Missense and truncating</td>
</tr>
<tr>
<td>DYT6</td>
<td>THAP1</td>
<td>Adult-onset torsion dystonia with prominent cranio-cervical and laryngeal involvement</td>
<td>AD with reduced penetrance</td>
<td>Missense and truncating</td>
</tr>
<tr>
<td>DYT8</td>
<td>MR1</td>
<td>Paroxysmal non-kinesigenic dyskinesia</td>
<td>AD</td>
<td>Missense</td>
</tr>
<tr>
<td>DYT</td>
<td>Gene</td>
<td>Disorder/Description</td>
<td>Inheritance</td>
<td>Mutation</td>
</tr>
<tr>
<td>-----</td>
<td>------</td>
<td>----------------------</td>
<td>-------------</td>
<td>----------</td>
</tr>
<tr>
<td>10</td>
<td>PRRT2</td>
<td>Paroxysmal kinesigenic dyskinesia</td>
<td>AD</td>
<td>Mainly truncating and large deletions</td>
</tr>
<tr>
<td>11</td>
<td>SGCE</td>
<td>Myoclonic dystonia (often with alcohol responsiveness)</td>
<td>AD</td>
<td>Mainly truncating and large deletions</td>
</tr>
<tr>
<td>12</td>
<td>ATP1A3</td>
<td>Rapid onset dystonia parkinsonism and alternating hemiplegia of childhood</td>
<td>AD/de novo</td>
<td>Missense</td>
</tr>
<tr>
<td>16</td>
<td>PRKRA</td>
<td>Early-onset dystonia-parkinsonism</td>
<td>AR</td>
<td>Missense</td>
</tr>
<tr>
<td>18</td>
<td>SLC2A1</td>
<td>Paroxysmal exercise-induced dyskinesia +/- epilepsy</td>
<td>AD</td>
<td>Missense and truncating</td>
</tr>
<tr>
<td>23</td>
<td>CIZ1</td>
<td>Adult onset cervical dystonia</td>
<td>AD</td>
<td>Missense</td>
</tr>
<tr>
<td>24</td>
<td>ANO3</td>
<td>Tremulous cranio-cervical dystonia +/- upper limb tremor</td>
<td>AD</td>
<td>Missense</td>
</tr>
<tr>
<td>25</td>
<td>GNAL</td>
<td>Adult onset cervical dystonia</td>
<td>AD with reduced penetrance /de novo</td>
<td>Missense and truncating</td>
</tr>
</tbody>
</table>
**THAP1 mutations (DYT6)**

In 2009 mutations in the *THAP1* gene, coding for the thanatos-associated protein domain-containing apoptosis-associated protein 1, were identified in two Mennonite families as the genetic abnormality responsible for the DYT6 locus (Almasy *et al.*, 1997, Fuchs *et al.*, 2009). A range of missense, nonsense and frameshift mutations spread throughout the coding portion of the gene are associated with disease in several genetically diverse populations. The inheritance of DYT6-related dystonia is autosomal dominant with a penetrance of ~60% (Saunders-Pullman *et al.*, 2007, Fuchs *et al.*, 2009, Houlden *et al.*, 2010).

Clinically, mutation carriers tend to develop symptoms later than DYT1-positive patients, with an onset in adolescence or adulthood. DYT6-positive patients usually present with a predominant involvement of the upper body, in particular severe oromandibular, cranio-cervical, laryngeal or upper limbs involvement, and subsequent generalisation (Xiromerisiou *et al.*, 2012).

*THAP1* is an atypical zinc finger protein with DNA binding properties, involved in the regulation of gene transcription (Roussigne *et al.*, 2003). One hypothetical mechanism whereby *THAP1* substitutions could lead to dystonia is by dysregulated transcription of strategic genes. Mutations in *THAP1* have been shown to disrupt binding to the *TOR1A* promoter and result in decreased *TOR1A* gene expression, suggesting that under normal conditions *TOR1A* is negatively regulated by *THAP1* (Kaiser *et al.*, 2010).

**TUBB4A (DYT4-related Dystonia)**

The DYT4 family is a large kindred of British origin characterised by onset in the third decade of a severe form of dominantly inherited generalized dystonia (Parker, 1985). Several affected family members spread across five generations have been described. Affected members would present with a laryngeal dysphonia which progresses to a generalized dystonia. Some patients displayed a unique ‘hobby horse’ gait (Wilcox *et al.*, 2011). To date, no other pedigree with similar clinical features has been published.

Through a combination of linkage analysis and exome sequencing, the missense mutation p.Arg2Gly in the *TUBB4A* gene was independently identified as causal in the DYT4 kindred by our and Klein’s group (Hershson *et al.*, 2012, Lohmann *et al.*, 2012).
TUBB4A encodes β-tubulin-4a, a constituent of axonal microtubules. The mutation lies in the highly conserved autoregulatory MREI (Methionine–Arginine–Glutamic acid–Isoleucine) domain of the protein. The MREI tetrapeptide sequence at the start of the N-terminal domain is known to be necessary for the autoregulation of the β-tubulin mRNA levels and in vitro studies using site-directed mutagenesis have previously shown that mutations in this domain abrogate the autoregulatory ability (Yen et al., 1988, Yen et al., 1988).

Screening of the gene in a cohort of 394 unrelated dystonia patients revealed one further possibly pathogenic variant (p.Ala271Thr) in an individual who exhibited autosomal dominant spasmodic dysphonia (Lohmann et al., 2012). We could not detect other mutation carriers in a cohort of ~200 dystonia cases with prominent laryngeal involvement (Hersheson et al., 2013). A more recent study investigating 575 subjects with primary laryngeal, segmental or generalized dystonia failed to detect any further TUBB4A mutations (Vemula et al., 2014), indicating overall that TUBB4A mutations are an exceptionally rare cause of dystonia.

Importantly, it was subsequently recognised that de novo mutations in TUBB4A are also the cause of hypomyelination with atrophy of the basal ganglia and cerebellum (H-ABC syndrome) (Simons et al., 2013, Hamilton et al., 2014). H-ABC is a rare neurodegenerative disorder with onset in infancy and childhood, characterised by the specific combination of hypomyelination, cerebellar atrophy, and absence or disappearance of the putamen (van der Knaap et al., 2002). Clinically, H-ABC cases show a combination of severe dystonia, cerebellar ataxia, and variable degrees of intellectual disability. There is ongoing debate about whether DYT4-dystonia and H-ABC are distinct allelic conditions (brain imaging of members of the DTY4 kindred does not show the classic H-ABC findings) or whether the DYT4-phenotype simply represents a milder clinical presentation of the same disease, possibly questioning TUBB4A as a primary dystonia gene (Erro et al., 2015).

CIZ1 mutations (DYT23)

Linkage analysis and WES were recently used to identify a missense mutation in CIZ1 (c.790A>G p.Ser264Gly) as the possible causal variant in a large Caucasian kindred with adult-onset autosomal dominant cervical dystonia (Xiao et al., 2012). Screening of other unrelated cases with adult-onset dystonia identified 2 additional missense
mutations in 3 individuals (p.Pro47Ser and p.Arg672Met; segregation not confirmed for any of these mutations), but also a similar number of novel variants in matched controls. Two subsequent studies have failed to identify additional mutations in CIZ1 (Ma et al., 2013, Dufke et al., 2015). This finding is still awaiting independent confirmation and several concerns have been raised questioning the real contribution of CIZ1 mutations to the pathogenesis of dystonia (Klein et al., 2012).

CIZ1 encodes Cip1-interacting zinc finger protein 1 and is a p21Cip1/Waf1-interacting zinc finger protein, expressed in brain and involved in DNA synthesis and cell-cycle control. The work detailing the features of a CIZ1 knock-out mouse model has been recently published (Xiao et al., 2016). Ciz1 showed the highest expression in the cerebellum and Ciz1-/- mice did not display features of dystonia, but exhibited mild motoric abnormalities on balance, open-field activity, and gait. Whole-genome gene expression analysis based on RNA extracted from cerebellum suggests that CIZ1 may be involved in the post-mitotic differentiation of neurons in response to external signals (Xiao et al., 2016).

ANO3 mutations (DYT24)

Our group identified a segregating missense mutation in ANO3 in a three-generation British family with autosomal dominant cranio-cervical dystonia (Charlesworth et al., 2012). Five further novel variants were detected by a subsequent screening of the gene in 188 probands with dystonia (Charlesworth et al., 2012).

Clinically, patients with ANO3 mutations exhibit focal or segmental tremulous dystonia, variably affecting the cranio-cervical, laryngeal and brachial regions. Some individuals showed isolated upper limb tremor and had been misclassified as essential tremor. Furthermore two patients exhibited myoclonic jerks of the head or the arms (Stamelou et al., 2014).

A recent screening in a large German cohort of sporadic cases with cervical dystonia identified two novel mutations in cases, but also a similar number of novel variants in controls, indicating that caution should be posed when assigning pathogenicity of individual variants (Zech et al., 2014). However, a screening conducted in the Chinese population identified two additional dominant families with cranio-cervical dystonia that carried a segregating ANO3 mutation (Ma et al., 2015), confirming the contribution
of ANO3 to dystonia pathogenesis.

ANO3 is very high and selectively expressed in the striatum, a key brain region in dystonia pathogenesis. ANO3 encodes a protein called anoctamin 3 and belongs to a family of closely related genes (ANO1-10) with different tissue expression patterns. ANO1 and ANO2 encode for Ca\textsuperscript{2+}-activated chloride channels. Hence, it is postulated that ANO3 may function in the same way and regulate striatal neurons excitability (Caputo et al., 2008).

**GNAL mutations (DYT25-related dystonia)**

In 2012 heterozygous mutations in GNAL were identified in two unrelated dominant families as a novel cause of familial adult-onset cervical dystonia (Fuchs et al., 2013). Screening of further 39 families revealed 6 additional segregating novel mutations, suggesting that GNAL mutations could be a relatively common cause for familial dystonia.

However, a number of subsequent mutational screening have shown a much lower mutational yield with a reported frequency across different populations of around 0.5-1%. Several loss-of-function mutations have been described, including missense and truncating mutations, (Dobricic et al., 2014, Dufke et al., 2014, Kumar et al., 2014, Zech et al., 2014). GNAL mutations are most frequently transmitted as a dominant trait, although de novo mutations have been found in sporadic cases (Dobricic et al., 2014, Kumar et al., 2014, Ziegan et al., 2014, Carecchio et al., 2016). More recently a family with severe childhood-onset dystonia harboring a recessive homozygous GNAL mutation has also been reported (Masuho et al., 2016).

Clinically, most GNAL mutation carriers present the first symptoms in the cervical region. Progression to other body segments is observed in ~50% of cases, although a frank generalisation of dystonia is rare (Fuchs et al., 2013).

GNAL encodes Go\textsubscript{olf}, the alpha subunit of triheteromeric G protein G\textsubscript{olf}, a striatal protein involved in dopamine (D1) signalling. Go\textsubscript{olf} is responsible for coupling D1 receptors to adenylyl cyclase 5 in striatal neurons and Go\textsubscript{olf} is required for D1-mediated behaviour and biochemical effects (Herve et al., 1993, Zhuang et al., 2000, Corvol et al., 2001).
HPCA-related dystonia (DYT2)

Our group has recently identified homozygous and compound heterozygous HPCA mutations in two pedigrees with autosomal recessive generalised dystonia. HPCA were initially identified through homozygosity mapping and WES in a pedigree with three affected subjects, product of a consanguineous marriage (parents are first cousins), presenting with early-onset generalised dystonia (Charlesworth et al., 2015). Subsequently, compound-heterozygous mutations in HPCA were also identified in a second independent kindred affected by AR isolated dystonia.

HPCA are likely to represent an extremely rare cause of dystonia, as a recent study assessing HPCA mutation frequency in large cohort of primary dystonia cases, failed to identify any additional mutation (Dobricic et al., 2016).

HPCA encodes hippocalcin, a neuronal calcium sensor protein present almost exclusively in the brain and expressed at particularly high levels in the striatum (Charlesworth et al., 2015).

Primary Dystonia-plus

The dystonia-plus syndromes (also known as combined dystonias, according to the most recent nomenclature system) represent a heterogeneous group of non-neurodegenerative diseases, where additional neurological features (non-epileptic myoclonus or parkinsonism) accompany dystonia (Asmus and Gasser, 2010).

DOPA-responsive dystonia (DYT5)

The term DOPA-responsive dystonia (DRD) defines a group of rare conditions, which share the clinic characteristic of a dramatic improvement in symptoms in response to treatment with oral L-dopa (Kurian et al., 2011). Mutations in three genes have incontrovertibly been shown to cause DRD: GCH1 (GTP Cyclohydrolase 1), TH (Tyrosine Hydroxylase) and SPR (Sepiapterin Reductase) (Ichinose et al., 1994, Brautigam et al., 1998, Bonafe et al., 2001). All three genes encode enzymes that are critically involved in the biosynthesis of dopamine in nigrostriatal cells and pathogenic mutations result in a dramatic decrease in dopamine synthesis (Figure 2-9).
Heterozygous GCH1 mutations are by far the most common cause of DRD, accounting for 60-80% of autosomal dominant DRD pedigrees (Furukawa, 2004). Penetrance is incomplete being lower for males (~40-50%) than females (~80%) (Wider et al., 2008).

Clinical presentation is typically with lower limb dystonia and gait disturbance, with a variable degree of generalisation over time (Nygaard et al., 1990). The response even to low doses of L-DOPA is excellent, sustained and not generally associated with dyskinesias that often follow chronic use of L-DOPA in PD (Clot et al., 2009). As I will discuss in more detail in chapter 4.4, parkinsonism may occur in the later stages of the disease.

TH and SPR mutations are associated with autosomal recessive DRD (Ludecke et al., 1996, Blau et al., 2001). Typically, patients with mutations in these two genes present with a severe combination of dystonia, parkinsonism, ptosis, seizures and truncal hypotonia in infancy. The clinical picture is overall more severe than in GCH1-related DRD and response to L-DOPA treatment is often less satisfactory.

Recently mutations in ATM, the gene usually mutated in ataxia- telangiectasia, have
been identified in an Indian pedigree with recessive DOPA-responsive cervical dystonia (Charlesworth et al., 2013). The mechanisms whereby ATM mutations may lead to this phenotype are unknown.

SGCE-related Myoclonus-Dystonia (DYT11)

M-D clinical and genetic aspects will be discussed in detail in chapter 5.2.

ATP1A3 Mutations (DYT12)

Mutations in ATP1A3 were initially reported as the cause of rapid-onset dystonia-parkinsonism (RDP), a rare autosomal dominant condition characterized by abrupt onset in adolescents or young adults of fixed dystonic and parkinsonian symptoms (de Carvalho Aguiar et al., 2004, Brashear et al., 2007).

It was subsequently recognized that de novo ATP1A3 mutations are also responsible for alternating hemiplegia of childhood (AHC) (Heinzen et al., 2012, Rosewich et al., 2012). AHC is an infantile disorder characterized by a fluctuating course and recurrent, and alternating, hemiplegic or hemidystonic episodes, followed in most cases by the development of permanent neurological symptoms (i.e. epilepsy, neurodevelopmental delay with intellectual disability, chorea, dystonia, parkinsonism, bulbar symptoms and cerebellar ataxia) (Heinzen et al., 2014).

More recently, a single missense mutation was shown to cause another distinct autosomal dominant condition, featuring progressive cerebellar atrophy with relapsing episodes of cerebellar ataxia, areflexia, pes cavus, optic atrophy, and sensori-neural deafness (CAPOS) (Demos et al., 2014).

Importantly, a growing number of cases with overlapping RDP/AHC (Roubergue et al., 2013, Rosewich et al., 2014, Rosewich et al., 2014, Panagiotakaki et al., 2015, Termsarasab et al., 2015) and CAPOS/AHC (Rosewich et al., 2014) have been reported in the literature.

ATP1A3 encodes the catalytic unit of the sodium pump that uses ATP hydrolysis to exchange Na+ and K+ across the cell membrane to maintain ionic gradients.
**PRKRA Mutations (DYT16)**

Mutations in *PRKRA* a rare cause of autosomal recessive dystonia-parkinsonism, were initially identified in two apparently unrelated consanguineous Brazilian families featuring young onset generalised dystonia with additional parkinsonian features (Camargos *et al.*, 2008). Autozygosity mapping and subsequent mutational screening revealed the *PRKRA* homozygous missense mutation p.Pro222Leu in both families (Camargos *et al.*, 2008). Only recently three more families with bi-allelic *PRKRA* mutation have been reported (Zech *et al.*, 2014, de Carvalho Aguiar *et al.*, 2015, Quadri *et al.*, 2016), confirming the contributory role of *PRKRA* to dystonia pathogenesis. Two European families (one Italian and one Polish) carried the same homozygous missense change, while the third family carried two novel mutations, p.Cys77Ser and p.Cys213Ph, in the compound heterozygous state.

*PRKRA* encodes protein kinase, interferon-inducible double-stranded RNA-dependent activator, a protein involved in signal transduction, cell differentiation, cell proliferation, antiviral response and apoptosis (Patel *et al.*, 2000).

**The Primary Paroxysmal Dystonias**

The paroxysmal dystonias are a clinically and genetically heterogenous group of movement disorders characterized by temporally defined dyskinetic episodes with interictal neurological normality. Paroxysmal dyskinesia can be subdivided into three main clinical syndromes: paroxysmal kinesigenic dyskinesias (PKD), paroxysmal exercise-induced dyskinesias (PED), and paroxysmal non-kinesigenic dyskinesias (PNKD) (Erro *et al.*, 2014).

**Paroxysmal kinesigenic dyskinesias**

Mutations in *PRRT2*, encoding for the synaptic protein proline-rich transmembrane protein 2, were identified as the cause of PKD in 2011 (Chen *et al.*, 2011, Wang *et al.*, 2011).

Affected individuals have frequent (up to 100 times per day) and short (seconds to minutes) dyskinetic attacks, triggered by sudden movements. PKD attacks respond dramatically to anticonvulsant therapy with carbamazepine.

*PRRT2* mutations have also been described as the cause of benign familial infantile
epilepsy and of other childhood-onset movement disorders (i.e. episodic ataxia, paroxysmal torticollis of childhood), headache disorders, and intellectual disability (Ebrahimi-Fakhari et al., 2015).

Most PRRT2 pathogenic mutations are truncating and by far the most common mutation is the recurrent frameshift c.649dupC, which is present in ~80-90% of the cases described to date (Cao et al., 2012, Friedman et al., 2012, Hedera et al., 2012, Steinlein et al., 2012).

The PRRT2 protein is highly expressed in the developing nervous system and localised to the synapse (Lee et al., 2012). Truncating mutations result in a protein lacking the transmembrane domain, leading to altered subcellular localization. The subsequent reduction of PRRT2 protein may lead to altered synaptic neurotransmitter release and dysregulated neuronal excitability (Chen et al., 2011).

More recently, a single missense mutation (p.Glu1483Lys) in SCN8A, encoding a voltage gated Na-channel subunit widely expressed in the CNS, has recently been linked to paroxysmal kinesigenic dyskinesia and benign familial infantile seizures (Gardella et al., 2016). This observation expands the phenotypic spectrum associated with mutations in this gene, which also includes severe epileptic encephalopathy and a neurodevelopmental disorder (Larsen et al., 2015).

Paroxysmal non-kinesigenic dyskinesia

Paroxysmal non-kinesigenic dyskinesia (PNKD) symptoms classically starts in the first two decades of life, featuring attacks of dystonic and choreatic dyskinesias, often triggered by stress or alcohol and/or caffeine consumption (Bruno et al., 2007). PNKD episodes usually last from minutes to hours and occur with a frequency that can vary between 20 attacks per day to only few attacks per year.

PNKD is caused by heterozygous mutations in the myofibrillogenesis regulator (MR-1) gene and mutations are inherited as an autosomal dominant trait. Only three mutations have been described to date, all clustered in the N-terminus of the protein: p.Ala7Val, p.Ala9Val and p.Ala33Pro (Lee et al., 2004, Rainier et al., 2004, Chen et al., 2005, Djarmati et al., 2005, Hempelmann et al., 2006, Ghezzi et al., 2009).

PRRT2 mutations have been found in small number of patients with a PNKD
presentation (Gardiner et al., 2015).

**Paroxysmal exercise-induced dyskinesia**

Mutations in the *SLC2A1* gene, which codes for the glucose transporter GLUT1, cause paroxysmal exercise-induced dyskinesia (PED). PED is characterized by attacks of combined chorea, athetosis, and dystonia, which tend to last from a few minutes to an hour (Bruno et al., 2004). GLUT1 is the main glucose transporter in the brain and PED may be caused by reduced glucose transport into the brain, particularly when energy demand is high after prolonged exercise.


2.3 Genetics of Choreas

2.3.1 Introduction

Chorea (from the Greek word χορεία = dance) is a hyperkinetic movement disorder characterized by continuous and brief involuntary movements. Chorea can affect any body segment, including limbs, trunk, neck, face, and mouth. Typically choreic movements flow from one body part to another and are unpredictable in terms of timing, speed and direction. Different from dystonia, choreic movements are not repetitive or patterned. Chorea is usually worsened by anxiety and execution of voluntary movements, and it is absent during sleep. When chorea is accompanied by athetosis, a term defining slower and continuous writhing involuntary movements, it is referred to as choreo-athetosis. Chorea is commonly observed in combination with other hyperkinesias, such as dystonia (Donaldson et al., 2012).

Focal lesions of the striatum and degeneration and/or functional dysregulation of medium spiny neurons (MSNs), which constitute ~95% of the striatal cells and form the striatal output projections, are considered to underlie the pathophysiology of choreic movements [2].

Chorea is a major feature of several neurological disorders that can be either genetically inherited or acquired.

There are many non-inherited causes of chorea. These include focal lesions of the striatum (e.g. strokes, occupying-space lesions), infection/immune-related conditions (e.g. Sydenham’s chorea, systemic lupus erythematosus or anti-phospholipid syndrome), use of drugs such as L-DOPA, neuroleptics, antiemetics and oral contraception, and several metabolic, haematological, or endocrinological disorders (e.g. polycythaemia rubra vera, hyperthyroidism, hypo/hyperparathyroidism and hypo/hyperglycemia) (Wild and Tabrizi, 2007). Acquired causes of chorea will not be further discussed in this chapter.

Clinically, inherited choreas may be classified according to the age at onset of symptoms, as mutations in different genes tend to cause disease with onset either in childhood/juvenile years (< 20 years) or in adulthood. A further important distinction can be made between progressive (i.e. neurodegenerative) and non-progressive causes of inherited chorea.
From a clinical point of view, it is also important to establish whether chorea is the only or the predominant neurological feature or whether other accompanying features are present. This is essential for guiding genetic diagnostic testing. Other neurological features that should be looked for in patients with chorea are cognitive impairment, ataxia, epilepsy, myoclonus, dystonia, peripheral neuropathy, spasticity, and eye movement disorders.

Of relevance, a number of childhood-onset metabolic (e.g. Lesch-Nyhan Syndrome, glutaric aciduria, propionic acidemia) or mitochondrial disorders (e.g. Leigh syndrome, MELAS) can present with a movement disorder (most frequently dystonia, followed by chorea), though typically in the context of a complex multi-systemic neurologic presentation (Hermann and Walker, 2015).

The genes that have been associated with choreic syndromes are summarized in Table 2-4. Mitochondrial diseases and inborn errors of metabolism associated with a choreic presentation are not listed in the table.
Table 2-4 Monogenic causes of chorea

<table>
<thead>
<tr>
<th>Gene</th>
<th>Main associated phenotype</th>
<th>Gene product</th>
<th>Inheritance</th>
<th>Age of onset</th>
<th>Diagnostic clues</th>
</tr>
</thead>
<tbody>
<tr>
<td>HTT</td>
<td>Huntington disease</td>
<td>Huntingtonin</td>
<td>AD (CAG expansion)</td>
<td>Childhood to late adulthood</td>
<td>Cognitive decline, psychiatric disturbances</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Progressive course</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>MRI: caudate nucleus head atrophy</td>
</tr>
<tr>
<td>PRNP</td>
<td>HDL1/ Prion disease</td>
<td>Prion protein</td>
<td>AD (octapeptide coding repeat expansion)</td>
<td>Adulthood</td>
<td>Dementia, seizures and psychiatric features</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Possible parkinsonism at onset and longer survival than HD</td>
</tr>
<tr>
<td>JPH3</td>
<td>HDL2</td>
<td>Junctophilin 3</td>
<td>AD (CAG/CTG expansion)</td>
<td>Adulthood</td>
<td>Parkinsonism may be first manifestation</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>High frequency in people with black African ancestry</td>
</tr>
<tr>
<td>TBP</td>
<td>HDL4/ Spinocerebellar ataxia type 17</td>
<td>TATA box-binding protein</td>
<td>AD (CAG expansion)</td>
<td>Childhood to adulthood</td>
<td>Ataxia, epilepsy and cognitive decline</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Frequent parkinsonism</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>MRI: cerebellar atrophy</td>
</tr>
<tr>
<td>ATN1</td>
<td>Dentato-rubral-pallido-luysian atrophy</td>
<td>Atrophin-1</td>
<td>AD (CAG expansion)</td>
<td>Childhood to adulthood</td>
<td>Seizures, myoclonus and cognitive decline</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>MRI: Cerebellar and brainstem atrophy (especially pons)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>High frequency in Japan</td>
</tr>
<tr>
<td>C9orf72</td>
<td>FTD/MND</td>
<td>Chromosome 9 Open Reading Frame 72</td>
<td>AD (GGGGCC expansion)</td>
<td>Childhood to adulthood</td>
<td>Prominent cognitive and psychiatric features</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Pyramidal signs</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>MRI: diffuse cerebral atrophy</td>
</tr>
<tr>
<td>FTL</td>
<td>Neuroferritinopathy</td>
<td>Ferritin light chain</td>
<td>AD</td>
<td>Teenage to late adulthood</td>
<td>Action-specific facial dystonia</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Reduced ferritin plasma levels</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>MRI: iron deposition in basal ganglia and cortical pencil lining</td>
</tr>
<tr>
<td>VPS13A</td>
<td>Chorea-acanthocytosis</td>
<td>Chorion</td>
<td>AR</td>
<td>Early adulthood</td>
<td>Severe oromandibular dystonia with lip and tongue biting</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Head drops</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Peripheral axonal neuropathy</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Elevated serum CK</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>MRI: caudate nucleus head atrophy</td>
</tr>
<tr>
<td>XK</td>
<td>Macleod syndrome</td>
<td>Kell blood group protein</td>
<td>X-linked recessive</td>
<td>Adulthood</td>
<td>Peripheral sensorimotor neuropathy</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Cardiomyopathy</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Elevated serum CK</td>
</tr>
<tr>
<td>Gene</td>
<td>Disorder</td>
<td>Mutated Gene</td>
<td>Inheritance</td>
<td>Age of Onset</td>
<td>Associated Findings</td>
</tr>
<tr>
<td>------</td>
<td>----------</td>
<td>--------------</td>
<td>-------------</td>
<td>--------------</td>
<td>---------------------</td>
</tr>
<tr>
<td>ATM</td>
<td>Ataxia-telangiectasia</td>
<td>Ataxia-telangiectasia mutated gene</td>
<td>AR</td>
<td>Childhood to adulthood</td>
<td>Oculocutaneous telangiectases, Sensorimotor neuropathy, Elevated serum alpha-fetoprotein, Predisposition to malignancy, MRI: cerebellar atrophy</td>
</tr>
<tr>
<td>APTX SETX</td>
<td>Ataxia with oculomotor apraxia (AOA) type 1, 2</td>
<td>Aprataxin, Senataxin</td>
<td>AR</td>
<td>Childhood to adulthood</td>
<td>Sensorimotor neuropathy, Hypoalbuminemia in AOA1, Hypercholesterolemia in AOA1, Elevated alpha-fetoprotein in AOA2, MRI: cerebellar atrophy</td>
</tr>
<tr>
<td>FRDA</td>
<td>Friedreich’s ataxia</td>
<td>Frataxin</td>
<td>AR</td>
<td>Childhood to adulthood</td>
<td>Cardiac involvement, diabetes, skeletal deformities, Pyramidal signs</td>
</tr>
<tr>
<td>RNF216</td>
<td>Gordon-Holmes syndrome</td>
<td>Ring finger protein 216</td>
<td>AR</td>
<td>Adulthood</td>
<td>Behavioural problems, and severe dementia, Hypogonadism, MRI: white matter lesions and cerebellar atrophy</td>
</tr>
<tr>
<td>NKX2-1</td>
<td>NKX2-1-related chorea (Benign Hereditary Chorea)</td>
<td>Thyroid transcription factor 1</td>
<td>AD/De novo</td>
<td>Infancy</td>
<td>Non-progressive course, Hypotonia and early falls, Learning difficulties, Frequent pulmonary and thyroid involvement</td>
</tr>
</tbody>
</table>

AD: autosomal dominant; AR: autosomal recessive; BFIS: Benign familial infantile seizures; HDL: Huntington’s disease-like.
2.3.2 Huntington’s disease

HD is by far the most common cause of chorea and represents the archetype of progressive forms of chorea.

HD is a fatal neurodegenerative disorder inherited as an autosomal dominant trait and is caused by heterozygous expansion of a CAG trinucleotide repeat in the *HTT* gene (encoding the huntingtin protein). In healthy individuals the CAG repeat ranges between 10 and 29 copies on normal chromosomes, whereas is expanded to 36–121 in HD patients. CAG repeats equal or above 40 are fully penetrant, whereas expansion between 36 and 39 are associated with reduced penetrance.

Clinically, HD features a triad of relentlessly progressive motor, cognitive and psychiatric symptoms, generally with onset in adult years, and inevitably leading to death after ~15-20 years of disease (Ross and Tabrizi, 2011). Onset of disease is typically in adult years, with less than 10% of patients presenting at an age below 20 (Koutsis et al., 2013). Juvenile HD (Westphal variant) tends to be associated with little choreic movements, while rigidity, parkinsonism, dystonic features, myoclonus and cognitive impairment would dominate the phenotype (Quinn and Schrag, 1998).

Most of the current research efforts in HD genetics are aimed at identifying disease modifiers, which may influence the disease progression and determine the age at onset (AAO) of motor symptoms (Gusella et al., 2014). The length of the CAG expansion is well known to be the most relevant determinant of the age at onset, with longer repeats associated with an earlier onset (Andrew et al., 1993). However, the CAG repeat size accounts for only ~50% of the variation in age at onset (Langbehn et al., 2004) and a substantial portion of the remaining variance in age at onset is highly heritable, strongly indicating the existence of other critical genetic determining factors (Gusella et al., 2014). Neither the size of the non-expanded *HTT* allele, nor the presence of a second smaller CAG pathological expansion, is able to significantly influence age at onset (Lee et al., 2012). A recent study showed that a variant (rs13102260; G>A) in the *HTT* promoter, located in the site that regulates binding of the transcription factor NF-κB, exerts a bidirectional effect on HD age at onset (Becanovic et al., 2015). The authors showed in *vitro* and in *vivo* that the presence of the A allele determined a lower NF-κB-mediated *HTT* transcriptional activity, resulting in delayed age at onset when inherited on the same allele of the pathological expansion (reduced expression of the pathological...
allele). On the contrary, the A allele was associated with an earlier age at onset when located on the non-expanded allele (reduced expression of the normal \textit{HTT}). An important corollary of these results is that therapeutic strategies aimed at lowering the expression of the pathological CAG expansion should take into account that non allele-specific silencing of \textit{HTT} could bear undesired effects by decreasing the expression of the normal allele. The most relevant advance toward the discovery of HD genetic modifiers is the recent publication of GWAS performed by the Genetic Modifiers of Huntington Disease (GeM—HD) Consortium (Lee \textit{et al.}, 2015). The authors identified two GWAS-significant loci, one on chromosome 15 and one on chromosome 8 that significantly modified the age at onset of motor symptoms as predicted solely by the CAG expansion length. Other suggestive associations, though not passing the stringent GWAS-significance threshold, were observed on chromosomes 3, 5 and 21. Genes located on chromosome 15 locus are \textit{MTMR10} and \textit{FAN1} and on the chromosome 8 locus are \textit{RRM2B} and \textit{UBR5}. Pathway analysis of the GWAS results indicates that HD modifiers may be involved in control of DNA handling and repair mechanisms. Supporting this view, the chromosome 3 locus centred on \textit{MLH1}, a gene previously identified in a HD mouse model as a modifier of somatic instability of the CAG repeats (Pinto \textit{et al.}, 2013).

2.3.3 Huntington’s disease-like syndromes

~1\% of families with a HD-like clinical presentation do not carry pathogenic expansion in the HD gene and are therefore labelled as HD phenocopies or HD-lookalike (HDL) disorders (Schneider \textit{et al.}, 2007).

Importantly, HD phenocopies are clinically and genetically heterogeneous and a conclusive genetic diagnosis is currently possible in only a small minority of cases. In the Queen Square series, a genetic diagnosis could be achieved in only 8/285 HDL cases (2.8\%) (Wild \textit{et al.}, 2008). Other studies report even lower diagnostic yields (Stevanin \textit{et al.}, 2003, Keckarevic \textit{et al.}, 2005, Costa Mdo \textit{et al.}, 2006, Koutsis \textit{et al.}, 2012).

\textbf{HDL1}

HDL1 is an exceptionally rare autosomal dominant neurodegenerative disorder, caused by an 8-octapeptide repeats expansion in the prion protein (PrP) gene (\textit{PRNP}) (Moore \textit{et al.}, 2001).
The clinical features of carriers of the \textit{PRNP} octapeptide repeats are indistinguishable from classic HD and include chorea, personality changes, psychiatric symptoms and cognitive decline. Moreover, few members of a very large British pedigree with a \textit{PRNP} 6-octapeptide repeats expansion presented with a combination of chorea and cognitive decline, mimicking HD (Mead \textit{et al.}, 2006). In the Queen Square series only one HDL case was found to carry a \textit{PRNP} mutation, and was subsequently recognized to be a member of the large family previously reported by Mead and colleagues (Wild \textit{et al.}, 2008).

\textbf{HDL2}

HDL2 is caused by a CTG-CAG triplet repeat expansion in the junctophilin 3 (\textit{JPH3}) gene (Holmes \textit{et al.}, 2001). The expansion is virtually absent in Caucasian population (Margolis \textit{et al.}, 2004), but is common in patients of black South Africans of sub-Saharan descent, where the expansion is responsible for up to \textasciitilde 50\% of patients with a HDL presentation (Stevanin \textit{et al.}, 2003, Krause \textit{et al.}, 2015). Abnormally expanded repeat expansions range from 40 to 59 triplets and the penetrance of the expansion is high (Govert and Schneider, 2013). Most HDL2 patients are clinically indistinguishable from HD cases (Margolis \textit{et al.}, 2001). However, Schneider and colleagues reported a large African-American kindred carrying a \textit{JPH3} expansion and presenting with a progressive akinetic-rigid syndrome and severe dementia but very little chorea. This indicates that \textit{JPH3} should be considered in the differential diagnosis of early-onset dementia and hypokinetic-rigid syndromes in individuals of African descent (Schneider \textit{et al.}, 2012)

\textbf{HDL3}

The HDL3 locus was mapped to 4p15.3 in a single consanguineous pedigree from Saudi Arabia with a recessive HDL presentation (Kambouri \textit{et al.}, 2000). However no other pedigrees mapping to this locus have been subsequently identified and the causative gene has not been reported as yet.

\textbf{HDL4}

HDL4 is caused by heterozygous CAG triplet repeat expansions in the TATA box-binding protein (TBP) gene located on chromosome 6q27 (Stevanin and Brice, 2008). The same triplet expansion is also responsible for SCA type 17 (SCA17) (Stevanin \textit{et al.}, 2003). SCA17 expansions account for approximately 1\% of the HDL cases, which
makes it one of the most important genetic mutations found in HDL Caucasian subjects (Wild and Tabrizi, 2007). However, chorea is present only in 20% of SCA17 patients, with ataxia, seizures, dementia and parkinsonism being the most frequently detected features (Zuhlke and Burk, 2007). Interestingly, within given SCA17 positive pedigrees, homogenous HDL presentation can be observed in multiple individuals, suggesting that other co-inherited genetic factors may influence the presentation (HDL vs ataxic) (Schneider et al., 2006).

Global cortical atrophy, putaminal enhancement or cerebellar atrophy can be inconsistently found in SCA17-associated HDL. Importantly, other SCA expansions, in particular those associated with SCA1, SCA2, and SCA3, can cause hyperkinetic movement disorders that can sometimes include chorea (Schols et al., 2000, van Gaalen et al., 2011).

**Dentato-rubral-pallido-luysian atrophy**

Dentato-rubral-pallido-luysian atrophy (DRPLA) is an autosomal dominant neurodegenerative disease. DRPLA is caused by CAG triplet repeat expansions in the gene encoding atrophin-1 (Nagafuchi et al., 1994). DRPLA is particularly frequent in Japan, where it is almost as frequent as HD, but is exceptionally rare in other ethnicities. Clinically, DRPLA presents with a variable combination of chorea, ataxia, cognitive decline, and, especially in young individuals with larger expansions, progressive myoclonic epilepsy (Becher et al., 1997).

**C9orf72-Related HDL**

A hexanucleotide repeat expansion in a non-coding portion of the C9orf72 gene was recently shown to be the most important cause of familial and sporadic frontotemporal lobar degeneration and amyotrophic lateral sclerosis (DeJesus-Hernandez et al., 2011, Renton et al., 2011). Hensman Moss and colleagues recently screened a cohort of 514 HDL patients for the presence of the C9orf72 expansion and 10 subjects (1.95%) carrying the expansion, making this mutation the most common identified cause of HDL syndromes in a UK cohort (Hensman Moss et al., 2014). Not surprisingly, a movement disorder was a prominent feature in these cases, with 3 exhibiting chorea, 4 dystonia, 4 myoclonus, and 3 tremor. Furthermore, 6 cases had rigidity and 5 bradykinesia. Although, none of the cases had been formally diagnosed with fronto-temporal dementia, behavioural, psychiatric and cognitive difficulties were observed in most of
the carriers. Upper motor neuron signs were present in 4 of the 10 subjects, while lower motor neuron signs were not detected.

Other groups have been subsequently confirmed C9orf72 expansions to be a relevant cause of HDL in other populations (Kostic et al., 2014, Koutsis et al., 2015).

Chorea-Acanthocytosis

Bi-allelic mutations in VPS13A, coding for the large protein chorein, are responsible for the neurodegenerative condition chorea-acanthocytosis (CA) (Rampoldi et al., 2001). CA is mainly characterized by orofacial and cervical dyskinesias (chorea, dystonia and tics) combined with abnormal erythrocyte morphology (acanthocytosis) that can be detected with blood film microscopic examination. Highly classical neurological signs are tongue protrusion, biting of the tongue and lips and neck and trunk flexion drops and extension spasms (Schneider et al., 2010). Progressive cognitive decline and psychiatric features are often prominent as in HD. Seizures, ataxia, and a peripheral neuropathy with distal amyotrophy and increased serum creatine kinase are also common features that should prompt a diagnosis of CA (Schneider et al., 2007).

Macleod Syndrome

Macleod syndrome is a recessive X-linked neuro-acanthocytosis syndrome caused by mutations in XK encoding Kell antigen (Wiethoff et al., 2014). Clinical presentation is similar to CA, though often with a later onset and milder and longer course. Dystonia tends to be less pronounced than in CA and cardiomyopathy and haemolytic anaemia may be relevant distinctive features (Walker et al., 2007).

Neurodegeneration with Brain Iron Accumulation

Neurodegeneration with brain iron accumulation (NBIA) is a group of inherited neurological disorders in which iron accumulates in the basal ganglia, which classically present with movement disorders and cognitive decline (Kara et al., 2013). Brain magnetic resonance imaging (MRI) usually shows the presence of iron deposition in the basal ganglia allowing the diagnosis. Most of the NBIA syndromes are recessive and start early in life and therefore are easily distinguishable from HD.

However, mutations in the FTL1 gene, encoding ferritin light chain, are an exception to this rule as they cause a late-onset dominant disease, called neuroferritinopathy, which may be often misdiagnosed as HD (Curtis et al., 2001). So far, only a small number of
FTL1 mutations have been reported and the frameshift insertion c.460InsA by far the most common in UK cohorts.

Neuroferritinopathy typically presents with chorea and/or dystonia. Other clinical features are variably reported, including parkinsonism, cerebellar signs, dysarthria, frontal lobe syndrome and dementia. An action-specific facial dystonia is reported in ~65% of the patients (Chinnery et al., 2007). Our group has recently shown that cortical pencil lining, which reflecting excessive cortical iron deposition, is a useful radiological sign of neuroferritinopathy (Batla et al., 2015).

Other recessive NBIA (in particular aceruloplasminemia) should be considered too, although chorea alone is rarely a sole prominent feature of these conditions (Walker, 2016).

**Chorea in Patients with Mutations in Cerebellar Ataxia-Related Genes**

Hyperkinetic movement disorders, including both chorea and dystonia, are frequently observed in a range of genetic neurodegenerative disorders usually presenting with predominant cerebellar ataxia (other than the aforementioned SCA17 and DRPLA).

Ataxia-telangiectasia (A-T) is an autosomal recessive multi-system disorder caused by mutation in the ataxia-telangiectasia mutated gene (ATM). A-T is usually characterized by progressive cerebellar ataxia, oculocutaneous telangiectasia, oculomotor apraxia, immunodeficiency with recurrent lung infections, and increased risk for cancer. Patients with bi-allelic ATM mutations, the cause of ataxia-telangiectasia (A-T), may present with a broad spectrum of movement disorders, including chorea (Klein et al., 1996, Hardy and Singleton, 2009, Worth et al., 2013), isolated dystonia (Saunders-Pullman et al., 2012, Claes et al., 2013), DOPA-responsive dystonia (Charlesworth et al., 2013), and myoclonus-dystonia (Cheng et al., 2002, Cummins et al., 2013, Termsarasab et al., 2015). Meneret and colleagues recently reported a total of 14 consecutive adults with variant A-T and compared them to patients with typical A-T (Meneret et al., 2014). Patients with variant A-T have milder mutations, which allow a degree of residual protein activity (Gilad et al., 1998), and clinically present with a later age at onset, a milder disease course and longer survival. Dystonia and subcortical myoclonus were the most common movement disorders observed (84%), followed by tremor (48%) (Meneret et al., 2014). Of relevance, patients with ATM-related chorea and dystonia may completely lack the classic clinical features of A-T (Kuhm et al., 2015).
Friedrich ataxia (FA) is the most common form of recessively inherited ataxia. FA is caused by mutations in the frataxin gene; 98% of cases have homozygous expansions of a GAA trinucleotide in intron 1 of the frataxin gene (Parkinson et al., 2013). Rarely, chorea can be observed in bi-allelic carriers of frataxin mutations, sometimes preceding the onset of frank cerebellar symptoms (Hanna et al., 1998, Zhu et al., 2002, Wild et al., 2008). Chorea has been rarely described also in cases with ataxia with oculomotor apraxia type 1 and 2, caused respectively by mutation in the genes encoding aprataxin and senataxin (Salvatore et al., 2008, Anheim et al., 2009).

Recessive mutations in RNF216 have recently been found in families with hypogonadotropic hypogonadism and cerebellar ataxia (Margolin et al., 2013). Expanding the phenotype of RNF216 mutations-associated diseases, Santens and colleagues identified bi-allelic mutations in the RNF216 in 2 recessive pedigrees with a HDL presentation, including chorea, behavioural problems, and severe dementia. In the affected cases brain imaging showed white matter lesions and subclinical evidence of hypogonadotropic hypogonadism was detected (Santens et al., 2015).

### 2.3.4 Benign Hereditary Chorea

Clinical and genetic aspects of benign hereditary chorea will be discussed in detail in chapter 6.2.
Chapter 3. Materials and Methods

3.1 DNA Extraction

3.1.1 Extraction from Whole Blood
Genomic DNA was extracted from whole blood (taken in ethylenediaminetetraacetic acid bottles) in the Diagnostic Genetics Laboratory in the UCL Institute of Neurology using a FlexiGene® kit (Qiagen), according to manufacturer instructions.

In brief, 300μl of whole blood is mixed with 750μl of buffer (FG1), which lyses the cells. The sample is centrifuged for 20 seconds at 10,000xg and then the supernatant was discarded. 150μl of Buffer (FG2 – which contains a protease) is added to the pellet and the tube vortexed until the pellet was homogenized. The sample is subsequently centrifuged for 5 seconds and then placed for 5 minutes in a heating block at 65°C. 150μl of isopropanolol (100%) is then added and the tube inverted several times, allowing the DNA to precipitate. The sample is centrifuged for 3 minutes at 10,000xg and then the supernatant was discarded. 150μl of 70% ethanol is added and the sample vortexed for 5 seconds and centrifuged for 3 minutes at 10,000xg. The supernatant is discarded and the pellet air dried. 200μl of Buffer (FG3) is added and the sample vortexed at 5 seconds. The DNA is then dissolved by heating the sample at 65°C in a heating block.

3.1.2 Extraction from Saliva
Genomic DNA was extracted from saliva samples using the Oragene© DNA kit according to the manufacturers’ instructions.

Briefly, the saliva sample and Oragene solution are incubated at 50°C in an air incubator for 2 hours. Then, 500μl of the solution is transferred to a 1.5mL microcentrifuge tube and 20μl of Oragene DNA purifier is added and then vortexed for a few seconds. The sample is then incubated on ice for 10 minutes and subsequently centrifuged at room temperature for 15 minutes at 13,000 rpm. The clear supernatant is transferred into a fresh microcentrifuge tube and the pellet containing impurities discarded. 500μl of 100% ethanol is added to the solution and inverted 10 times. The sample is then left at room temperature for 10 minutes allowing the DNA to precipitate. The sample was then
centrifuged at room temperature for 2 minutes at 13,000 rpm and the supernatant is removed and discarded. 250μl of 70% ethanol was added. The sample is allowed to stand at room temperature for 1 minute. The ethanol is then removed, taking care not to disturb the pellet. 100μl of DNA buffer was then added to dissolve the DNA pellet.

After extraction, the concentration of DNA samples was then measured using the Nanodrop (Thermo Scientific).

### 3.2 Polymerase-Chain Reaction

PCR primers were designed using Primer3 (http://primer3.ut.ee/). The nucleotide sequence form each exon, along with flanking ~200 downstream and upstream intronic base pairs, was inputted in the program. DNA sequence of each exon was obtained from a suitable transcript (generally the longest coding transcript available with a consensus CDS (CCDS) identifier) downloaded from UCSC Genome Browser (https://genome.ucsc.edu/).

PCR was performed using 10 μl Faststart PCR Master (Roche Applied Sciences), 2 μl of forward and reverse primers at a dilution of 5 pmol/μl, 5 μl of double distilled water and genomic DNA at a final concentration of 20-30ng. DMSO was added to the mix for amplification of amplicons GC-rich. Each PCR reaction was performed using a program that had been demonstrated to optimally amplify the product.

### 3.3 Agarose Gel Electrophoresis

Agarose gel electrophoresis was used to check whether the PCR reaction was successful and verify the quality and size of PCR amplification product. 10x TBE (tris-borate-EDTA solution) solution was prepared using 121.1g of Trizma base (Sigma), 61.8g Boric Acid (Sigma), 7.4g of Ethylenediaminetetraacetic acid (EDTA) (Sigma) and dissolved in 1 litre of distilled water. A 1.5% gel was prepared using Ultrapure Agarose (Invitrogen) and TBE 1X buffer and stained with gel red (Cambridge bioscience). 3μl of PCR product and 3μl of X6 Orange DNA loading dye (Thermo scientific) were mixed together and loaded into one well of the gel. A DNA ladder (Midrange 100-2000 bp) (Qiagen) was run alongside each lane of the gel and then the samples were run at 100 mV for 45 minutes. DNA fragments are visualized using UV light.
3.4 PCR Purification

PCR cleanup was performed using an enzymatic method. The enzymatic cleanup solution was prepared mixing 200μl of Fast-Alkaline phosphatase (Thermo scientific), which removes unused dNTPs, 50μl Exonuclease I (Thermo scientific), which removes single strand DNA from PCR products, and 750μl of purified water. 5μl of PCR product was added to 2μl of the enzymatic solution and then run on a thermal cycler at 37°C for 30 minutes followed by 80°C for 15 minutes.

3.5 Sanger Sequencing

The sequencing Reaction mix was prepared combining 0.5μl BigDye Terminator v3.1 (Applied Biosystems), 2μl 5x Sequencing Buffer (Applied Biosystems), 1μl forward or reverse primers at 5 pmol/μl and 3μl of purified PCR product and 2.15μl of PCR grade water. The standard thermo-cycler program recommended by Applied Biosystems was used to run the Sequencing reaction. Sequencing purification was performed through the Sephadex method. A hydrated solution of Sephadex was prepared mixing 40ml of autoclaved water and 2.9 grams of Sephadex G-50 powder (Sigma-Aldrich). It was then vigorously mixed and allowed to hydrate, resting for at least 30 minutes at room temperature. Next, 350μl of the Sephadex solution was placed into each well of a Corning FiltrEXTM 96 well filter plates (0.66 mm glass fibre filter). The Corning FiltrEXTM filter plate was placed on top of an empty collection plate and centrifuged for 3 minutes at 750xg. The Corning FiltrEXTM was then placed on top of a new PCR plate. The entire sequencing reaction volume was then pipetted onto the Sephadex columns and the plates were then centrifuged again for 5 minutes at 910xg.

Sanger sequencing reaction products were then read on a 3730 DNA analyser (Applied Biosystems, Foster City, CA, USA) and electropherograms were visualised using Sequencher software (Gene Codes Corporation, MI, USA, version 5.0.1).

3.6 Multiplex Ligation-Dependent Probe Amplification

Multiplex Ligation-Dependent Probe Amplification (MLPA) analysis is a multiplex PCR method used to determine exon copy number of genes of interest.

In this thesis I used the MLPA probemix P051-D1 (MRC-Holland, Amsterdam, the
Netherlands), which screens for the presence of exon deletions and duplications in SNCA, PARK2, PINK1, PARK7 (Djarmati et al., 2007). In addition, it contains probes that will only generate a signal when the SNCA p.Ala30Pro and LRRK2 p.Gly2019Ser mutation are present.

The MLPA reaction can be divided in five major steps (Figure 3-1): 1) DNA denaturation and hybridisation of MLPA probes. During this step, the DNA of the probands is first denatured at 98°C and then incubated overnight with a mixture of MLPA probes and MLPA buffer. MLPA probes consist of two separate oligonucleotides, each containing one of the PCR primer sequences. 2) Ligation reaction; this reaction can occur only if the two probe oligonucleotides are both hybridised to their adjacent targets. 3) PCR reaction; because only ligated probes will be exponentially amplified during this reaction, the number of probe ligation products is a measure for the number of target sequences in the sample. 4) Separation of PCR amplification products, after combining it with a mix of Hi-Di Formamide (Thermo Fisher Scientific) and marker Liz500 (Thermo Fisher Scientific) by electrophoresis on the 3730 DNA analyser. 5) Data
analysis; the results of the analysis are checked and interpreted using Genemarker software (Softgenetics).

### 3.7 Analysis of Microsatellites

Microsatellites are DNA sequences located in non-coding genomic regions and formed by mono, di, tri or tetra-nucleotides repeated in multiple tandem copies. Microsatellites are abundant and dispersed through the genome. The number of repeats is highly variable amongst different individuals in a population, whereas they tend to be stable between generations and are transmitted in a Mendelian fashion. Microsatellites are therefore widely used as genetic markers in genetic familial study. In this thesis microsatellites were used to define and compare the haplotype surrounding specific mutations in different families. Microsatellites were selected in UCSC genome browser. Primers to amplify the regions containing the microsatellite of interest were designed using Primer3. One primer for each reaction was labelled with a fluorescent dye (6-FAM). PCR reactions were performed as previously described. 0.5μl of undiluted PCR product was mixed with 12 μl of Hi-Di Formamide and 0.3μl of Liz500, denatured at 98°C for 3 minutes and analysed by electrophoresis on the 3730 DNA analyser. The size of the microsatellites was assessed using the GeneMapper software (Thermo Fisher Scientific).

### 3.8 DNA Array SNP Analysis

Two types of DNA array SNP beadchips were used. For whole genome linkage analysis, I used Human CytoSNP beadchips (Illumina), which contains probes for ~220,000 markers. For detection of copy number variants or small chromosomal rearrangements, I used HumanOmni2.5-8 BeadChips (Illumina), which contains probes for ~2,500,000 markers.

For each sample, 10 μl of DNA at a concentration of ~75 ng/μl were prepared. Samples were processed, hybridized and scanned in accordance with the manufacturers instructions at UCL Genomics. Clustering, normalization and genotype calls were performed using Genome Studio 2010.3 Genotyping module (Illumina).

A Genome Studio project was created for each family project with default parameters. CNVs analysis was performed using the plugin CNV partition 2.4.4 und results were...
applied as bookmarks in genome viewer. PLINK output files, to be used for whole-genome linkage analysis, were created in Genome Studio through the PLINK report plug-in.

### 3.9 Genome-Wide Parametric Multipoint Linkage Analysis

Quality control checks were performed through PLINK. SNPs with low call rates (<90%) and monomorphic SNPs (MAF <0.5%) were removed using the commands — maf 0.05 —geno 0.1. ~6000 single nucleotide polymorphisms (SNP) equally distributed across the genome were then randomly selected using the command —thin.

Linkage analysis was performed using the software MERLIN ([http://www.sph.umich.edu/csg/abecasis/merlin/tour/linkage.html](http://www.sph.umich.edu/csg/abecasis/merlin/tour/linkage.html)) (Abecasis et al., 2002).

A .ped, .dat, .map and .model files were generated following the instructions. The .ped file contains information about the affection status of each individual included in the analysis (specified with a number: 1=unaffected and 2=affected 0=unknown), the sex of the individual (1=male, 2=female). Furthermore, the .ped file contains information about the pedigree structure, indicated by assigning every individual a unique number and specifying for each subject the identifiers of his/her father and mother. The .map files contains the list of SNP markers used for the linkage analysis and their genetic position expressed in centimorgans (cM). The model file contains information regarding the assumed mode of inheritance (autosomal dominant vs recessive) together with the estimated disease allele frequency and the probability of being affected for individuals with 0, 1 and 2 copies of the disease allele (penetrances).

The PEDSTATS function ([http://csg.sph.umich.edu/abecasis/PedStats/](http://csg.sph.umich.edu/abecasis/PedStats/)) was used to validate the input formats and pedigree consistency using the command:

```
pedstats –d file.dat –p file.ped
```

To identify Mendelian errors (genotyping errors), which can lead to misleading inferences about gene flow in pedigrees and reduce the power of pedigree analysis, I used the command:

```
merlin -d file.dat -p file.ped -m file.map –error
```
Then, to erase from the input files the genotypes flagged as problematic I run the command:

```
pedwipe -d error.dat -p error.ped
```

MERLIN was then used to perform genome wide parametric linkage analysis using the command:

```
merlin -d file.dat -p file.ped -m file.map -model file.model -markernames --pdf --tabulate
```

This produces an output file containing LOD scores for each marker and a PDF graphical representation of the results. At least three randomly generated sets of markers were generated to perform linkage analysis.

Haplotype analysis allows to reconstructing the haplotype of a region of interest and checking segregation in the pedigree of the disease-associated haplotypes with disease status. Furthermore, it allows to possibly identifying recombination event, which could further narrow down the genomic intervals (fine mapping).

Fine mapping and haplotype reconstruction were performed for chromosomal loci identified by linkage analysis with LOD scores >0. First, for each locus, a list of SNPs spanning the linked region was compiled and this was used to generate a new of input files (.ped, .map, and .dat) using PLINK.

Haplotype analysis was then performed using the MERLIN command:

```
merlin -d haplo.dat -p haplo.ped -m haplo.map --best --horizontal
```

### 3.10 Whole-Exome Sequencing

WES was performed using the next-generation sequencing platform at Institute of Neurology. WES experiments were performed by Deborah Hughes and preliminary bioinformatics analysis by Dr Alan Pittman. The Truseq Exome Enrichment Kit (Illumina) was used for WES and run on an Illumina Hiseq 2000. This kit targets 64Mb of genomic sequence, including exons, exons-introns boundaries and UTRs and micro-RNAs. This kit offers on average more than 90% coverage of the exons, as defined in the latest version of the CCDS and Refseq database.
The first of WES is the library preparation (Figure 3-2). In brief, 3μg of genomic DNA must be sheared to form fragments, either by mechanical (Covaris) or enzymatic methods. This was followed by blunting of ragged ends of the DNA fragments and binding of the sequencing adapters. The latter contains unique indexes, which are unique for each sample and are essential in the next steps of the data analysis. After a series of PCR amplification, purification steps (to eliminate leftover adapters, enzymes and primers) and normalization (to ensure equal coverage of each sample in the total data), the libraries are ready for the enrichments stage.

![Figure 3-2 Scheme of the library preparation using the Truseq Illumina kit](http://tucf-genomics.tufts.edu/home/faq)

The first step of enrichment is hybridization, during which the DNA fragments are first denatured to single strands and then bound to biotin-labelled probes specific for the target exome region. This is followed by a series of capture and washing steps through a magnetic beads system and then the library is amplified through a final PCR step. After a last washing step, the libraries are ready for cluster generation, which is completed.
within the Illumina cBOT. To form the clusters, the single-stranded, single-molecule template are primed and extended and subsequently the immobilized templates undergo bridge amplification with immediately adjacent primers (Figure 3-3).

A flow cell consists of eight lanes and is covered with a lawn of specific oligos to capture libraries as they are passed through the flow cell. Multiple samples can be pooled together and loaded in the same lane. The more libraries are pooled together the poorer will be the coverage.

Following cluster generation and bridge amplification, the flow cell is loaded into the HiSeq 2000 for sequencing and imaging. After sequencing and imaging are terminated, the individual sample reads are de-multiplexed based on their unique indexes (Casava 1.8.2).

### 3.11 Bioinformatic Analysis of Whole-Exome Sequencing

In order to obtain a final list of variants the WES raw data was run and analysed through a bioinformatics pipeline, constructed by Dr. Vincent Plagnol and optimized by Dr Alan Pittman. The pipeline is schematically represented in Figure 3-4.
The WES data generated by the HiSeq 2000 is stored in a FASTQ file, a text-based format for storing nucleotide sequence.

Paired-end sequence reads are aligned with Novoalign (www.novocraft.com) against the reference human genome (UCSC hg19). Generation of duplicate DNA sequencing reads, which are due to PCR-induced duplication, is one of the most relevant technical artifacts of the capture-sequencing procedure. Duplicate read needs to be removed so that they do not influence variant calling and this step is performed with Picard software package (http://picard.sourceforge.net/).

Figure 3-4. Bioinformatic pipeline for WES analysis.

Information regarding the project, the genome and an alignment section, which includes mapping position and information on mismatches, are contained into tab-delimited text files, called SAM files (.sam). A BAM file (.bam) is the binary version of a SAM file. BAM files, rather than SAM, are the recommended format for visualisation software.

After the alignment is completed, variant calling is performed comparing at each genomic position the aligned SAM/BAM file to the reference genome. SAMTOOLS package (http://samtools.sourceforge.net/) is used to recalibrate base quality scores, perform local realignments around possible indels, and to call and filter the variants.
The outcome of the analysis is a Variant Call Format (VCF) file, which holds the list of identified genetic variants and the associated quality scores that take into account both base calling and mapping scores.

ANNOVAR software is used to perform functional annotation the variants (Wang et al., 2010). Annotated variant files are generated including a comparison to several publicly available reference databases of genetic variation, including the Exome Aggregation Consortium (ExAC; http://exac.broadinstitute.org/), dbSNP version 129 and 137 (www.ncbi.nlm.nih.gov/projects/SNP), 1000 Genomes project (1kGP; www.1000genomes.org), NHLBI Exome Variant Server (EVS; evs.gs.washington.edu) and Complete Genomics 69 (www.completegenomics.com/public-data/69-Genomes).

In silico prediction of pathogenicity is assessed using SIFT (http://sift.jcvi.org/) (Kumar et al., 2009), PolyPhen2 (http://genetics.bwh.harvard.edu/pph2/) (Adzhubei et al., 2010), Provean (http://provean.jcvi.org/) (Choi et al., 2012), MutationTaster (http://www.mutationtaster.org/) (Schwarz et al., 2014) and CADD (http://cadd.gs.washington.edu/home) (Kircher et al., 2014). Evolutionary conservation of nucleotides involved by variants is score using PhiloP (Pollard et al., 2010) and GERP++ (Davydov et al., 2010).

Copy number variants (CNV) were called in WES data using the Exome depth algorithm (Plagnol et al., 2012).

### 3.12 Regional Gene Expression Profiling in Brain Tissue

To add further support to the causal role of newly identified gene variants, regional gene expression profile across various brain regions in humans was assessed using an in-house and other publicly available datasets. The in-house dataset originated from the work of Dr Mina Ryten and Dr Daniah Trabzuni. This dataset was generated using Affymetrix Exon 1.0 ST Arrays and brain tissue originating from 134 control individuals, collected by the Medical Research Council (MRC) Sudden Death Brain and Tissue Bank, Edinburgh, UK, and the Sun Health Research Institute (SHRI), an affiliate of Sun Health Corporation, USA. Ten brain regions were analyzed: putamen (PUTM), frontal cortex (FCTX), temporal cortex (TCTX), hippocampus (HIPP), substantia nigra (SNIG), medulla (specifically inferior olivary nucleus, MEDU), intralobular white matter (WHMT), thalamus (THAL), and cerebellar cortex (CRBL). A full description of
the samples used and the methods of RNA isolation and processing can be found in (Trabzuni et al., 2011).

As described therein, all arrays were preprocessed using Robust Multi-array Average (RMA) quantile normalisation with GC background correction and log2 transformation in Partek’s Genomics Suite v6.6 (Partek Incorporated, USA). Regional differences in gene-level expression were investigated using Partek's mixed-model ANOVA with gender and batch effects (date of hybridization and brain bank) included as co-factors.

RNA expression changes during the course of human brain development were assessed accessing the data available through the Human Brain Transcriptome (HBT) database (Johnson et al., 2009, Kang et al., 2011). The brain regions analyzed are the striatum (STR), amygdala (AMY), neocortex (NCX), hippocampus (HIP), mediodorsal nucleus of the thalamus (MD), and cerebellar cortex (CBC). The data is publicly available accessing the HBT website (http://hbatlas.org/).
Chapter 4. Exploring the Genetics of Parkinson Disease

4.1 Outline Of The Chapter

This chapter describes the work I conducted to expand our knowledge of the genetic basis of PD.

The specific aims of the study were:

1) To assess the contribution of genetic variation in GBA in the Queen Square cohort of early-onset PD

2) To characterise genetically the Queen square cohort of PD patients who underwent treatment with DBS and assess whether the genetic status influence the outcome of the treatment and disease progression;

3) To assess the relationship between mutations in GCH1, the gene responsible for the childhood-onset disease DOPA-responsive dystonia, and PD.

4.2 GBA Mutational Analysis In Early Onset Parkinson Disease

4.2.1 Statement Of Contribution

I participated to the design of this study and ascertained the cases to be included in this study. I performed the GBA sanger-sequencing analysis in the PD cases. Dr Raquel Duran performed the GBA sanger-sequencing analysis in the Gaucher disease (GD) cohort and in control samples.

4.2.2 Background

Patients with EOPD, in comparison to patient with late-onset PD (LOPD) tend to progress more slowly in terms of motor features and generally have a longer disease course, often with preservation of cognitive function. On the other hand, L-DOPA induced complications, including motor fluctuations and dyskinesias, tend to appear earlier and be more severe (Schrag and Schott, 2006).
As previously discussed in this thesis, mutations in several genes have been associated with EOPD, in particular autosomal recessive genes, including PARK2, PINK1 and PARK7 (Bonifati, 2012). PARK2 is by far the most commonly mutated gene in EOPD, whereas mutations in PINK1 and PARK7 are significantly more rare (Abou-Sleiman et al., 2003, Healy et al., 2004, Alcalay et al., 2010).

PARK2 mutations are particularly common (frequency ~50%) in familial cases and in patients with very early onset (e.g. onset at age < 30), whereas their frequency drops (frequency ~10-15%) in sporadic cases and patients with later onset (e.g. onset at age > 30 and < 50) (Lucking et al., 2000).

The different clinical course of EOPD, together with pathological reports showing in most PARK2 mutations carriers the absence of the characteristic LBs (Doherty et al., 2013), suggest that EOPD may represent a distinct clinical and pathological entity.

Loss-of-function mutations in the acid β glucocerebrosidase gene (GBA), which are responsible for the autosomal recessive lysosomal disorder GD (Grabowski, 2008), represent the most prevalent genetic risk factor for PD identified to date (Sidransky et al., 2009). So far genetic analysis has suggested that all GBA mutations causing GD can equally predispose to PD, the assumption being that variants that do not cause GD in the homozygous would not predispose to PD.

In this study, a complete sequence analysis of the open reading frame of GBA was performed in a cohort of 185 PD cases with early onset PD (EOPD; defined by an age at onset of ≤50 years). These data were compared with those obtained from 283 controls and 73 patients with type 1 GD from the same geographic region (Duran et al., 2012).

This study had two main aims; first, to assess and quantify the contribution of GBA mutations in EOPD; second, to compare the spectrum of GBA mutations found in PD cases versus that of GD cases. Indeed, possible differences in the type of mutations found in the two disease groups may shed light onto the underlying pathogenic mechanisms responsible for PD and GD.

4.2.3 Subjects, Materials and Methods

Subjects
The cohort included in this study was formed by 185 unrelated PD cases with an onset age ≤ 50 who attended the Movement Disorders clinics at the National Hospital Queen
Square. They were all diagnosed with PD following the UK Brain Bank Clinical Criteria, presenting with at least 2 of 3 cardinal signs of tremor, rigidity, and bradykinesia, as well as a positive response to L-DOPA therapy (Hughes et al., 1992). The age-at-onset was on average $40 \pm 7.2$ years (range 16 – 50). A positive family history compatible with PD in at least a first or second-degree relative was reported in 27.5% of patients (51/185) reported.

Controls were age-matched subjects. This series of controls has been previously described (Neumann et al., 2009) with the addition of twenty-six age-matched spouse controls. Controls were neurologically intact. All cases and controls were of UK Caucasian origin with no reported Jewish Ashkenazim background. GD patients were 73 unrelated subjects, affected with the type 1 form of the disease (non-neuronopathic), attending the National GD clinics at the Royal Free and Addenbrookes Hospitals. 16 were Ashkenazi Jewish (22%) and 4 had Eastern European ancestry (5.7%); the remainder were white UK citizens without Ashkenazi ancestry. The results of the GBA mutational screening of this cohort have been published independently (Duran et al., 2012). Written informed consent was taken from each participant.
**Genetic Analysis**

The open reading frame (11 exons) of the *GBA* gene (RefSeq accession number NM_000157.3) was entirely sequenced in cases and controls. For PCR amplification of *GBA*, three different long-range PCR reactions were performed. To avoid amplification of the pseudogene, primers were designed to bind to DNA regions exclusively present within *GBA*. Primer sequences are listed in the appendix. Allele names refer to the processed protein, excluding the 39-residue signal peptide.

**Statistical Analysis**

Frequencies of coding and splice-site *GBA* variants in cases and controls were compared by means of Fisher’s exact (statistical significance set at P value < 0.05 using a two-tailed test) and odds ratios (OR) and 95% confidence intervals (CI) were calculated. Analyses were performed using the statistical analysis program R (http://www.r-project.org/).

4.2.4 Results

The results of the genetic analysis are summarised in Table 4-1. In this population of early-onset PD, a surprisingly high proportion of cases carried a *GBA* variant (25.94%; 48/185 individuals; odds ratio for carriers of any variant in cases versus controls: 7.9, 95% confidence interval 4.1 - 15.4, P < 0.0001). 14 of the *GBA* mutation carriers (29.7%) had a positive family history for PD, though we could not check segregation of the mutations in other affected family members due to unavailability of genetic material.

Among the *GBA* mutation carriers, 22 cases had a known GD-causing pathogenic mutation (20 carried one mutation and two each carried two mutations), a proportion significantly higher than in a late-onset PD cohort from the same population reported by our group (11.9%; 22/185 individuals versus 3.8%; 30/790, p = 0.0001) (Neumann et al., 2009).

The most frequent GD-causing mutations found in the PD cohort was the N370S (p.Asn409Ser) that was observed in 5 cases, followed by R463C (p.Arg502Cys) and the recombinant allele RecNciI (L444P [p.Leu483Pro] + A456P [p.Ala495Pro] + V460V [p.Val499Val]), which were each present in 3 cases, and by L444P (p.Leu483Pro) and R131C (p.Arg170Cys), each found in 2 cases. All the other GD-causing mutations were each found in single individuals. In the control group the N370S (p.Asn409Ser) and the
R257Q (p.Arg314Gln), each found in 1 subject, were the only GD-causing mutations detected.

Of relevance, one PD case carried the N370S (p.Asn409Ser) mutation in the homozygous state and, although this genotype is known to cause GD, he did not display any feature of GD. This is not entirely surprising, as the clinical spectrum in GD is wide and there are studies suggesting that patients homozygous for mild mutations, such as N370S, may actually never develop the disease (Cormand et al., 1997). Furthermore, individuals with two pathogenic GBA variants, presenting with typical PD but showing no signs of GD, have already been reported in the literature (Lesage et al., 2011).

The second important finding is that the single most common mutation in the PD group was E326K (p.Glu365Lys), which was detected at a frequency of 7.57% in cases (14/185 individuals; 12 were heterozygote, including one in combination with L444P and 2 homozygote) versus 2.47% in controls (7/283 individuals). To confirm this observation, a further 202 controls were screened for this variant and the frequency in this new cohort was 2.97 % (6/202). Taking into account both sets, the total frequency in controls was 2.68% (13/485 individuals), indicating a significant association of this variant with PD (p = 0.0059, odds ratio 2.97, 95% confidence interval 1.3-6.4).

However, despite the high frequency in both PD and control groups, E326K (p.Glu365Lys) was not detected at all in the GD cohort (Duran et al., 2012). Consistent with previous reports (Park et al., 2002), this finding clearly indicates that this variant does not cause GD. Further supporting this conclusion, the two PD cases homozygote for the E326K did not have any of the clinical or haematological features of GD and their white blood cell GCase enzymatic activity was at the lower limit of normal range. Interestingly, a number of other variants, not present in GD group and never associated with GD, were also detected in the PD cohort. Of these E388K (p.Glu427Lys), G113A (p.Gly152Ala), T369M (p.Thr408Met), S465P (p.Ser504Pro) have been previously reported in both PD cases or controls, whereas L(-14)V (p.Leu25Val), V172L (p.Val211Leu), S177T (p.Ser216Thr), L217P (p.Leu256Pro), L354P (p.Leu393Pro), V375G (p.Val414Gly), c.1506-4 C>T, c.1506-12 C>T are novel. Collectively, these variants of undetermined clinical significance were more frequent in cases than in controls (OR 7.0 - confidence interval 1.8-31.6; p = 0.001).
Table 4-1 Proportion of cases and controls with GBA mutations

<table>
<thead>
<tr>
<th>GBA variants</th>
<th>% in PD (n=185)</th>
<th>% in controls (n=283)</th>
<th>% in type 1 GD (n=73)</th>
<th>OR (95%CI)*</th>
</tr>
</thead>
<tbody>
<tr>
<td>N370S</td>
<td>2.70 (5)</td>
<td>0.35 (1)</td>
<td>86.30 (63)</td>
<td>7.8 (0.9-67.6)</td>
</tr>
<tr>
<td>RecNc/i (L444P+A456P+V460V)</td>
<td>1.6 (3)</td>
<td>0 (0)</td>
<td>10.96 (8)</td>
<td>Ø</td>
</tr>
<tr>
<td>R463C</td>
<td>1.62 (3)</td>
<td>0 (0)</td>
<td>5.48 (4)</td>
<td>Ø</td>
</tr>
<tr>
<td>L444P</td>
<td>1.08 (2)</td>
<td>0 (0)</td>
<td>24.66 (18)</td>
<td>Ø</td>
</tr>
<tr>
<td>Other Gaucher’s causing mutations</td>
<td>5.4 (10)</td>
<td>0.35 (1)</td>
<td>45.2 (33)</td>
<td>16.1 (2.04-126.97)</td>
</tr>
<tr>
<td>E326K</td>
<td>7.57 (14)</td>
<td>2.47 (7)</td>
<td>0 (0)</td>
<td>3.2 (1.2-8.1)</td>
</tr>
<tr>
<td>Other variants with unproven pathogenicity</td>
<td>7.03 (13)²</td>
<td>1.06 (3)</td>
<td>0 (0)</td>
<td>7.0 (1.8-31.6)</td>
</tr>
<tr>
<td>Total</td>
<td>25.94 (48)</td>
<td>4.24 (12)</td>
<td>100 (100)</td>
<td>7.9 (4.1-15.4)</td>
</tr>
</tbody>
</table>

The number of carriers for each GBA variant is in brackets.

*OR: Odds ratio calculated for PD patients versus controls. CI: confidence interval. Ø: OR cannot be calculated, as all carriers were cases.

1 Additionally further 202 individuals were screened for just this variant and the frequency in these additional samples was 2.97 % (6/202). The total frequency, taking into account both sets of controls, was 2.68% (13/485 individuals).

2 IVS2+1 (1), R131C (2), W184R (1), N188S (1), H255Q (1), R257Q (1), D409H (2), RecTL (1) 3 R257Q (1)

3 R257Q (1)


5 E340A (1), T369M (1) and V458L (1).
4.2.5 Discussion

This study reports the largest series of early-onset PD cases in which the \textit{GBA} gene was entirely sequenced and shows that the frequency of \textit{GBA} mutations is much higher than in late-onset patients from the same geographic area. Prior to this work, only one study had screened the entire \textit{GBA} open-reading frame in EOPD cases, but in a much smaller cohort (Clark et al., 2007). Consistent with the results of this project, Clark and colleagues found that the frequency of \textit{GBA} mutations was 22.2\% in subjects with EOPD, compared with 9.7\% in cases with classic LOPD. This indicates that \textit{GBA} mutations represent by far the most common genetic alteration in EOPD, in particular in subjects with an onset of symptoms at ages between 30 and 50 years. As will be discussed more extensively in the next chapter, this finding bears important clinical consequences. It has been reported that EOPD cases often have a more benign disease course, in particular with regard to the risk of developing dementia (Schrag et al., 1998). However, this may not be true for carriers of \textit{GBA} mutations, as it is now recognized that the risk of progression both to dementia and major motor difficulties (i.e. balance loss) is significantly greater in \textit{GBA} mutation carriers (Alcalay et al., 2012, Winder-Rhodes et al., 2013).

Importantly, the \textit{GBA} variant E326K was by far the most common variant found in this cohort (7.5\% of cases screened), indicating that in north European populations this mutation may be the single most frequent predisposing factor for PD.

The possibility that the E326K could be a risk factor for PD was raised before, but discarded as in previous studies it was \textit{a priori} considered a benign polymorphism or the association with PD was not as robust as we describe in this study (Nichols et al., 2009, Lesage et al., 2011). These data confirm those initial observations and put uncertainty beyond doubt, demonstrating that indeed this variant does increase the risk of PD. Consistent with this results, similar data were presented by Pankratz and colleagues as part of a large meta-analysis of the existing available PD GWAS datasets, in which the E326K reached genome-wide significance (OR=1.71; p=5 × 10^{-8}) and hence was highlighted as a susceptibility allele for PD (Pankratz et al., 2012).

These results are in apparent contrast with the large meta-analysis of \textit{GBA} in PD performed by Sidransky et al., which failed to demonstrate this association (Sidransky et al., 2009). This difference could be due to the heterogeneity of populations included
in that study. Indeed, it was enriched with subjects from populations (e.g. Asian, Jewish, Portuguese) in which the E326K is absent or very rare (Clark et al., 2007, Bras et al., 2009, Mitsui et al., 2009).

Although biochemical observations have demonstrated that the E326K is a subtle loss-of-function variant (Montfort et al., 2004, Horowitz et al., 2011), this reduction of the activity is clearly not severe enough to lead to disease in the tissues where GD is associated with lysosomal storage dysfunction. However, when E326K is detected on the same allele with another mutant variant (e.g. L444P, N188S, or N370S), it does not act as a neutral allele, and instead contributes to GD severity by further reducing the residual enzymatic activity (Grace et al., 1999, Chabas et al., 2005, Rozenberg et al., 2006, Liou and Grabowski, 2012).

The association of the E326K with PD and its absence in the GD cohort, together with the presence of many other novel rare GBA variants in the PD group not associated with GD, may suggest a possible dissociation of the pathogenic molecular mechanisms that underlie these two diseases.

However, it is not clear whether the mild loss-of-function of the allele E326K is the cause of the increased risk for PD or whether other, still undetermined, effects of the mutation on lysosomal function are contributing to nigral neuron degeneration. Recent studies suggest that heterozygous GBA mutations may contribute to the pathogenesis of PD through a gain of function effect that may include endoplasmic reticular stress (Gegg et al., 2012).

Interestingly, an overall enrichment of variants of undetermined clinical relevance was detected in the PD group. Although further population studies and functional analyses are warranted to substantiate this observation, these data indicate that there may be other GBA variants that predispose to PD but are not associated with GD. It is also possible that, as PD is a frequent disorder and a higher number of PD cases, compared to GD, are screened for GBA mutations, some of these variants are actually GD-causing mutations, simply not observed yet in GD cases.

In conclusion, this study shows that GBA mutations are the most common genetic abnormality in patients with PD with an age at onset below 50. These results strongly link the pathogenesis of EOPD to that of LOPD and suggest that both forms are part of the disease spectrum associated with the same underlying pathogenic process.
Furthermore, these results incontrovertibly establish a pathogenic role for the allele E326K. The latter extends the relevance of GBA dysfunction to a much larger number of PD patients than previously thought, bearing crucial implications for the development and application of future treatment targeting lysosomal dysfunction in PD. Finally, these data reflects the need to accurately explore the contribution of all GBA rare variants to the increased risk of PD and not just alleles that have been previously associated to GD.
4.3 Genotype-Phenotype Correlation in A Cohort Of Parkinson Disease Patients Treated With Deep-Brain Stimulation

4.3.1 Statement of Contribution
I performed all the genetic experiments and analysed the genetic data. Dr Aikaterini Angeli performed the clinical assessment and Dr Thomas Foltynie performed the clinical data analysis.

4.3.2 Background
Patients with PD may present significant heterogeneity in their clinical features, in particular with relation to their motor phenotype, their rate of disease progression and the development of non-motor symptoms (e.g. cognitive decline) (Foltynie et al., 2002). From a clinical viewpoint, the purpose of identifying PD subtypes is to aid individually tailored prognosis and to understand whether different subtypes of PD may show different responses to treatments.

Previous authors have attempted to classify subgroups of PD patients based on clinical features, such as “Tremor Dominant PD” or “Akinetic-Rigid PD” (Jankovic et al., 1990). However, despite the recognition of the existence of PD heterogeneity, there has been little translation of theories into practice, largely because of the lack of consensus regarding how best to sub-classify PD based on clinical phenotype. Techniques other than clinical examination have been used to help explore PD heterogeneity, including neuro-physiological techniques (Morgante et al., 2006), CSF examination (Jellinger, 2012), structural (Jubault et al., 2011) & functional imaging (Brooks and Pavese, 2011) biomarkers as well as clinico-pathological correlations have all been explored (Selikhova et al., 2009). However, to date there have been no convincing demonstrations that any of these methods can reliably distinguish subtypes of disease.

It is possible that some features of PD phenotype heterogeneity could depend on the underlying genetic background, with superimposed effect of other environmental or stochastic inter-individual modifiers. A number of studies have described the phenotype of specific genetic mutations carriers and compared it to non-carriers (Lesage et al., 2011, Marras et al., 2011), but a direct comparison of PD phenotype between carriers of
mutations in different genes are few (Gan-Or et al., 2010).

Functional neurosurgery has embraced PD heterogeneity in day-to-day practice for many years. PD tremor responds very well to Deep Brain Stimulation (DBS) of the motor thalamus (Benabid et al., 1993), severe disabling dyskinesia responds very well to Globus Pallidum pars interna (GPi) DBS (Krack et al., 1998), whereas the Akinesia/Rigidity of PD is best relieved through targeting the Subthalamic nucleus (STN) (Limousin et al., 1995). Given that the risks of neurosurgery become higher with advancing age, patients undergoing DBS are a very skewed group of PD patients, necessarily having had many years of PD but nevertheless still remaining young and judged fit enough to undergo neurosurgery. Consequently PD patients undergoing DBS tend to have relatively young onset disease and may contain an over-representation of genetic forms of PD. Furthermore, as part of the pre-operative workup, patients undergo a systematic evaluation of PD severity both in the absence and presence of medication, as their drug responsiveness is a helpful predictor of subsequent DBS response.

I performed an extensive genetic characterization of a series of patients with PD undergoing DBS surgery in a single centre. Cases were screened for point mutations in PARK2, GBA, for exonic rearrangements in SNCA and PARK2 and for the presence of the LRRK2 variants c.6055 G>A; p.Gly2019Ser.

The clinical phenotype among the carriers of different genetic mutations both on and off dopaminergic medication, as well as assessing the range of responses to DBS, was compared. In addition, a subgroup of these patients with complete longitudinal follow up assessments allows us to evaluate the natural history of PD following DBS according to genotype with respect to, perhaps the most important long term outcome, cognitive decline.

4.3.3 Subjects, Materials and Methods

Subjects
94 patients who underwent DBS for PD were included in the study. The mean age at PD onset of the cohort was 40.4±8.2 (range 7-58). Of these 31 (32.9%) had a positive family history for PD in a first or second degree relative.

Clinical assessment
All patients included in this series were followed-up at the National Hospital for
Neurology and Neurosurgery (NHNN), Queen Square. As part of their routine assessment, details were recorded regarding their basic demographics such as age at onset, initial symptoms and disease duration. The L-DOPA equivalent dose was derived using standard formulae (Tomlinson et al., 2010).

As part of the decision making process regarding appropriateness of DBS, all patients underwent detailed assessments including a formal L-DOPA challenge comparing scores on motor part of the Unified Parkinson’s disease rating scale (UPDRS) in the practically defined “Off state”- (overnight free from dopaminergic medication) and the “On state” after a L-DOPA challenge (250mg of oral dispersible L-DOPA). Additional assessments included a detailed battery of Neuropsychological assessments including the Mattis Dementia Rating Scale (DRS-2). From the UPDRS motor scale, sub-scores were derived to quantify tremor (based on the sum of items 20 and 21), akinesia and rigidity (based on the sums of items 22-26 and 31), and axial features (based on the sum of Items 18, 19, & 27-30). Part 4 of the UPDRS was divided into 2 sub-scores; items 32 to 34 for dyskinesia and items 35-39 to reflect OFF periods.

The optimal DBS target was chosen according to the frequency and severity of OFF-symptoms, dyskinesias, tremor, speech-intelligibility, and cognition. STN was targeted in the majority of patients requiring DBS. However, the presence of severe dyskinesias was an indication to target the bilateral internal globus pallidus (GPi), whereas severe tremor, especially if unresponsive to L-DOPA and in the absence of other disabling PD symptoms, is an indication for DBS of motor thalamus (VIM DBS).

Post-operative assessments were performed at 12 months after DBS implantation. A subgroup of patients with longer follow-up after DBS surgery was recruited to repeat cognitive assessment, including the DRS-2. Their mean annual change in cognitive performance between baseline and follow up was calculated as follows (DRS-2 score at baseline – DRS-2 score at follow up)/ number of years between assessments.

**Genetic analysis**

The coding regions and the flanking intronic sequences of PARK2 and GBA were analysed by PCR amplification and subsequent direct Sanger sequencing, as described in Materials and Methods. Primer sequences are listed in the appendix. Exonic rearrangements (deletions/duplications) in SNCA, PARK2, PINK1, PARK7 and the presence of the p.Gly2019Ser mutation in LRRK2 were screened using a multiplex
ligation-dependent probe amplification (MLPA), as described in Materials and Methods. Patients found by MLPA to be positive for \textit{LRRK2} p.Gly2019Ser mutation had this result confirmed by Sanger sequencing.

\textit{Statistical analysis}
Chi-squared test was used to make comparisons between categorical data. Continuous data were checked for normality using Shapiro-wilk test. For normally distributed variables, one way analysis of variance (ANOVA) was used to compare genetic subgroups. Post hoc pair wise comparisons were performed using Sidak's method. For non-normally distributed variables, Kruskal-Wallis & non-parametric post hoc pair wise comparisons were used. Analyses were performed using the statistical analysis program R (http://www.r-project.org/).

4.3.4 Results

\textit{Genetic results}
27 out 94 patients included in the study (29\%) carried at least one mutation in the genes I tested. The results of the genetic screening are summarised in Table 4-2.

Eight individuals (8.5\% of the cohort) had at least one pathogenic mutation in \textit{PARK2}. Of these six had homozygous or compound heterozygous mutations (phase of the mutations unknown), while two patients had a single mutation only (one of whom also had a \textit{GBA} T369M mutation). Given the controversial role of single heterozygous \textit{PARK2} mutations (Klein \textit{et al.}, 2007), only patients with bi-allelic \textit{PARK2} mutations were included in the phenotypic analysis.

Eighteen patients (19.1\% of the cohort) had at least one mutation in the \textit{GBA} gene. Amongst the \textit{GBA} mutation carriers, ten carried GD-causing mutations (of these one had the L444P combined with E326K) and five carried the pathogenic common risk factor E326K (four heterozygotes and one homozygote; see previous section of this chapter for further discussion about this variant). One case carried the novel variant L271P, predicted pathogenic by \textit{in silico} prediction scores and therefore likely to be a deleterious variant. Two cases carried the variants T369M and E388K, respectively. Given the uncertainty regarding the clinical relevance of the two variants (Walker \textit{et al.}, 2003) (Paciotti \textit{et al.}, 2012), these two cases were excluded as well from further phenotypic analysis.
Five patients (5% of the cohort) carried the common pathogenic mutation in \textit{LRRK2} p.Gly2019Ser. One of the cases with p.Gly2019Ser also carried the E326K \textit{GBA} mutation. This patient developed PD symptom at age 35, the youngest age at onset from the group of \textit{LRRK2} positive patients reported in this study and overall much earlier than the average age at onset of the \textit{LRRK2} mutation carriers (Healy \textit{et al.}, 2008). This finding possibly suggests a synergistic effect of the two variants on the onset anticipation.

The remaining 67 patients had no mutation detected using the genetic tests performed.

\textbf{Phenotypes}

The mean age at onset of the bi-allelic \textit{PARK2} carriers (26.6 years) was significantly lower than \textit{GBA} mutation carriers (p=0.0009), \textit{LRRK2} p.Gly2019Ser mutation carriers (p=0.0008) and gene negative (p=0.006) patients (Table 4-3). There was no significant difference in the age at onset between the other groups.

The mean duration of PD in patients undergoing DBS surgery was 15.0±6.6 years (range 3.8-38.0 years). Patients with \textit{PARK2} mutations had a longer duration of disease at the time of surgery than all other subgroups (p=0.007). On the contrary, \textit{GBA} mutation carriers had a significantly shorter mean duration of disease at the time they needed DBS surgery compared to all other subgroups (p=0.001) (Table 4-3). There was no difference in duration of disease between the \textit{LRRK2} p.Gly2019Ser and gene negative groups.
Table 4-2 Description of the genetic mutations in PARK2, GBA and LRRK2 in a cohort of PD patients with DBS surgery.

<table>
<thead>
<tr>
<th>Genetic test results</th>
<th>n</th>
<th>Description</th>
<th>Age at symptom onset ± SD</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>PARK2</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(Compound heterozygotes/homozygotes)</td>
<td>6</td>
<td>Homozygous c.101_102delAG</td>
<td>30</td>
</tr>
<tr>
<td></td>
<td></td>
<td>c.337_376del+c.465_466del</td>
<td>36</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Homozygous del exons 3-4</td>
<td>27</td>
</tr>
<tr>
<td></td>
<td></td>
<td>p.Arg275Trp+het dup of exon 6</td>
<td>7</td>
</tr>
<tr>
<td></td>
<td></td>
<td>c.337_376del +p.Ala82Glu</td>
<td>39</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(Mean 26.5±11.6)</td>
<td></td>
</tr>
<tr>
<td><strong>PARK2</strong></td>
<td>2</td>
<td>p.Arg334Cys</td>
<td>41</td>
</tr>
<tr>
<td>(Single mutation carriers)</td>
<td></td>
<td>p.Pro437Leu + GBA T369M&lt;sup&gt;a&lt;/sup&gt;</td>
<td>39</td>
</tr>
<tr>
<td><strong>GBA&lt;sup&gt;a&lt;/sup&gt;</strong></td>
<td>16</td>
<td>R463C</td>
<td>45</td>
</tr>
<tr>
<td></td>
<td></td>
<td>L444P+E326K</td>
<td>34</td>
</tr>
<tr>
<td></td>
<td></td>
<td>N370S</td>
<td>45</td>
</tr>
<tr>
<td></td>
<td></td>
<td>D409H</td>
<td>39</td>
</tr>
<tr>
<td></td>
<td></td>
<td>recNciI</td>
<td>40</td>
</tr>
<tr>
<td></td>
<td></td>
<td>R463C</td>
<td>45</td>
</tr>
<tr>
<td></td>
<td></td>
<td>N188S</td>
<td>49</td>
</tr>
<tr>
<td></td>
<td></td>
<td>R275Q</td>
<td>42</td>
</tr>
<tr>
<td></td>
<td></td>
<td>IVS2+1 G&gt;A</td>
<td>41</td>
</tr>
<tr>
<td></td>
<td></td>
<td>L444P</td>
<td>45</td>
</tr>
<tr>
<td></td>
<td></td>
<td>E326K/E326K</td>
<td>42</td>
</tr>
<tr>
<td></td>
<td></td>
<td>E326K</td>
<td>36</td>
</tr>
<tr>
<td></td>
<td>E326K</td>
<td>51</td>
<td></td>
</tr>
<tr>
<td>----------------</td>
<td>-------------</td>
<td>----</td>
<td></td>
</tr>
<tr>
<td></td>
<td>E326K</td>
<td>58</td>
<td></td>
</tr>
<tr>
<td></td>
<td>E326K+LRRK2 p.Gly2019Ser(^a)</td>
<td>35</td>
<td></td>
</tr>
<tr>
<td></td>
<td>L217P</td>
<td>39</td>
<td></td>
</tr>
<tr>
<td>(Mean 42.9±6.2)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LRRK2</td>
<td></td>
<td>5</td>
<td></td>
</tr>
<tr>
<td></td>
<td>p.Gly2019Ser</td>
<td>40</td>
<td></td>
</tr>
<tr>
<td></td>
<td>p.Gly2019Ser</td>
<td>36</td>
<td></td>
</tr>
<tr>
<td></td>
<td>p.Gly2019Ser</td>
<td>49</td>
<td></td>
</tr>
<tr>
<td></td>
<td>p.Gly2019Ser</td>
<td>55</td>
<td></td>
</tr>
<tr>
<td></td>
<td>p.Gly2019Ser +GBA E326K(^a)</td>
<td>35</td>
<td></td>
</tr>
<tr>
<td>(Mean 43±8.7)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No mutation found</td>
<td>67</td>
<td>(Mean 40.8±7.2)</td>
<td></td>
</tr>
</tbody>
</table>

Numbers add up to 96 since 2 individuals marked with \(^a\) are represented twice in the table in view of their carrying 2 mutations.

* Allele names refer to the processed protein, excluding the 39-residue signal peptide
As part of the preoperative assessment for DBS, all patients underwent an L-DOPA challenge. All patients in this series had a good response to L-DOPA (mean UPDRS part 3 improvement of 68% in the group as a whole), independently from the mutational status. This is not surprising, as a positive L-DOPA challenge represents an essential part of the basis for predicting a beneficial response from DBS surgery and therefore only patients with an excellent L-DOPA responsiveness were included in the cohort.

Dyskinesia scores were higher in the PARK2 bi-allelic mutations carriers, in spite of the lower doses of L-DOPA they were taking.

All patients in this series had baseline neuropsychological performance that was judged to be clinically acceptable for DBS surgery.

During pre-operative assessment in the off phase, there were no overt differences in the motor phenotype according to mutational status. Tremor was present in all subgroups, and the proportion of tremor compared with akinesia/rigidity or axial features was similar in each group.

Chi-squared analysis showed a significant difference in targeted brain region for DBS according to mutational status (p<0.001), with an excess of PARK2 and GBA patients that received bilateral GPi DBS rather than STN DBS.

There were no significant differences in the level of improvement obtained with STN DBS between gene negative patients and patients with PARK2, GBA or LRRK2 mutations.

Longitudinal 5-year follow up data regarding cognitive performances were available for 35 subjects, all treated with STN DBS, and amongst them six were GBA mutations carriers. The mean reduction in Mattis DRS-2 scores in patients with GBA mutations was 4.4±7.3 points per year, compared with 0.5±0.9 points per year among gene negative patients, although this did not reach threshold for significance.
Table 4-3 Pre-operative UPDRS scores and response to L-dopa according to genetic subgroup

<table>
<thead>
<tr>
<th>Genetic results</th>
<th>Duration of PD at DBS assessment yrs (SD)</th>
<th>L-dopa Equivalent Dose; mg (SD)</th>
<th>UPDRS part 1 (SD)</th>
<th>UPDRS part 2 OFF meds (SD)</th>
<th>UPDRS part 2 On meds (SD)</th>
<th>UPDRS Part 3 Off meds (SD)</th>
<th>UPDRS Part 3 On meds (SD)</th>
<th>Percentage improvement UPDRS part 3 with L-dopa</th>
<th>UPDRS Part 4 Dyskinesia score (SD)</th>
<th>UPDRS Part 4 Off Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>PARK2 (N=6)</td>
<td>25.2 (12.8)</td>
<td>960 (611)</td>
<td>1.7 (2.0)</td>
<td>24.3 (4.0)</td>
<td>6 (6.2)</td>
<td>57.0 (11.2)</td>
<td>21.0 (6.4)</td>
<td>61.0 (18.3)</td>
<td>5.3 (3.9)</td>
<td>3.5 (1.3)</td>
</tr>
<tr>
<td>GBA (N=16)</td>
<td>11.2 (5.0)</td>
<td>1143 (540)</td>
<td>1.5 (1.5)</td>
<td>24.6 (11.3)</td>
<td>10.9 (12.3)</td>
<td>51.3 (14.0)</td>
<td>18.0 (15.4)</td>
<td>66.9 (18.6)</td>
<td>3.2 (2.7)</td>
<td>4.9 (1.9)</td>
</tr>
<tr>
<td>LRRK2 (N=5)</td>
<td>12.1 (1.8)</td>
<td>1317 (803)</td>
<td>1.3 (1.3)</td>
<td>26.8 (7.9)</td>
<td>4.5 (4.3)</td>
<td>65.4 (14.9)</td>
<td>10.8 (5.1)</td>
<td>84.9 (5.5)</td>
<td>3.5 (3.1)</td>
<td>4.8 (0.9)</td>
</tr>
<tr>
<td>No mutations found N=67</td>
<td>15.1 (5.5)</td>
<td>1280 (548)</td>
<td>2.0 (1.7)</td>
<td>22.7 (7.9)</td>
<td>8.0 (7.3)</td>
<td>47.4 (14.7)</td>
<td>15.6 (11.3)</td>
<td>68.5 (19.3)</td>
<td>3.2 (2.3)</td>
<td>4.5 (1.2)</td>
</tr>
</tbody>
</table>
4.3.5 Discussion

This study shows the usefulness of studying a cohort of cases undergoing DBS as a valuable resource for detecting genetic forms of PD. DBS is usually performed in patients with advanced PD that has been refractory to non-invasive therapies, and yet patients are also selected on the basis that they are still young enough to tolerate invasive neurosurgery. Therefore, it is not surprising that EOPD are over-represented in this cohort.

While some of the Mendelian forms of PD can present in later decades, it is indeed well established that, overall, mutations in most Mendelian genes tend to be associated with an earlier onset of symptoms (Alcalay et al., 2010). This could explain in part why the frequency of PD genes mutation carriers is much higher in this cohort (27/94 patients; 29%), than in population representative cohorts of PD (Winder-Rhodes et al., 2013). Although it is widely known that both GBA mutations (Sidransky et al., 2009) and LRRK2 mutations (Ozelius et al., 2006) are more common in Jewish populations, this was not the reason for the high rate of gene positive subjects in our cohort.

While the use of DBS cohorts represents a clear advantage in terms of facilitating a rapid identification of gene positive patients, there are some biases to consider that may impact on the relevance of these data to the broader PD population. Patients with major cognitive, main psychiatric problems or lack of L-DOPA response are excluded during the selection process, and therefore this limits the ability to comment whether these features are more or less frequent according to mutational status. However, in this study it was therefore possible to explore the underlying PD phenotype and disease progression in patients undergoing with DBS with different genetic forms of PD, without the major bias of dopamine replacement. Patients were indeed routinely systematically evaluated On and Off medication as part of their DBS work-up, and thus it was possible to observe and evaluate their Off medication phenotype.

Apart from confirming the known PARK2-related feature of younger age at onset and earlier and more severe L-DOPA induced dyskinesias (Doherty et al., 2013), any consistent phenotypic difference between carriers of mutations in different genes and gene negative cases were not identified at a single time point. This includes evaluating the relative distribution of tremor, akinesia/rigidity and axial features as well as the degree of L-DOPA responsiveness.
However, a number of interesting observations were made through analysis of the longitudinal data. Duration of disease at time of needing DBS surgery was significantly different between genetic subgroups. As mentioned, \textit{PARK2} mutation carriers had an earlier age at onset of disease and earlier dyskinesias, however they also have a longer duration until requiring DBS, which indicates a more benign form of disease. As observed in this study, \textit{PARK2} patients have been previously reported to respond well to DBS (Capecci \textit{et al.}, 2004, Romito \textit{et al.}, 2005, Lohmann \textit{et al.}, 2008, Moro \textit{et al.}, 2008), though these series restricted their reports to response to STN DBS. In this series the GPi target was favoured, given the greater severity of dyskinesias in \textit{PARK2} cases.

Importantly, PD patients carrying \textit{GBA} mutations needed DBS significantly earlier in their disease course than any other mutation group. At the time of needing DBS, \textit{GBA} mutations carriers presented off-medication comparable clinical features to other groups, while with a significantly shorter duration of disease. This suggests a more aggressive form of the disease in PD patients with \textit{GBA} mutations. This is in keeping with data obtained from an unselected population of PD cases from the UK showing that PD patients with \textit{GBA} mutations have a faster rates of progression to major motor endpoints (Hoehn and Yahr stage 3) in cases followed longitudinally (Winder-Rhodes \textit{et al.}, 2013). At one-year follow-up, the response to both STN DBS and GPi DBS was not different between \textit{GBA} patients and the other groups. However, longer follow up showed that patients undergoing STN DBS with \textit{GBA} mutations may develop far greater cognitive impairment than any other group. This possibly suggests there may be an interaction between \textit{GBA} status and risk of subsequent decline following DBS. Consistent with this observation, a previous report of three \textit{GBA} patients undergoing STN DBS also identified a more aggressive process that lead to both cognitive impairment and axial impairments (Weiss \textit{et al.}, 2012).

Patients with \textit{LRRK2} mutations have been previously shown to have a good response to STN DBS (Schupbach \textit{et al.}, 2007, Gomez-Esteban \textit{et al.}, 2008, Breit \textit{et al.}, 2010). Consistent with previous data, this study confirms that G2019S mutation carriers have indeed an excellent response to STN DBS, without risk of major cognitive decline. The particular suitability of \textit{LRRK2} G2019S mutation carriers to DBS treatment may explain the surprisingly high frequency of the mutation in this cohort (\textasciitilde5\%), compared to unselected PD case from the UK (\textasciitilde1\%) (Healy \textit{et al.}, 2008).
In conclusions, from a longitudinal perspective, differences in phenotype between carriers of mutations in different genes appear to clearly emerge. *PARK2* mutation carriers have a younger onset of disease and more severe dyskinesias, but overall a more indolent disease course. On the other hand, patients with *GBA* mutations have a more aggressive motor disease needing earlier intervention with therapies such as DBS and are likely to have more rapid cognitive impairment following STN DBS implantation.
4.4 Mutational Analysis of the GTP cyclohydrolase-1 Gene in Parkinson Disease

4.4.1 Statement of Contribution
I designed the research, performed the genetic analysis and collected the clinical information concerning the four pedigrees reported in the first section of the study, performed clinical assessment and DNA and RNA analysis in family 1, and analysed the whole-exome sequencing data, generated and provided by the International Parkinson Disease Genetic Consortium (IPDGC) and UCL-exomes. Dr Ioannis Isaias performed the analysis and quantification of the dopaminergic imaging data.

4.4.2 Background
GTP cyclohydrolase 1 (GTPCH), encoded by the GCH1 gene (14q22.1-q22.2; OMIM 600225), is the enzyme controlling the first and rate-limiting step of the biosynthesis of tetrahydrobiopterin (BH4), the essential cofactor for the activity of tyrosine hydroxylase, and for dopamine production in nigrostriatal cells (Kurian et al., 2011).

Mutations in GCH1 are the commonest cause of dopa-responsive dystonia (DRD) (Clot et al., 2009), a disorder that presents typically in childhood with lower limb dystonia and subsequent generalisation (Nygaard, 1993). An excellent and sustained response to small doses of L-DOPA, generally without the occurrence of motor fluctuations generally observed in patients with PD, is the classic hallmark of the disease (Trender-Gerhard et al., 2009).

The diagnosis of GTPCH-deficiency may be supported by low cerebrospinal fluid (CSF) levels of pterins, dopamine and serotonin metabolites (Assmann et al., 2003) or by an abnormal phenylalanine-loading (PLT) (Bandmann et al., 2003).

Inheritance is usually autosomal dominant with incomplete penetrance (Furukawa et al., 1998), though recessive cases have been described (Opladen et al., 2011). Dominant GCH1 mutations result in a significant reduction of GTPCH activity through a dominant negative effect of the mutant protein on the normal enzyme (Hwu et al., 2000).

Neuropathological examination, in a limited number of DRD cases, revealed marked reduction of melanin pigment and dopamine content in nigrostriatal neurons, but no
evidence of nigral cell loss or degeneration (Furukawa et al., 1999).

Parkinsonian features are common in patients with DRD (Tassin et al., 2000) and carriers of GCH1 mutations may develop adult-onset parkinsonism in absence of dystonia (Nygaard et al., 1990). Based on previous studies, the prevailing hypothesis was that parkinsonism represented an atypical, age-specific, presentation of DRD without nigral degeneration (Nygaard and Wooten, 1998).

The aim of this work was to explore the relationship between GCH1 mutations and parkinsonism and investigate whether adult GCH1 mutation carriers are actually at increased risk of developing true neurodegenerative PD.

### 4.4.3 Subjects, Materials and Methods

**Family Study**

**Pedigrees**

DRD pedigrees were included in the study, where family members affected with adult-onset parkinsonism were available for clinical and genetic examination and in whom dopaminergic studies had been performed. Local ethics committees approved the study and informed consent for genetic testing was obtained in all cases.

**Genetic Analysis**

Probands were screened for GCH1 mutations (NCBI transcript NM_000161.2) by standard bi-directional Sanger sequencing of all 6 coding exons and exon-intron boundaries. Dosage analysis for GCH1 exonic deletions and duplications was performed by MLPA. Sanger sequencing and MLPA were performed as described in the Materials and Methods.

To test the effect of the splice-site variant c.343+5G>C identified in family-A, messenger RNA (mRNA) was extracted from peripheral blood of individuals II-1 and III-1 using the PAXgene Blood RNA System (Qiagen/PreAnalytix). Following RNA extraction, complementary DNA (cDNA) was synthesized using the Applied Biosystems High-Capacity cDNA reverse transcription kit. The GCH1 transcript was amplified using exonic primers in exon 1 and exon 3 in both cases and two controls (supplementary Fig. 1A). Following separation by agarose gel electrophoresis, the mutant bands from the patient samples were extracted and sequenced.
Transcript PCR primers

<table>
<thead>
<tr>
<th>Primer</th>
<th>Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>GCH1_exonic_1F</td>
<td>GGGAGTGTGATCTAAGCAG</td>
</tr>
<tr>
<td>GCH1_exonic_3R</td>
<td>AAGTTTGCTGAGGCCAAGG</td>
</tr>
</tbody>
</table>

A SYBR Green PCR assay was designed to establish if the c.343+5G>C mutation caused retention of GCH1 intron 1. Two PCRs were employed, each using the same reverse primer in exon 3 (GCH1_rt_Exon3_R), and a forward primer in either intron 1 (GCH1_rt_Intron1_F) or exon 1 (GCH1_rt_Exon1_F).

rtPCR primers

<table>
<thead>
<tr>
<th>Primer</th>
<th>Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>GCH1_rt_Intron1_F</td>
<td>catgttactaaagcaagctctgta</td>
</tr>
<tr>
<td>GCH1_rt_Exon1_F</td>
<td>TCTTCACCAAGGGCTACCAG</td>
</tr>
<tr>
<td>GCH1_rt_Exon3_R</td>
<td>AAGTTTGCTGAGGCCAAGG</td>
</tr>
</tbody>
</table>

Real-Time PCR was performed on a RotorGene 6000, measuring product accumulation after each amplification cycle. All PCRs for the standard curves and samples were performed in triplicate, the results for sample replicates were averaged. Serial dilutions of control cDNAs were prepared for each PCR over a 27-490 fold range. Standards and patient samples were amplified on the same experimental run. Rotor-Gene 6000 Series Software 1.7 was used to calculate the cycle threshold (Ct) at which exponential amplification was occurring in the PCRs and a standard curve was plotted for each PCR (log DNA concentration vs. Ct) from the serial dilutions. The relative concentration of each transcript was then calculated for all samples by interpolating from the standard curves. The ratio of immature/mutant transcript : Wild-Type transcript abundance was then derived for each sample.

Dopamine transporter imaging studies

Dopaminergic striatal innervation was evaluated as dopamine reuptake transporter (DAT) density by means of Single Photon Computed Tomography (SPECT) and [\(^{123}\text{I}\)]N-o-fluoropropyl-2β-carbomethoxy-3β-(4-iodophenyl) tropane ([\(^{123}\text{I}\)]FP-CIT).
SPECT data acquisition and reconstruction has been described in details elsewhere (Isaias et al., 2010). To obtain comparable measurements among different centers, $[^{123}]$FP-CIT binding values for the caudate nucleus and putamen were calculated by means of the Basal Ganglia Matching Tool (Nobili et al., 2013).

**Whole-exome sequencing study**

**Participants and study design**

The study initially involved 1337 unrelated subjects with PD and 1764 controls of European origin or North American of European descent that underwent whole-exome sequencing.

Patients, originating mainly from the USA, UK, Holland and France, were recruited by the IPDGC. A further 190 PD cases were recruited through a community-based epidemiological study of PD in Estonia (University of Tartu, Estonia).

Patients with PD were clinically diagnosed according to the UK Parkinson disease Society Brain Bank (UKPDSBB) criteria (Hughes et al., 1992).

Control samples were collected by the UCL-exomes. Controls had no diagnosis of PD, DRD or any other movement disorder. Whole-exome sequencing data from additional 4300 North American individuals of European descent were analysed from the publicly available from the NHLBI Exome Sequencing Project Exome Variant Server (EVS) database.

**Procedures**

Whole exome-sequencing was performed as described in the Materials and Methods.

**Statistical analysis**

Frequencies of coding and splice-site GCH1 variants in cases and controls were compared by means of Fisher’s exact (statistical significance set at $P$ value < 0.05 using a two-tailed test) and odds ratios (OR) and 95% confidence intervals (CI) were calculated. Analyses were performed using the statistical analysis program R (http://www.r-project.org/).
4.4.4 Results

Family study

Family A

The proband (III-1, Figure 4-1A) is a British 18-year-old man who had a difficult caesarean birth, with perinatal distress and subsequent developmental delay. At 18 months he developed inward turning of his feet with walking difficulties and frequent falls. He was diagnosed clinically with DRD at the age of 3 and administration of L-DOPA (300mg/day) markedly improved his symptoms. [$^{[123]}$I]FP-CIT SPECT, performed at the age 17, was normal. The proband’s father (II-1), who was initially thought to have cerebral palsy due to a birth injury, was subsequently diagnosed at age 42 with DRD. The proband’s grandfather (I-1) is a 65 year-old man with a six-year history of progressive asymmetric rest tremor in the right upper limb. Examination showed signs of typical PD with hypomimia, unilateral rest tremor and asymmetric bradykinesia. He did not present signs of dystonia. [$^{[123]}$I]FP-CIT SPECT showed bilateral reduced tracer uptake more marked on the left, consistent with nigrostriatal dopaminergic denervation. He responded well to L-DOPA therapy (300mg/day).
Figure 4-1 Pedigrees and $[^{125}]$FP-CIT SPECT scan images of the four families with $GCH1$ mutations involved in this study

P=Parkinson disease. D=DOPA-responsive dystonia
GCH1 analysis revealed a heterozygous splice site mutation (c.343+5G>C) in the three affected individuals. We previously detected c.343+5G>C in a recessive pedigree, carried by the unaffected mother of two very severely affected children who also inherited the K224R mutation from their unaffected father (Bandmann et al., 1996, Trender-Gerhard et al., 2009). However the c.343+5G>C mutation has not been previously described in DRD dominant pedigrees, making its pathogenicity uncertain.

To test the pathogenic effect of this variant on splicing, messenger RNA (mRNA) was extracted from peripheral blood of individuals II-1 and III-1 and the GCH1 transcript was amplified using exonic primers in exon 1 and exon 3 in both cases and two controls (Figure 4-2A). Following separation by agarose gel electrophoresis, the mutant bands from the patient samples were extracted and sequenced. Sequencing of the upper mutant band (352 bp) revealed a mutant transcript whereby exon 2 is spliced to base c.50 of exon 1, splicing-out 293 bp of exon 1 that would be predicted to cause a premature truncation of the protein after 8 out-of-frame amino acids (Figure 4-2A and B). Sequencing of the second mutant transcript at ~200 bp failed.

The low intensity of the mutant bands in Figure 4-2A prompted us to consider other possible effects of this mutation. A SYBR Green PCR assay was designed to establish if the c.343+5G>C mutation caused retention of GCH1 intron 1. As illustrated in figure 1, the Intron 1 – Exon 3 PCR captured any transcripts in which at least some splicing had occurred (exon 2 to exon 3), but in which intron 1 had not been spliced-out. This is referred to as 'immature/mutant transcript'. The ratio of the abundance of the transcripts was derived and indicates that the mutant/immature transcript is significantly more abundant relative to WT transcript in the patients than controls (Figure 4-2C), confirming that the c.343+5G>C mutation abolishes splicing of intron 1 and is therefore likely to lead to haploinsufficiency.
A) Transcript PCR: PCR from GCH1 exon 1 to exon 3 revealed additional bands in the patients carrying the c.343+5G>C mutation compared to controls (Con1 and Con2) at 352 bp and ~200 bp. (B) Mutant Transcript Sequencing: Sequencing electropherogram screenshot showing mutant GCH1 transcript from patient III-1 whereby exon 2 is aberrantly spliced to c.50 of exon 1. If translated, this would result in 16 in-frame amino acids, followed by 8 out-of-frame amino acids before a premature stop codon. (C) Intron 1 Retention: Two GCH1 cDNA PCRs each used a common reverse primer in exon 3 and a forward primer either in intron 1 (to detected incompletely-spliced or mutant transcript) or exon 1 (to detect wild-type (WT) transcript). Transcripts are shown with exons 1 to 6 alternately shaded and numbered, the translational start (ATG) and stop (X) sites are indicated with vertical lines. Each SYBR Green PCR was performed on cDNA from patients II-1 and III-1 (two independent extractions from each) and 6 controls. The ratio of the abundance of the transcripts was derived and indicates that the mutant/immature transcript is significantly more abundant relative to WT transcript in the patients than controls (Mean +/- SE: Student’s t-test).

Figure 4-2 RNA analysis performed in family-A
Family B

The proband (III-1; Figure 4-1B) is a 12-year-old right-handed female of German origin with DRD, with an onset at age 11 with writing and foot dystonia. Her mother (II-1) presented at age 39 with progressive loss of dexterity and slowness in her right arm and dystonic posturing of the right foot. Examination showed an asymmetric rigid-akinetic parkinsonian syndrome without tremor and severe right foot fixed dystonia. L-DOPA therapy resulted in marked improvement of both dystonic and parkinsonian symptoms. \[^{123}\text{I}]\text{FP-CIT SPECT}\) revealed an asymmetric bilateral reduced tracer uptake, more marked in the left striatum. There was sustained response to L-DOPA therapy although there was an increase in dose requirement (up to 800mg/day). L-DOPA -induced dyskinesias developed 6 years after initiation of L-DOPA. Examination of the proband’s 66-year-old grandmother (I-1) revealed oromandibular dyskinesias and upper limb dystonic features. She declined a trial of L-DOPA. Her \[^{123}\text{I}]\text{FP-CIT SPECT}\) displayed border-line reduced DAT values in both putamens.

\textit{GCH1} screening in this family revealed two variants: c.68C>T; p.Pro23Leu (carried by III-1 and II-1) and c.312C>A;p.Phe104Leu (carried by II-1 and I-1). There were no \textit{GCH1} exonic rearrangements. F104L is absent in public control datasets and has been previously reported in association with DRD (Clot \textit{et al.}, 2009). p.Pro23Leu (rs41298432) is a benign polymorphism present in population controls at a frequency of 1-2% (Jarman \textit{et al.}, 1997, Hauf \textit{et al.}, 2000).

To confirm GTPCH deficiency, PLT (100 mg/kg) was performed in individuals I-I and II-I and showed pathologically elevated phenylalanine/tyrosine ratios in both cases (Figure 4-3). CSF analysis, performed in individual III-I, displayed low levels of BH\(_4\) (13 nmol/l; 18-53 nmol/l) and neopterin (6nmol/l; 10-31 nmol/l), consistent with GTPCH deficiency. Given the benign nature of p.Pro23Leu, we hypothesize that the GTPCH deficiency confirmed in this patient may be the result of an as yet unidentified non-coding causative mutation.
Figure 4-3 Phenylalanine-loading test performed in family B
Abnormally elevated ratios of phenylalanine (Phe) and tyrosine (Tyr) plasma levels at 1, 2, 4, and 6 hours after oral loading with phenylalanine (100mg/kg) in individual I-1 (c.312C>A;p.Phe104Leu), II-1 (c.68C>T;p.Pro23Leu and c.312C>A;p.Phe104Leu) of family B compared to healthy subjects (HS).

Family C
The proband (II-1, Figure 4-1C) is a German 41-year-old lady, affected by DRD, who presented at age four with bilateral foot inversion on walking. Her father (I-1) is a 67-year old man with a one-year history of typical PD with left hand rest tremor, bilateral rigidity and bradykinesia and mild gait difficulties. There was no dystonia. [123I]FP-CIT SPECT examination revealed asymmetrically reduced DAT-density in the striatum. Rasagiline and pramipexole were started with good response. The mother (I-2), aged 62, had a normal neurological examination.

The proband was compound heterozygous for two GCH1 missense variants, c.610G>A; p.Val204Ile, inherited from the asymptomatic mother, and the novel variant c.722G>A; p.Arg241Gln, which was paternally inherited. p.Arg241Gln is absent in all public control datasets, is predicted deleterious by all in silico prediction tools and involves an amino acid residue conserved down to invertebrate species. Furthermore a pathogenic mutation at the same residue has already been reported (Bandmann et al., 1998).

CSF analysis in the parkinsonian case supported a pathogenic effect of the p.Arg241Gln mutation on GTPCH activity: pterin analysis revealed low BH$_4$ (8 nmol/L; 18-53), but
normal neopterin (24 nmol/L; 10-31); neurotransmitter analysis showed low homovanillic acid (95 nmol/L; 115-455) and 5-hydroxyindolacetic acid (59 nmol/L; 61-204), which are metabolites of dopamine and serotonin respectively.

Family D

The proband is an Italian 58-year-old lady (II-1, Figure 4-1D), who developed progressive tremor and clumsiness in the right arm at age 44. Clinical examination showed typical PD with hypomimia, hypophonia and asymmetrical bradykinesia and rigidity. Action dystonic tremor (right>left), poor postural reflexes and slow gait were also evident and there was a sustained response to L-DOPA. The dose was gradually increased up to 400 mg/day after which rotigotine 4 mg/day was added. Dyskinesias and wearing-off symptoms developed 6 years after L-DOPA initiation. $^{123}$I FP-CIT SPECT revealed asymmetrically reduced DAT binding values in the striatum.

Her sister (II-2; Figure 4-1D), aged 60, had a childhood onset of mild walking difficulties. At age 55, she developed exercise-induced left foot dystonia and dystonic tremor in both arms. She had no bradykinesia or other parkinsonian signs. Low-dose L-DOPA (100mg alternate days) was started with excellent symptom control. $^{123}$I FP-CIT SPECT was normal. Their father was reported to have a tremulous condition but was not available for clinical or genetic examination. GCH1 sequencing revealed that both sisters were heterozygous for the previously reported pathogenic mutation c.626+1G>C (Garavaglia et al., 2004).

The structure of the four pedigrees is shown in Figure 4-1. DAT binding values are reported in Table 4-4. The main clinical features of the GCH1 mutation carriers with adult-onset parkinsonism and abnormal $^{123}$I FP-CIT SPECT imaging are summarized in Table 4-5. Their clinical features fully met the UKPDSBB criteria for definite PD diagnosis. None of these cases presented significant diurnal fluctuations, worsening of symptoms in the evening or substantial sleep benefit, features often recognised in DRD cases (Kurian et al., 2011).
Table 4-4 DAT binding values of *GCH1* mutation carriers evaluated in this study

<table>
<thead>
<tr>
<th>Subject</th>
<th>Phenotype</th>
<th>L-putamen</th>
<th>R-putamen</th>
<th>L-caudatus</th>
<th>R-caudatus</th>
</tr>
</thead>
<tbody>
<tr>
<td>Family A/ I-1</td>
<td>PD</td>
<td>2.08&lt;sup&gt;a&lt;/sup&gt;</td>
<td>2.74&lt;sup&gt;a&lt;/sup&gt;</td>
<td>3.51</td>
<td>4.28</td>
</tr>
<tr>
<td>Family B/ II-1</td>
<td>PD</td>
<td>2.63&lt;sup&gt;a&lt;/sup&gt;</td>
<td>3.73&lt;sup&gt;a&lt;/sup&gt;</td>
<td>4.72</td>
<td>4.94</td>
</tr>
<tr>
<td>Family B/ I-1</td>
<td>DRD</td>
<td>2.85&lt;sup&gt;a&lt;/sup&gt;</td>
<td>2.96&lt;sup&gt;a&lt;/sup&gt;</td>
<td>3.95</td>
<td>3.95</td>
</tr>
<tr>
<td>Family C/ I-1</td>
<td>PD</td>
<td>1.86&lt;sup&gt;b&lt;/sup&gt;</td>
<td>1.64&lt;sup&gt;b&lt;/sup&gt;</td>
<td>2.52&lt;sup&gt;b&lt;/sup&gt;</td>
<td>2.19&lt;sup&gt;b&lt;/sup&gt;</td>
</tr>
<tr>
<td>Family D/ II-1</td>
<td>PD</td>
<td>1.53&lt;sup&gt;b&lt;/sup&gt;</td>
<td>2.63&lt;sup&gt;b&lt;/sup&gt;</td>
<td>3.29&lt;sup&gt;a&lt;/sup&gt;</td>
<td>4.5</td>
</tr>
<tr>
<td>Family D/ II-2</td>
<td>DRD</td>
<td>5.38</td>
<td>5.27</td>
<td>5.71</td>
<td>5.82</td>
</tr>
</tbody>
</table>

Values are expressed as the ratio of [123I]FP-CIT specific uptake in the region of interest/background. [123I]FP-CIT binding values for the caudate nucleus and putamen were calculated by means of the Basal Ganglia Matching Tool.

<sup>a</sup> Value below 97% inferior confidence limit of healthy controls

<sup>b</sup> Value below 90% inferior confidence limit of healthy control
Table 4-5 Characteristics of PD cases with GCH1 pathogenic variants and abnormal dopaminergic imaging described in this study and present in the literature

<table>
<thead>
<tr>
<th>Origin</th>
<th>Sex/Age at scan/Age at onset (y)</th>
<th>Mutation</th>
<th>Relatives with DRD</th>
<th>Age at L-DOPA start (y)</th>
<th>Current treatment dose (mg/day)</th>
<th>Parkinsonian features</th>
<th>Dystonic features</th>
<th>L-DOPA-induced complications</th>
<th>Scan Result</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>UK</td>
<td>M/65/59</td>
<td>c.343+5G&gt;C/w</td>
<td>Son and grandson</td>
<td>60</td>
<td>L-dopa 300</td>
<td>Hypomimia, R hand rest and re-emergent postural tremor, and bilateral rigidity and bradykinesia (R&gt;L)</td>
<td>No</td>
<td>No</td>
<td>Bilateral (L&gt;R) reduced DAT density</td>
<td>Present study (Family A)</td>
</tr>
<tr>
<td>Germany</td>
<td>F/47/39</td>
<td>p.Phe104Leu/p.Pro23Leu</td>
<td>Daughter, mother (?)</td>
<td>41</td>
<td>L-dopa 800</td>
<td>Hypomimia, bilateral rigidity, bradykinesia, reduced arm swinging (R&gt;L), and mild gait difficulties</td>
<td>R foot dystonia</td>
<td>Dyskinesias after 6 y of therapy</td>
<td>Bilateral (L&gt;R) reduced DAT density</td>
<td>Present study (Family B)</td>
</tr>
<tr>
<td>Germany</td>
<td>M/67/66</td>
<td>p.Arg241Gln/w</td>
<td>Daughter</td>
<td>/</td>
<td>Rasagiline 1 Pramipexole 0.375</td>
<td>Hypomimia, L hand rest tremor, bilateral bradykinesia and rigidity (L&gt;R), and mild gait difficulties</td>
<td>No</td>
<td>No</td>
<td>Bilateral (R&gt;L) reduced DAT density</td>
<td>Present study (Family C)</td>
</tr>
<tr>
<td>Italy</td>
<td>F/58/44</td>
<td>c.626+1G&gt;C/w</td>
<td>Sister</td>
<td>53</td>
<td>L-dopa 400 Rotigotine 4</td>
<td>Hypomimia, bilateral rigidity and bradykinesia (R&gt;L), mild postural instability, and gait difficulties</td>
<td>Bilateral (R&gt;L) upper limb dystonic tremor</td>
<td>Dyskinesias after 6 y of therapy</td>
<td>Bilateral (L&gt;R) reduced DAT density</td>
<td>Present study (Family D)</td>
</tr>
<tr>
<td>Japan</td>
<td>M/54/39</td>
<td>p.Arg184His/w</td>
<td>No</td>
<td>40</td>
<td>L-dopa 600</td>
<td>Cogwheel rigidity, akinesia, and postural instability</td>
<td>Dystonic posture in the four limbs (R&gt;L)</td>
<td>Wearing-off and dyskinesias after 10 y of therapy</td>
<td>Bilateral reduced FD intake</td>
<td>Kikuchi et al.</td>
</tr>
<tr>
<td>Country</td>
<td>Age/ID</td>
<td>Mutation/Genotype</td>
<td>Relative</td>
<td>Age (y)</td>
<td>Treatment</td>
<td>Symptoms</td>
<td>Motor Complications</td>
<td>Imaging Findings</td>
<td>Outcome</td>
<td>Other Findings</td>
</tr>
<tr>
<td>---------</td>
<td>---------</td>
<td>-------------------</td>
<td>----------</td>
<td>---------</td>
<td>-----------</td>
<td>----------</td>
<td>---------------------</td>
<td>----------------</td>
<td>---------</td>
<td>---------------</td>
</tr>
<tr>
<td>Denmark</td>
<td>M/38/28</td>
<td>p.Pro199Ser/w</td>
<td>Brother</td>
<td>33</td>
<td>L-dopa 350 Entacapone Selegiline 5</td>
<td>Bradykinesia and rigidity in the L arm</td>
<td>Dystonia of neck, trunk and four limbs, action tremor (L&gt;R)</td>
<td>Dyskinesias after 2 y of therapy</td>
<td>Bilateral (R&gt;L) reduced DAT density</td>
<td>Hjermand et al.</td>
</tr>
<tr>
<td>Germany</td>
<td>F/65/50</td>
<td>Complete deletion of the GCH1 gene/w Daughter</td>
<td>60 (for 10 y on dopamine agonist only)</td>
<td>L-dopa 200 Selegiline 5</td>
<td>Tremor in the R hand, reduced dexterity and mild gait disturbance</td>
<td>No</td>
<td>No</td>
<td>Bilateral (L&gt;R) reduced DAT density</td>
<td>Eggers et al.</td>
<td></td>
</tr>
<tr>
<td>Italy</td>
<td>M/59/NA</td>
<td>Deletion of exons 5-6/w Son with DRD, sister with MSA</td>
<td>NA</td>
<td>NA</td>
<td>Hypomimia, L hand rest tremor, bradykinesia (L&gt;R), mild gait difficulties</td>
<td>No</td>
<td>Dyskinesias after 10 y of therapy</td>
<td>Bilateral reduced DAT density</td>
<td>Ceravolo et al.</td>
<td></td>
</tr>
</tbody>
</table>

NA=not available; DRD=dopa-responsive dystonia; H&Y=Hoehn and Yahr; F=female; M=male L=left; R=right; MSA=multiple-system atrophy; y=years; w=wild-type.
Whole-exome sequencing study

To investigate if pathogenic variants in \textit{GCH1} could be found in PD subjects without a family history for DRD, I analysed whole-exome sequencing data of a large cohort of patients predominantly affected by early-onset or familial PD and controls.

After quality control checks (removal of gender mismatches, duplicate, related and non-Caucasian samples, samples with low call rate or excess of heterozygosity), 1318 PD cases and 1635 controls remained. Additional control data (\(n = 4300\)) were obtained from the publically available EVS dataset.

In total 1318 cases and 5935 controls were analysed for the presence of \textit{GCH1} coding (including small insertions/deletions, missense and stop-gain changes) or splice-site variants (±5 base pairs from the coding exons).

The mean age of subjects with PD was 55.7 ± 13.9 years (range 17-101; data available for 970 cases) and the mean age at onset was 46.7 ± 13.8 years (range 6-98; data available for 1194 cases). 423/1194 (35.4%) were early-onset cases (age at onset ≤ 40 years) and ~630 were familial cases (positive family history for PD in a first or second degree relative).

Coverage of the six \textit{GCH1} coding exons (NCBI transcript NM_000161.2) was similar in the three datasets as shown in Table 4-6.

Coverage of exon 1 was sub-optimal, being the average read depth only 12 in the IPDGC data, 17.9 in the UCL-ex data and 6 in the EVS data. Visual inspections of aligned reads of IPDGC and UCL-ex data (BAM files opened with Genome Browse software, Golden Helix) revealed that the first ~150 coding nucleotides of exon 1 were barely covered (average depth 1-2 reads) in the majority of samples. Similar coverage data were evident for the EVS data (Figure 4-4). The coverage of the rest of the exon was good (read depth ≥ 8). No common variants (frequency >1%) were identified. The benign polymorphisms p.Pro23Leu (rs41298432) and p.Pro69Leu (rs56127440), detected at similar frequencies in cases and controls, were not included in the analysis.
Table 4-6 Average exome sequencing coverage of the six *GCH1* exons.

<table>
<thead>
<tr>
<th>Exon</th>
<th>Genomic coordinates</th>
<th>IPDGC (n=1318)</th>
<th>UCL-ex (n=1645)</th>
<th>EVS (n=4300)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>chr14:55369016-55369405</td>
<td>12 (65.4%)</td>
<td>18 (61.8%)</td>
<td>6 (59.1%)</td>
</tr>
<tr>
<td>2</td>
<td>chr14:55332045-55332154</td>
<td>89 (99.6%)</td>
<td>48 (87.4%)</td>
<td>98 (99.7%)</td>
</tr>
<tr>
<td>3</td>
<td>chr14:55326326-55326528</td>
<td>58 (99.6%)</td>
<td>45 (97.1%)</td>
<td>85 (86.3%)</td>
</tr>
<tr>
<td>4</td>
<td>chr14:55313798-55313934</td>
<td>14 (81.8%)</td>
<td>35 (77.3%)</td>
<td>20 (83.3%)</td>
</tr>
<tr>
<td>5</td>
<td>chr14:55312427-55312630</td>
<td>94 (100%)</td>
<td>114 (98.9%)</td>
<td>93 (95%)</td>
</tr>
<tr>
<td>6</td>
<td>chr14:55310683-55310913</td>
<td>53 (99.2%)</td>
<td>22 (80%)</td>
<td>133 (99.5%)</td>
</tr>
</tbody>
</table>

IPDGC=International PDGenetic Consortium; UCL-ex= University College of London exomes consortium; EVS= NHLBI Exome Variant Server; chr=chromosome

Data are expressed as the average read depth per region of interest. In brackets is the percentage of samples that had a read depth ≥8 reads.
Figure 4-4 Visual representation of coverage depth for *GCH1* exon 1

NHLBI exome variant server sample average coverage depth, showing significantly reduced coverage is evident at the 5' end of the exon, most likely due to the high GC content (greater than 70%)
The main results of the GCH1 analysis are summarized in Table 4-7.

Combining cases and controls, 11 unique heterozygous GCH1 variants (ten missense and one stop-gain mutation) were identified in 16 individuals. Six variants were found only in PD cases (p.Gln110*, p.Gln110Glu, p.Ala120Ser, p.Asp134Gly, p.Gly217Val and p.Met230Ile), three in controls alone (p.Thr112Ala, p.Ile154Val and p.Arg198Gln) and two were detected in both groups (p.Val204Ile, p.Lys224Arg).

The frequency of GCH1 variants was significantly higher in PD cases (10/1318; 0.75%) than in individual (UCL-ex controls 1/1635; 0.06%; P = 0.003; OR 12.4 95% CI 1.7-541.1; EVS database 5/4300; 0.11%; P = 0.0004; OR 6.5, 95% CI 2.0-24.5) and combined datasets of controls (6/5935; 0.1%; P = 0.0001; OR 7.5, 95% CI 2.4-25.3).

All carriers of variants in GCH1 were negative for pathogenic mutations in all known genes associated with Mendelian forms of parkinsonism. The presence of copy number variants in the SNCA, PARK2, PARK7, and PINK1 genes was excluded by MLPA in all cases. One case carried the GBA mutation E326K.

The main features of the ten PD cases with pathogenic or possibly pathogenic GCH1 variants are listed Table 4-8. The age at onset of GCH1-mutated cases was 43.2 ± 13.4 years (range 17-61). Seven had a positive family history of PD. DNA of other family members was available for only one case and we showed segregation of the same GCH1 mutation (p.Gln110*) in the affected sister of the index case. All cases exhibited a variable combination of asymmetrical bradykinesia, rigidity, rest and postural tremor, walking difficulties, postural instability and excellent response to dopaminergic treatment, consistent with a clinical diagnosis of PD (Lees et al., 2009).

The two subjects with the youngest age at onset of symptoms (case 4 and 9, who developed symptoms at age 32 and 17 respectively) presented with dystonic features in the lower limbs at onset, a well-recognised characteristic of young-onset PD cases (Bozi and Bhatia, 2003). One case (case 5) developed lower limb dystonia in off periods over the course of the disease. The remainder did not present with any symptoms or signs of dystonia.

Detailed information about treatment was available for eight cases: the two cases (case 1 and 7) with the shortest disease duration (≤ 5 years) were treated only with a
dopamine-agonist, whereas the other cases were taking a combination of L-DOPA and other anti-parkinsonian drugs.

Mean disease duration was 17.6 ± 15.4 years (range 4-56). Patients with longer disease duration displayed a more severe clinical picture with some degree of postural instability (Hoehn and Yahr score ≥ 3), indicating disease progression in spite of the dopaminergic treatment.

In those patients taking L-DOPA and for whom follow up information was available (n = 7), all developed clinically relevant motor complications of chronic L-DOPA treatment, including wearing off, motor fluctuations and dyskinesias. Dyskinesias in case 4 were so disabling that he required treatment with deep brain stimulation of the subthalamic nuclei at age 60.

Most cases exhibited some of the typical non-motor features often recognised in PD, such as cognitive difficulties (case 5), hyposmia (case 3, 4, 5, 6 and 10) constipation (case 4 and 10), urinary problems (case 5, 6 and 9), fatigue (case 2 and 5) and sleep disturbances (case 4, 5, 6 and 10).
<table>
<thead>
<tr>
<th>Mutation</th>
<th>Exon</th>
<th>dbSNP</th>
<th>Prediction score</th>
<th>Previously described in DRD?</th>
<th>PD patients (n=1318)</th>
<th>UCL-ex controls (n=1635)</th>
<th>OR (95% CI)</th>
<th>p-value</th>
<th>EVS controls (n=4300)</th>
<th>OR (95% CI)</th>
<th>p-value</th>
<th>Total controls (n=5935)</th>
<th>OR (95% CI)</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>All variants</td>
<td>10</td>
<td>0.75%</td>
<td>1 (0.06%)</td>
<td>12.4 (1.7-541.1)</td>
<td>0.003</td>
<td>5 (0.11%)</td>
<td>6.5 (2.0-24.5)</td>
<td>0.0004</td>
<td>6 (0.1%)</td>
<td>7.5 (2.4-25.3)</td>
<td>0.0001</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>c.328C&gt;T; p.Gln110*</td>
<td>1</td>
<td>NA</td>
<td>Yes, in dominant and recessive pedigrees</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>c.328C&gt;G; p.Gln110Glu</td>
<td>1</td>
<td>2</td>
<td>No</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>c.334A&gt;G; p.Thr112A</td>
<td>2</td>
<td>2</td>
<td>No</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>c.358G&gt;T; p.Ala120S</td>
<td>2</td>
<td>4</td>
<td>No</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>c.401A&gt;G; p.Aspl134Gly</td>
<td>2</td>
<td>4</td>
<td>No</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>c.460A&gt;G; p.Ile154V</td>
<td>3</td>
<td>2</td>
<td>No</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>c.593G&gt;A; p.Arg198Gln</td>
<td>5</td>
<td>0</td>
<td>No</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>c.610G&gt;A; p.Val204Ile</td>
<td>5</td>
<td>2</td>
<td>Yes, in sporadic and recessive cases</td>
<td>3</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>c.650G&gt;T; p.Gly217Val</td>
<td>6</td>
<td>4</td>
<td>No</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gene</td>
<td>rsID</td>
<td>Frequency</td>
<td>Yes, in dominant and recessive pedigrees</td>
<td>1</td>
<td>0</td>
<td>2</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>---------</td>
<td>----------</td>
<td>-----------</td>
<td>------------------------------------------</td>
<td>----</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>c.671A&gt;G; p.Lys224Arg</td>
<td>rs41298442</td>
<td>2</td>
<td></td>
<td>1</td>
<td>0</td>
<td>2</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>c.690G&gt;A; p.Met230Ile</td>
<td>6</td>
<td>4</td>
<td>Yes, in a sporadic case</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

NA=not applicable; DRD=dopa-responsive dystonia; PD=Parkinson disease; UCL-ex= University College of London exomes consortium; EVS=Exome Variant Server; OR=odds ratio; CI=confidence intervals. \( p \)-values were calculated by means of Fisher’s exact test.

\(^a\) NCBI transcript NM_000161.2. This count includes all detected coding and splice-site variants at any frequency, but the two benign variants P23L and P69L.

\(^b\) This score, ranging from 0 to 4, indicates the number of tools (Polyphen-2, SIFT, LRT and Mutation Taster) predicting a pathogenic effect on the protein function.
<table>
<thead>
<tr>
<th>Case</th>
<th>Sex/Age / Age at onset (y)</th>
<th>GCH1 mutation</th>
<th>Family history of PD</th>
<th>Age at L-dopa start (y)</th>
<th>Current treatment (mg/day)</th>
<th>Parkinsonian features</th>
<th>H&amp;Y score</th>
<th>Cognitive symptoms</th>
<th>Other non-motor features</th>
<th>Dystonia</th>
<th>Levodopa-induced complications</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>F/47/43</td>
<td>p.Met230Ile</td>
<td>No</td>
<td>/</td>
<td>Pramipexole 0.75</td>
<td>Asymmetric onset, bilateral involvement with rest and postural tremor, bradykinesia and rigidity, mild gait difficulties</td>
<td>2</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>2</td>
<td>M/55/37</td>
<td>p.Lys224Arg</td>
<td>Yes (father)</td>
<td>NA</td>
<td>NA</td>
<td>Asymmetric onset, moderate bilateral involvement with rest tremor, bradykinesia and rigidity, postural instability and gait difficulties</td>
<td>3</td>
<td>No</td>
<td>Fatigue</td>
<td>No</td>
<td>Dyskinesias and wearing-off</td>
</tr>
<tr>
<td>3</td>
<td>M/49/35</td>
<td>p.Gly217Val</td>
<td>No</td>
<td>43</td>
<td>L-dopa 600, Tolcapone 400, Pramipexole 3.15</td>
<td>Asymmetric onset, slurred speech, mild L arm rest and postural tremor, moderate bilateral bradykinesia and rigidity, postural instability</td>
<td>3</td>
<td>Subjective loss of memory (MMSE 29/30)</td>
<td>Hyposmia, ICD</td>
<td>No</td>
<td>Initial dyskinesias and wearing-off</td>
</tr>
<tr>
<td>4</td>
<td>M/63/32</td>
<td>p.Val204Ile</td>
<td>Yes (1st degree cousin)</td>
<td>36</td>
<td>DBS, L-dopa 200 Amantadine 100 Rotigotine 8</td>
<td>Asymmetric onset, hypomimia, slurred speech, hypophonia, marked bilateral rest and postural tremor, moderate bilateral rigidity and bradykinesia, postural instability</td>
<td>4</td>
<td>No</td>
<td>Hyposmia, constipation, RBD</td>
<td>Right foot exercise-induced dystonia at onset</td>
<td>Disabling dyskinesias and on-off fluctuations</td>
</tr>
<tr>
<td>5</td>
<td>M/75/61</td>
<td>p.Val204Ile</td>
<td>Yes (mother)</td>
<td>61</td>
<td>L-dopa 400, Pramipexole 3.15</td>
<td>Asymmetric onset, rest and postural tremor (R&gt;L), bradykinesia and rigidity, mild gait disorder, hypomimia</td>
<td>3</td>
<td>Mild cognitive impairment</td>
<td>Hyposmia, fatigue, sleep and bladder disorder</td>
<td>Lower limb off-dystonia</td>
<td>On-off fluctuations (30% of waking day in off-state)</td>
</tr>
<tr>
<td>6</td>
<td>M/72/59</td>
<td>p.Val204Ile</td>
<td>Yes (mother)</td>
<td>65</td>
<td>L-dopa 600 Entacapone 800</td>
<td>Asymmetric onset, bilateral bradykinesia and rigidity (L&gt;R), no tremor. Mild gait difficulties and postural instability</td>
<td>3</td>
<td>No</td>
<td>Hyposmia, sleep and bladder disorder</td>
<td>No</td>
<td>Dyskinesias (30-40% of waking day), wearing-off</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>M/57/52</td>
<td>p.Asp134Gly</td>
<td>Yes (father and paternal aunt)</td>
<td>Ropinirole 14</td>
<td>Asymmetric onset, unilateral left arm rest tremor, bradykinesia and rigidity. Reduced arm swing</td>
<td>1</td>
<td>No</td>
<td>NA</td>
<td>No</td>
<td>No</td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>F/59/51</td>
<td>p.Ala120Ser*</td>
<td>Yes (mother)</td>
<td>NA</td>
<td>NA</td>
<td>Asymmetric onset, bilateral bradykinesia and rigidity. No tremor. Mild gait difficulties</td>
<td>2</td>
<td>No</td>
<td>NA</td>
<td>No</td>
<td>NA</td>
</tr>
<tr>
<td>9</td>
<td>F/73/17</td>
<td>p.Gln110X*</td>
<td>Yes (sister; father had tremor)</td>
<td>L-dopa 600</td>
<td>Bilateral rest and postural tremor (L&gt;R), bilateral rigidity and bradykinesia. Some postural instability</td>
<td>3</td>
<td>No</td>
<td>Urinary urgency</td>
<td>Lower limb dystonia at onset</td>
<td>Marked limb and truncal dyskinesias, off phases in the morning</td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>M/58/45</td>
<td>p.Gln110Glu</td>
<td>No</td>
<td>L-dopa 400</td>
<td>Bilateral severe bradykinesia and rigidity, postural instability, mild tremor, hypomimia</td>
<td>3</td>
<td>No</td>
<td>Hyposmia, constipation, fatigue, sleep disorder</td>
<td>No</td>
<td>Mild dyskinesias and wearing-off</td>
<td></td>
</tr>
</tbody>
</table>

NA = information not available; M = male; F = female; PD = Parkinson disease, y = years, ICD = Impulse control disorder, DBS = deep brain stimulation, RBD = REM behavioural sleep disorder, H&Y = Hoehn and Yahr. *This case also carries in the heterozygous state the GBA E326K variant
4.4.5 Discussion

Family study

I report here four unrelated DRD pedigrees in which loss-of-function GCH1 mutations (two splice-site mutations and two missense mutations, confirmed to be pathogenic by metabolic or CSF studies) were found in individuals, asymptomatic for DRD during childhood, who developed adult-onset parkinsonism. They all met the UKPDSBB clinical criteria for a definite diagnosis of PD and had imaging evidence of a PD-like nigrostriatal dopaminergic denervation.

A parkinsonian syndrome in the absence of dystonia has been reported in adults who are first-degree relatives of children with DRD. In a series of 21 families, Nygaard showed that 7/50 (14%) individuals older than 40 years had parkinsonism (Nygaard, 1993) and Hagenah and colleagues reported that 8/23 (34.7%) patients of their series had a positive family history for PD (Hagenah et al., 2005). GCH1 mutations have also been shown to segregate in pedigrees with multiple individuals affected by isolated parkinsonism (Irie et al., 2011).

This study provides evidence that in most of the cases the parkinsonian phenotype in adult GCH1 mutation carriers is likely due to nigrostriatal degeneration, rather than being part of the phenotypic spectrum of metabolic GTPCH-related striatal dopamine deficiency. This is consistent with other previous isolated reports of adult-onset parkinsonism in GCH1 mutation carriers with abnormal nigrostriatal imaging (features summarized in table 1) (Kikuchi et al., 2004, Hjerminde et al., 2006, Eggers et al., 2012, Ceravolo et al., 2013).

Our imaging findings are however in apparent contrast to a previous report by Nygaard and colleagues (Nygaard et al., 1992). The authors described a large DRD pedigree, in which three subjects had a late-onset benign parkinsonism, two of which had normal nigrostriatal dopaminergic function determined by means of 18F-fluorodopa (FD) PET.

Compensatory mechanisms at the presynaptic level (e.g. increased dopamine-intake and dopamine-decarboxylation activity) may result in relatively higher striatal FD uptake in the initial phase of PD, underestimating the degree of nigral cell decrease (Nandhagopal et al., 2011). DAT values are therefore a more precise indicator of dopaminergic innervation loss (Ito et al., 1999). I speculate that GCH1-parkinsonian cases with
normal FD-PET scan could have up-regulated compensatory dopaminergic activity at the presynaptic level, possibly masking the presence of striatal denervation.

In agreement with these findings, Gibb and Lees reported a case that presented with juvenile-onset parkinsonism and dystonia with good response to L-DOPA (commenced at the age of 30) and occurrence of disabling dyskinesias after one year of treatment. The patient died at 39 years and pathological examination showed a striking combination of low melanin content in nigral neurons and devastating neuronal loss with reactive gliosis. Furthermore, LBs were found in surviving nigral cells and in the locus coeruleus (Gibb et al., 1991). This case was subsequently demonstrated to be carrier of a heterozygous mutation in GCH1 (c.276delC) (Segawa et al., 2004).

**Whole-exome sequencing study**

I subsequently showed, in a large cohort of PD patients without family history of DRD, that rare GCH1 coding variants are associated with PD and increase the disease risk by seven-fold on average.

Among the GCH1 variants identified by exome sequencing, two (p.Gln110* and Lys224Arg) have been shown to cause GTPCH deficiency and DRD in dominant pedigrees (Leuzzi et al., 2002, Saunders-Pullman et al., 2004) and two (p.Val204Ile and p.Met230Ile) have been reported in heterozygous sporadic or in recessive cases with DRD (Segawa et al., 2004, Trender-Gerhard et al., 2009, Opladen et al., 2011).

The other heterozygous variants identified in this study could not be functionally investigated (e.g. PLT or CSF analysis), therefore their effect on GTPCH activity remains undetermined. However, three of the four novel variants (p.Ala120Ser, p.Asp134Gly and p.Gly217Val) detected in PD cases were located at amino acid positions that are fully conserved through species down to invertebrates and were predicted to be pathogenic by all *in silico* prediction tools, whereas this was not the case for any of the novel mutations present in controls.

Nevertheless, the limitations of prediction tools in reliably distinguishing benign from pathogenic missense changes are well known and therefore we did not exclude any variant from the association test based on predictions scores, possibly underestimating the effect size of GCH1 pathogenic variants.
The p.Pro23Leu (rs41298432) and p.Pro69Leu (rs56127440) variants were excluded from the exome sequencing analysis as several data argue against their pathogenic role.

p.Pro23Leu has been reported in population controls at a frequency of 1-2% (Jarman et al., 1997, Hauf et al., 2000), it affects a non-conserved amino acid residue and functional studies have shown that it does not affect GCH1 enzymatic activity (Furukawa et al., 1999). It was also shown not to segregate with the clinical phenotype in a large dominant dystonia kindred and GTPCH deficiency was excluded in affected carriers by CSF analysis (Hauf et al., 2000). In this study it was detected in only 1/1318 (0.07%) cases and 4/5915 controls (0.06%; difference not significant), due to very poor sequencing coverage of the first part of exon 1.

The p.Pro69Leu variant, despite its relatively high frequency in the general population, has been described only in two published DRD cases; Clot et al. detected it in the heterozygous state in a sporadic case, but in the same study it was also found at a similar frequency in controls (Clot et al., 2009); it was also found in an apparently recessive pedigree, in trans with a pathogenic frameshift mutation, but the phenotype of the carriers was not as severe as usually reported in recessive GTPCH deficient cases (Furukawa et al., 2004). In this study it was detected in 3/1318 (0.22%) cases and 10/5935 controls (0.17%; difference not significant). Among the three cases identified in this study with p.Pro69Leu, two had a positive family history for PD. Segregation analysis was performed in one pedigree for which DNA of multiple affected members was available. The mutation did not segregate with the clinical phenotype and no carriers presented features of DRD.

Previous studies investigating the contribution of rare coding GCH1 variants in small cohorts of PD cases have reported negative results although these were insufficiently powered to draw conclusions (Bandmann et al., 1996, Hertz et al., 2006, Cobb et al., 2009).

The association between GCH1 variants and PD is further strengthened by the recent identification, through the meta-analysis of GWAS data deriving from approximately 13,000 cases and 95,000 controls, that GCH1 is also a low-risk susceptibility locus for PD (Nalls et al., 2014). This finding potentially extends the role of GTPCH deficiency in the pathogenesis of PD beyond carriers of rare deleterious coding mutations.
The mechanism whereby \textit{GCH1} mutations could predispose to nigral cell degeneration is a matter of speculation. Biochemical evidence of GTPCH deficiency and reduced dopamine production has been reported in asymptomatic carriers of \textit{GCH1} mutations (Takahashi et al., 1994, Furukawa et al., 2004). It is therefore possible that GTPCH deficiency and the consequent chronic dopamine deficiency could directly predispose to nigral cell death. This would suggest that normal levels of dopamine exert a protective role on the survival of nigral neurons. There is increasing evidence that L-DOPA is not toxic to nigral neurons as was previously thought (Parkkinen et al., 2011). Furthermore, activation of dopamine receptors may have a strong antiapoptotic effect and increase survival of dopaminergic neurons (Nair et al., 2003, Vaarmann et al., 2013) (Nair et al., 2003, Vaarmann et al., 2013). In animal models, L-DOPA has been shown to promote recovery of nigrostriatal denervation (Datla et al., 2001).

Another possibility is that \textit{GCH1} mutation carriers who do not develop symptoms of DRD in childhood may have compensatory mechanisms that allow for normal nigrostriatal dopaminergic transmission. The maintenance of these mechanisms may increase nigral cell vulnerability to aging or other environmental and genetic factors, favouring degeneration.

It is also possible that the reduced striatal basal dopamine levels found in \textit{GCH1} mutation carriers may simply lower the threshold of nigral cell loss before parkinsonian symptoms are exhibited.

Lastly, we cannot exclude that other yet unrecognized cellular pathways, not related to dopamine synthesis, may be disrupted by GTPCH and BH$_4$ deficiency.

With this regard, Ryan and colleagues have recently proposed that different cellular mechanisms secondary to BH$_4$ deficiency, other than reduced dopamine levels, could be contributing to the death of nigral dopaminergic neurons (Ryan et al., 2015). BH$_4$ acts as an antioxidant itself (Milstien and Katusic, 1999) and is an essential cofactor for nitric oxide synthases (NOS) activity (Crabtree et al., 2009). Furthermore decreased BH$_4$ levels have been demonstrated to lead to NOS uncoupling which results in increased oxidative and nitrative stress (Chen et al., 2014). In addition to these mechanisms, Ryan and colleagues have recently demonstrated that α-synuclein modulates cellular BH$_4$ levels by regulating GTPCH activity (Ryan et al., 2014). This supports the compelling hypothesis that a pathogenic cascade may occur in nigral
neurons, whereby increased levels of α-synuclein and mitochondrial dysfunction lead to decreased GTPCH activity and BH₄ levels, which in turn may result in increased oxidative stress and cell death.

There are some important limitations in this study that need to be acknowledged.

Firstly, dopamine transporter imaging was not available for the PD cases with GCH1 variants identified in the exome sequencing study. It remains a possibility therefore that some of these cases (in particular case 9, who presented at age 17, with lower limb dystonia and parkinsonism) may represent DRD cases with a parkinsonian phenotype, which may have been misdiagnosed as PD. However, removal of the aforementioned case from the statistical analysis did not change substantially the significance of the association (P = 0.0003). Furthermore, most of the patients for whom clinical follow-up data were available showed a progressive disease course with increasing L-DOPA requirements, emergence of motor complications due to chronic treatment with L-DOPA and presence of classic non-motor features of PD, strongly supporting nigrostriatal cell loss as the underlying pathology. Although dyskinesias have been very rarely described also in DRD cases, these are significantly different from the ones generally observed in PD. Indeed they tend to appear at the beginning of the treatment and subside after dose reduction without reoccurring with subsequent slow dose increase (Furukawa et al., 2004, Lee et al., 2013).

Secondly, the effect of the GCH1 variants detected in the exome sequencing study on pterin and dopamine metabolism could not be determined at the individual level. Reduced penetrance of GCH1 pathogenic variants for the DRD phenotype is a well-established feature. Nevertheless it has been repeatedly reported, through analysis of brain tissue (Furukawa et al., 2002), CSF (Takahashi et al., 1994) and urine (Leuzzi et al., 2013), that even completely asymptomatic carriers of GCH1 mutations have abnormal metabolism of biopterins and dopamine, although to a lesser extent than DRD cases. This indicates the existence of a metabolic endophenotype, which we speculate could be the pathogenic mechanism underlying the increased risk for PD.

Thirdly, a cohort enriched with early-onset and familial PD cases was evaluated. Thus the frequency of detected GCH1 variants may not reflect the frequency in late-onset
sporadic cases.

Finally, I did not assess our samples for the presence of GCH1 copy number variants, possibly underestimating the frequency of GCH1 mutations.

In conclusion, I demonstrate that rare GCH1 coding variants should be considered as a risk factor for PD. This is derived both from imaging evidence of striatal dopaminergic denervation in GCH1 pathogenic variant carriers with a clinical diagnosis of definite PD (in DRD pedigrees) and from exome sequencing data which show a significant association between GCH1 coding variants and an increased risk for the disease.

These findings expand the clinical and biological relevance of GTPCH deficiency, suggesting a role not only in biochemical dopamine depletion and DRD, but also in nigrostriatal degeneration. The question as to how the same variants known to cause a Mendelian disease may also exist as risk alleles in PD may be explained by the well-known reduced penetrance of GCH1 pathogenic variants. Whether additional genetic or epigenetic factors play a role in determining the clinical phenotype of GCH1 variant carriers should be addressed by future studies.
Chapter 5. Exploring the Genetics Basis of Myoclonus-Dystonia

5.1 Outline of the Chapter

This chapter describes my efforts to identify novel genetic causes responsible for Myoclonus-Dystonia (M-D).

The chapter begins with an overview of the clinical and genetic features of M-D. The specific aims of the study were:

1) To identify the underlying causative mutation in a pedigree with genetically unexplained autosomal dominant M-D;

2) To identify the genetic cause in a recessive pedigree with genetically unexplained M-D;

3) To assess the role of the missense variant p.Arg1389His in CACNAIB, a newly identified M-D genetic cause, in a large international cohort of patients with M-D.

5.2 Overview of Myoclonus-Dystonia

M-D is a very rare condition with a suggested prevalence about 2 per million in Europe. M-D is clinically characterized by a variable combination of non-epileptic myoclonic jerks and dystonic movements (Asmus and Gasser, 2004).

The first locus for autosomal dominant M-D (DYT11) was mapped on chromosome 7q21-q31 (Nygaard et al., 1999). Heterozygous mutations (including nonsense, missense, deletions, and insertions) in the SGCE gene (MIM 604149) were subsequently identified as the genetic cause responsible for DYT11 and represent a major cause of inherited autosomal dominant M-D (Zimprich et al., 2001). Usually SGCE mutation carriers manifest the first symptoms in childhood with brief lightning-like myoclonic jerks affecting the upper body often being the most prominent symptom (Asmus et al., 2002). Dystonia, presenting as cervical dystonia or writer’s cramp, generally appears later and is generally of mild to moderate intensity. Cranial, laryngeal and lower limb
dystonia are conversely more rarely reported (Roze et al., 2008). Additional features such as alcohol responsiveness of jerks and psychiatric co-morbidities (depression, anxiety and obsessive-compulsive disorder, alcohol addiction and attention deficit and hyperactivity disorder) are frequently present in SGCE-mutation carriers (Peall et al., 2013).

Asmus and colleagues recognised that large heterozygous deletions of SGCE, undetectable by direct Sanger sequencing, may account for a substantial portion (~25%) of the genetically unexplained cases with familial M-D (Asmus et al., 2005). Large deletions on the 7q21.13-21.3 region, encompassing not only SGCE but also other neighbouring genes (e.g. KRIT1, COLIA2, SHFM1, DLX6 and DLX5), are associated with a M-D plus presentation, which include other clinical features, such as hearing loss, short stature, joint laxity and microcephaly, skeletal deformities and cavernous cerebral malformations on cranial MRI (Asmus et al., 2007, Grunewald et al., 2008, Saugier-Veber et al., 2010, Peall et al., 2014). Penetrance of mutations in the SGCE is reduced and related to maternal imprinting. Transcription occurs almost exclusively from the paternal allele, whilst transcription from the maternal allele is silenced by promoter methylation (Grabowski et al., 2003). Individuals with paternally inherited loss-of-function SGCE mutations exhibit M-D, whereas individuals with maternally inherited mutations are asymptomatic. This particular pattern of inheritance explains why patients with M-D due to SGCE-mutations may lack a positive family history.

The SGCE gene encodes the 438-aminoacid protein ε-sarcoglycan, member of the sarcoglycan gene family. In the brain SGCE mutations are thought to lead to mislocalisation of the protein from the plasma membrane to the endoplasmic reticulum and to promotion of its degradation by the proteasome (Esapa et al., 2007).

SGCE mutations, including both point mutations and large deletions, are detected in only ~30-50% of familial cases, suggesting genetic heterogeneity and the existence of other still unidentified disease-associated genes (Schule et al., 2004, Tezenas du Montcel et al., 2006, Carecchio et al., 2013, Peall et al., 2014). An additional locus on chromosome 18p11 (DYT15; MIM 607488) has been linked to M-D in a single large Canadian family, but the causative mutation has not been identified as of yet (Han et al., 2007). A frameshift mutation in TOR1A and a missense mutation in the DRD2 gene have been proposed as possible causes for M-D in single pedigrees, but then dismissed
as in both pedigrees pathogenic SGCE were subsequently detected (Klein et al., 2002).

Other autosomal dominant M-D syndromes include some cases of benign hereditary chorea due to NKX2-1 mutations, with additional signs such as delayed motor milestones and later chorea (Armstrong et al., 2011) and one family with GCH1 mutations, which presented with L-DOPA-responsive M-D and later developed parkinsonism (Leuzzi et al., 2002). With regard to AR disorders, M-D is described in one case with Vitamin E deficiency (Angelini et al., 2002), that later developed ataxia. When myoclonus and dystonia are only part of the phenotype a number of mitochondrial (Hanna and Bhatia, 1997, Garcia-Cazorla et al., 2008) or neurometabolic disorders (e.g. Lafora body disease, GM2 gangliosidosis, Niemann-Pick disease) should be considered in the differential diagnosis (Gouider-Khouja et al., 2010).

Recently, Groen and colleagues identified two missense variants, p.Arg1389His (rs184841813) in CACNA1B and p.Thr1904Met (rs114190729) in RELN, each in a single autosomal dominant Dutch pedigree with M-D lacking mutations in SGCE, as the likely causative mutations (Groen et al., 2015, Groen et al., 2015). However, the pathogenic role of these variants is not confirmed as yet.
5.3 Linkage Analysis and Exome Sequencing Identify A Novel Cause Of Familial Myoclonus-Dystonia

5.3.1 Statement of contribution
The index case of the family described in this chapter was recruited in Professor Kailash Bhatia’s clinic at the National Hospital for Neurology, London. I examined and recruited all other available family members. I performed all the genetic analysis and the functional studies in fibroblasts. I performed stable transfection of SH-SY5 cell lines and immunocytochemistry, in collaboration with Dr Anselm Zdebik. The WGCNA analysis was performed by Dr Conceicao Bettencourt based on the data generated by Dr Paolo Foraboschi, Dr Mina Ryten and Dr Trabzuni.

5.3.2 Introduction
I describe here the work that resulted in the identification of the genetic cause in a large British four-generation pedigree affected with autosomal dominant M-D, without mutations in SGCE. Through a combination of linkage analysis and WES, I identified a single missense mutation in KCTD17 as a novel cause for M-D.

5.3.3 Subjects, Materials and Methods

Index Family
The clinical characteristics of the patients and the family pedigree are described in the Results. Out of 19 living family members from the index family, 13 subjects were assessed and recruited. Six children of unaffected individuals were not examined. Assessment included a detailed medical interview and a full videotaped neurological examination, with focus on movement disorders. All videos were subsequently blindly reviewed by two experts in movement disorders (Professor Kailash P. Bhatia and Dr Tom Foltynie). After obtaining informed consent, a blood sample was collected and DNA was extracted from 13 family members and 4 spouses. Samples were collected with the written consent of participants and formal ethical approval by the relevant research ethics committee (UCLH Project ID number 06/N076).

Linkage Analysis
Genome-wide linkage analysis was performed in a standard manner as described in the Materials and Methods section.
**Whole-Exome Sequencing**

WES (TruSeq SBS chemistry sequenced on the Illumina HiSeq 2000) was performed in the two most distantly related affected individuals (V-3 and IV-14), as described in the Materials and Methods chapter.

**Sanger Sequencing**

Sanger sequencing was performed to validate WES results, check variants segregation in the index family and screen other M-D cases. Sanger sequencing was performed as described in Materials and Methods section. Primers were designed the entire coding region of *KCTD17* (9 exons; transcript NM_001282684.1). The primers used are listed in the Appendix.

**Selection of Additional Cases for Mutational Screening**

A cohort of M-D cases was selected to search for further confirmatory mutations in candidate genes. Cases were selected from a bank of DNA samples donated with research consent, which are held at NHNN Neurogenetics Department. Additional cases were provided by international collaborators (Prof. Thomas Gasser, University of Tubingen; Prof. Christine Klein, University of Lubeck; Prof. Nardo Nardocci, Istituto Neurologico Carlo Besta). All selected cases fulfilled the proposed diagnostic criteria for M-D (Kinugawa *et al.*, 2009). All samples included in this project had previously undergone diagnostic testing for mutations in *SGCE* (both point mutations and copy number variants) and were negative. In total 87 unrelated probands with familial M-D (defined by a family history of M-D in at least a first or second degree relative) and 358 sporadic cases of British, German and Italian origin were selected. The study was approved by the relevant local ethics committees and informed consent was provided by all participants in accordance with local guidelines.

**Expression Profiling of KCTD17 In Brain Tissue**

Brain expression profiling data was collated for *KCTD17* as described in Materials and Methods section.

**Generation of stably transfected SH-SY5 cell lines**

To assess subcellular distribution of KCTD17, a major splice variant from mouse brain was cloned. Total RNA was isolated from TRIZOL lysate, which was purified using the Zymo Direct-zol RNA mini kit (Cambridge Bioscience, UK) and quality assessed by non-denaturing agarose gel electrophoresis. Following poly-dT primed transcription
using Superscript III (Invitrogen, UK) on 5 ug total RNA, we amplified the full-length cDNA with primers containing unique restriction sites. The mutation c.434 G>A p.Arg145His was inserted by recombinant PCR. Either N- or C-terminal HA tagged wild-type and mutant cDNAs were inserted with a 1-step recombinant PCR into pcDNA3.1 constructs for expression in mammalian cells. Stable SH-SY5 cells were generated by electroporating 5 µg linearized tagged wild-type and mutant plasmids into ~1 million cells and G418 (InvivoGen) selection at 250 mg/l over at least 4 weeks and at least 6 passages. A control cell line expressing the empty vector was obtained in parallel.

**Immunocytochemistry**

Immunocytochemistry was performed on cells seeded on glass coverslips (13 mm diameter) coated with Cultrex basement membrane extract (R&D systems, USA). After fixation with either 4% PFA in PBS or ice-cold 50% methanol/50% acetone, cells were blocked in PBS+2% BSA%, 3% normal goat serum, 1% NP-40, 0.5% sodiumdesoxycholate for 30 minutes and primary antibodies. The Roche 3F10 or 12Ca5 monoclonal antibodies were used for detection of the HA tag, at 1:1000 or 1:300 dilutions, respectively. Subcellular marker antibodies were used at 1:1000 dilution. After washing, secondary detection used Alexa-dye labelled, highly cross-absorbed anti-rat (Invitrogen, UK) in 0.5x block, with 1 mg/l DAPI. Coverslips were mounted using Prolong Gold (Invitrogen, UK) after 4 washes with PBS. Microscopy was performed on a Zeiss confocal microscope.

**Weighted Gene Co-expression Network Analysis**

Weighted Gene Co-expression Network Analysis (WGCNA) was performed based on the UKBEC human brain mRNA expression data. In brief, this systems biology analytic approach uses brain regional whole-transcriptome gene expression data and establishes the degree of gene neighbourhood sharing, as defined on the basis of co-expression relationships. This approach allows to identify in an unsupervised and unbiased manner modules of genes that are highly co-expressed and co-regulated and therefore likely to be functionally related (Oldham et al., 2006). Microarray data on 19152 transcripts (corresponding to 17247 genes), generated from 788 brain samples obtained from 101 neuropathologically healthy individuals was used for weighted gene co-expression network analysis (WGCNA). The WGCNA network was constructed for each brain region to achieve a scale-free topology, as previously described (Forabosco et al.,
A dissimilarity matrix based on topological overlap measure was used to identify gene modules (i.e., densely interconnected and co-expressed genes) through a dynamic tree-cutting algorithm. A detailed description of the methods used to generate the dataset is available in the manuscript of Forabosco and colleagues.

**Calcium Imaging in Fibroblasts** 

After obtaining informed consent, fibroblasts were isolated from a skin biopsy taken from a carrier of the *KCTD17* Arg145His mutation (index family, III-2). Two unrelated age- and passage-matched controls were selected from in-house cell lines. The fibroblasts were cultured in Dulbecco’s modified Eagle’s medium GlutaMAX supplemented with 10% heat-inactivated fetal bovine serum and 1% penicillin-streptomycin. They were maintained at 37°C in a humidified atmosphere of 5% CO₂ and 95% air, media changed twice a week. Calcium homeostasis was assessed using the ratio-metric calcium (Ca²⁺) dye, Fura-2, AM (Molecular Probes, Paisley, UK), which indicates intracellular Ca²⁺ concentration and allows recordings of Ca²⁺ fluxes upon application of different pharmacological stimuli. Fibroblasts were loaded at room temperature for 30 minutes with 5mM fura-2 AM and 0.005% Pluronic in HBSS composed of (mM): 156 NaCl, 3 KCl, 2 MgSO₄, 1.25 KH₂PO₄, 2 CaCl₂, 10 glucose and 10 HEPES (pH adjusted to 7.35 with NaOH). Fluorescence measurements were obtained on an epifluorescence inverted microscope equipped with a 20x fluorite objective. Analysed areas were randomly chosen, and at least three independent experiments were performed for each condition. Ca²⁺ was monitored in single cells using excitation light provided by a Xenon arc lamp, with the beam passing monochromator centred at 340 and 380 nm (Cairn Research, Kent, UK). Emitted fluorescence light was reflected through a 515 nm long-pass filter to a cooled CCD camera (Retiga, QImaging, Surrey, BC, Canada). All imaging data were collected and analysed using software from Andor (Belfast, UK). ATP (100 mM) was used for stimulating Ca²⁺ signals in fibroblasts via purinoceptors and for releasing calcium from the endoplasmic reticulum (ER) via IP3 receptors. In a second round of experiments thapsigargin (1 mM), in Ca²⁺-free medium (plus 0.5 mM EGTA), was used for inducing the release of calcium from the ER to the cytosol and thus for estimating the size of the reticular Ca²⁺ pool. After stimulation with thapsigargin, Ca²⁺ (2mM) was added to the medium for stimulating the plasma membrane store-operated Ca²⁺ channels and assessing the store-operated calcium entry.
5.3.4 Results

Clinical phenotype of the index family

The family pedigree is shown in Figure 5-1. The 71-year-old index case (subject III-2) was first seen in clinic at age 56. She was born without complications after an uneventful pregnancy and had a normal development as a child. Her symptoms started during childhood with involuntary jerky movements of her arms. She always felt clumsier than other children. In her fifties, she developed constant head jerks and head deviation to the left. In her sixties her speech became involved, with changes of tone and stops during conversation. She also noticed that her balance was not as good as before. Involuntary movements worsened while performing voluntary actions, particularly while walking, and improved when relaxed. Touching her chin or the back of her head reduced the jerks markedly. Occasional use of alcohol did not have an impact on symptoms. On examination she had spasmodic dysphonia, perioral myoclonus and blepharospasm. She had a left torticollis and frequent irregular head myoclonic jerks. There was dystonic posturing of the hands with some writhing movements of the fingers and occasional low amplitude jerks. When she walked her trunk was slightly tilted to the left and backwards and she presented bilateral foot dystonia with feet turning inwards. She walked with a wide base and felt that walking backwards was easier. While standing she developed jerks in her legs.

A broad range of genetic investigations resulted negative, including testing for SCA 1,2,3,6 and 7, HD, DRPLA, DYT1, DYT6 and DYT11 (including sequencing of all coding exons and adjacent intronic regions and MLPA to detect copy number variants). Her brain MRI showed minor age-linked atrophy and some non-specific high signal spots in subcortical areas. She also had normal EEG, nerve conduction studies, EMG and central motor and sensory conduction times. She had a trial of oral L-DOPA (up to 600 mg/day for at least a month) without improvement. She was treated with local Botulinum Toxin injections for her cervical dystonia, with some improvement. Because of progressive worsening of her symptoms she was considered for globus pallidus internus (GPI) deep brain stimulation and, after the appropriate assessment, she was operated, at the age of 70. Following surgery her cervical dystonia improved considerably but her other symptoms, particularly her gait problems, remained unchanged.
Six other family members displayed signs of dystonia and/or myoclonus, and were accordingly categorized as affected (see Table 5-1). Age of onset of movement disorder symptoms ranged from 5 to 20 years. All affected family members initially presented with jerks or a jerky tremor, with mild dystonic features presenting later in life. All cases fulfilled the currently proposed clinical criteria for a definite diagnosis of M-D (Kinugawa et al., 2009), except individual IV-3, who upon examination displayed isolated cervical dystonia, although she reported intermittent jerky arm tremor. Myoclonus involved predominantly the arms. Dystonia predominantly affected the cranio-cervical region and upper limbs. Older individuals (>60 years; III-2 and III-5) were more severely affected and also showed laryngeal involvement. None of the affected subjects reported improvement of symptoms with alcohol. No other individuals presented with psychiatric symptoms. Case IV-12 aged 42 when clinically assessed, was considered to be unaffected with the family condition by both Professor Bhatia and Dr Foltynie. He had strabismus and benign congenital nystagmus. He showed no myoclonus or dystonia, but exhibited a mild bilateral postural hand tremor, consistent with a diagnosis of essential tremor. A similar tremor was also observed in his father, spouse of the affected individual III-5, suggesting that this feature is likely to be unrelated from the condition affecting the other family members. The remaining individuals were asymptomatic and had an entirely normal neurological examination.
Figure 5-1 Pedigree of the index family where the KCTD17 c.434 G>A p.Arg145His mutation was identified.
Table 5-1 Clinical features of cases carrying the *KCTD17* c.434 G>A p.Arg145His mutation

<table>
<thead>
<tr>
<th>Subject</th>
<th>III-2</th>
<th>III-5</th>
<th>IV-2</th>
<th>IV-3</th>
<th>IV-14</th>
<th>V-1</th>
<th>V-3</th>
<th>II.2 (Germany)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sex</td>
<td>F</td>
<td>F</td>
<td>M</td>
<td>F</td>
<td>M</td>
<td>F</td>
<td>M</td>
<td>M</td>
</tr>
<tr>
<td>Age at last examination (y)</td>
<td>70</td>
<td>65</td>
<td>46</td>
<td>44</td>
<td>35</td>
<td>20</td>
<td>19</td>
<td>62</td>
</tr>
<tr>
<td>Age at onset (y)</td>
<td>10</td>
<td>6</td>
<td>10</td>
<td>20</td>
<td>10</td>
<td>5</td>
<td>5</td>
<td>10</td>
</tr>
<tr>
<td>Presenting symptoms</td>
<td>Jerks; Arms</td>
<td>Jerks; Arms</td>
<td>Jerks; Arms</td>
<td>Jerky tremor; Arms</td>
<td>Jerky tremor; Arms</td>
<td>Jerks; Arms</td>
<td>Jerks; Arms</td>
<td>Jerks, Arms</td>
</tr>
<tr>
<td>Dystonic features</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>Involved areas</td>
<td>Cranial, Cervical, Brachial, Truncal</td>
<td>Cranial, Cervical, Brachial</td>
<td>Brachial</td>
<td>Cervical</td>
<td>Cervical, Brachial</td>
<td>Brachial</td>
<td>NA</td>
<td>Cranial, Cervical, Brachial, Truncal</td>
</tr>
<tr>
<td>Myoclonus</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Involved areas</td>
<td>Cranial, Cervical, Brachial</td>
<td>Cranial, Brachial</td>
<td>Brachial</td>
<td>NA</td>
<td>Cranial, Brachial</td>
<td>Brachial</td>
<td>Cervical, Brachial</td>
<td>Brachial</td>
</tr>
</tbody>
</table>
**Linkage analysis**

Genotype from seven affected individuals (III-2, III-5, IV-2, IV-3, IV-14, V-1 and V-3), five unaffected (III-4, IV-6, IV-8, IV-12 and V-4) and four spouses (III-1, III-6, IV-1, IV-4) was included in the analysis. The unaffected subject V-2 was not included as she was too young (17 when last examined) to exclude or confirm disease status. Linkage was performed under the assumption of an autosomal dominant inheritance with 90% penetrance and a disease allele frequency of 0.0001%. One single locus with a LOD score > 2 was identified on chromosome 22q13 (LOD score 2.4, the maximal expected value given the pedigree size; Figure 5-2). Fine mapping identified a segregating haplotype delimited by SNP markers rs926543 and rs3213584 and spanning 6.7Mb (chr22: 36989327-43716324; UCSC hg19 Genome Build), which contained 132 protein-coding genes. In addition, 5 other regions presented with uninformative multipoint LOD scores, ranging from -0.9 to +0.14 (see Table 5-2), but haplotype analysis excluded co-segregation of these regions with the disease.

<table>
<thead>
<tr>
<th>Chromosome</th>
<th>Start position (Hg19)</th>
<th>Stop position (Hg19)</th>
<th>LOD score</th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td>100410724</td>
<td>101569726</td>
<td>-0.8111</td>
</tr>
<tr>
<td>12</td>
<td>58842292</td>
<td>60028954</td>
<td>0.1474</td>
</tr>
<tr>
<td>12</td>
<td>70112707</td>
<td>70517723</td>
<td>-0.7314</td>
</tr>
<tr>
<td>12</td>
<td>70422480</td>
<td>71099936</td>
<td>-0.9131</td>
</tr>
<tr>
<td>12</td>
<td>72641784</td>
<td>73373089</td>
<td>-0.7143</td>
</tr>
<tr>
<td>22</td>
<td>36989327</td>
<td>43716324</td>
<td>2.4082</td>
</tr>
</tbody>
</table>
Figure 5-2 Linkage analysis results

LOD score plot for genome-wide linkage analysis in the British index pedigree showing a single linkage peak on chromosome 22q13 with a maximum LOD score of 2.4. An autosomal dominant model was specified with an estimated allele frequency of 0.00001 and 90% penetrance. The chromosome is indicated across the top of the plot; chromosomal distance in centiMorgans (cM) is indicated on the x-axis; and the LOD score is indicated on the y-axis.
Whole-exome sequencing

To identify the disease-causing variant, WES was performed in the two most distantly related affected individuals (V-3 and IV-14). In total, 83,572,847 (V-3) and 81,527,162 (IV-14) unique reads were generated. According to the Consensus Coding Sequences hg19 definition of the ‘TruSeq exome’, the average read depth of both exomes was > 70, > 95% of the target bases were covered at a read depth of 2x and > 90% at a depth of 10x. A total of 22,857 (V-3) and 22,946 (IV-14) exonic/splicing variants were detected. A summary of the WES metrics can be found in Table 5-3.

| Table 5-3 Summary of results from exome sequencing |
|---------------------------------|--------|--------|
| Patient                         | V-3    | IV-14  |
| Unique reads                    | 83,572,847 | 81,527,162 |
| Aligned Reads (%)               | 87.6   | 87.2   |
| Mean depth                      | 73     | 71     |
| Target covered at least 2x      | 96     | 95     |
| Target covered at least 10x     | 92     | 91     |
| Total variants                  | 22,857 | 22,946 |
| Excluding synonymous variants   | 11,710 | 11,737 |
| Novel variants                  | 154    | 140    |
| Shared novel variants           | 4      | 4      |
| Shared novel variants within the linkage peak on 22q | 1 | 1 |

All synonymous changes and those not shared by the two affected individuals were filtered out. Then, under the assumption that the mutation causing this rare autosomal dominant disease is extremely rare and not present in the general population, I also excluded variants that are present in databases of sequence variations publicly available databases of sequence variations (dbSNP version 129, 1000 Genomes project, NHLBI Exome Variant Server and Complete Genomics 69). Furthermore, I excluded variants found in our own in-house exomes (n=200) from individuals with unrelated diseases.
After applying filtering criteria, only four novel missense variants shared by the two affected individuals were left (see Table 5-4): c.10976 C>T; p.Ser3659Phe in FLG (filaggrin; RefSeq NM_002016.1), c.1055 T>G; p.Phe352Cys in OBSCN (obscurin; RefSeq NM_052843.3), c.1076 A>C; p.Lys359Thr in LRRC6 (leucine rich repeat containing 6; RefSeq NM_012472.4) and c.434 G>A; p.Arg145His in KCTD17 (potassium channel tetramerisation domain containing 17; RefSeq NM_001282684.1).

Of these variants, only the missense change in KCTD17 was located within the linked chromosomal locus on chromosome 22q. No shared rare copy number variants in exome sequencing data were detected using the Exome depth algorithm.

Sanger sequencing of the KCTD17 variant (see Figure 5-3) in all available family members confirmed perfect co-segregation of the variant with the disease-phenotype, being the nucleotide change present in all affected individuals and absent in all unaffected (including subject V-2, initially excluded from the linkage analysis). Segregation analysis of the other three novel variants, shared by individuals V-3 and IV-14 but located outside the identified linkage peak, was also performed. As expected, none of them fully segregated: the LRRC6 variant was not carried by affected individuals (IV-2 and V-1) and the OBSCN and FLG variants were carried by unaffected subjects V-4 (OBSCN and FLG) and IV-12 (FLG), aged 20 and 40 respectively when last examined. Although the KCTD17 p.Arg145His substitution falls in a functionally uncharacterized portion of the protein, it lies in an extremely conserved amino acid motif, not only completely conserved down to invertebrate species, but also identical in the KCTD17 human paralogs KCTD2 and KCTD5 (see Figure 5-4).
Figure 5-3 Sanger sequencing confirmation of the \textit{KCTD17} c.434G>A p.Arg145His mutation.

The variant is absent in over 3,700 individuals of European origin without movement disorders, who were exome sequenced by the UCL-exomes consortium, and in a further > 61,000 individuals listed in the Exome Aggregation Consortium database (last accessed in March 2016). All \textit{in silico} tools consistently predicted a deleterious effect of the substitution (see Table 5-4).

\textbf{Figure 5-4 Interspecies protein sequence alignment}

Multiple-sequence alignment showing complete conservation of protein sequence across species and human paralogs (KCTD2 and KCTD5) in the region of KCTD17 where the disease-segregating mutation p.Arg145His was identified. The amino acid involved by the mutation is marked in red. The residues that are not completely conserved are highlighted in yellow. In blue the residues completely conserved.
Table 5-4 Summary of novel variants detected by whole-exome sequencing and shared by individuals V-3 and IV-14.

<table>
<thead>
<tr>
<th>Chr</th>
<th>Position (hg19)</th>
<th>Gene (Transcript)</th>
<th>Variant</th>
<th>GERP score&lt;sup&gt;a&lt;/sup&gt;</th>
<th>CADD C-score&lt;sup&gt;b&lt;/sup&gt;</th>
<th>SIFT</th>
<th>Provean</th>
<th>PolyPhen-2 HumVar</th>
<th>Mutation Taster</th>
<th>Gene previously associated with disease?</th>
<th>Linkage analysis (LOD score)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>152276386</td>
<td>FLG (NM_002016.1)</td>
<td>c.10976 C&gt;T; p.Ser3659Phe</td>
<td>2.48</td>
<td>11.15</td>
<td>T (0.06)</td>
<td>N (-0.8)</td>
<td>D (0.78)</td>
<td>P (0.99)</td>
<td>Yes, skin diseases (e.g. ichthyosis vulgaris, and/or eczema)</td>
<td>&lt; -2</td>
</tr>
<tr>
<td>1</td>
<td>228401208</td>
<td>OBSCN (NM_052843.3)</td>
<td>c.1055 T&gt;G; p.Phe352Cys</td>
<td>5.58</td>
<td>21</td>
<td>D (0)</td>
<td>D (-3.8)</td>
<td>D (0.94)</td>
<td>D (0.99)</td>
<td>Hypertrophic cardiomyopathy</td>
<td>&lt; -2</td>
</tr>
<tr>
<td>8</td>
<td>133622476</td>
<td>LRRC6 (NM_012472.4)</td>
<td>c.1076 A&gt;C; p.Lys359Thr</td>
<td>3.5</td>
<td>15.74</td>
<td>D (0.01)</td>
<td>D (-3.1)</td>
<td>B (0.39)</td>
<td>D (0.99)</td>
<td>Recessive primary ciliary dyskinesia</td>
<td>&lt; -3</td>
</tr>
<tr>
<td>22</td>
<td>37453460</td>
<td>KCTD17 (NM_001282684.1)</td>
<td>c.434 G&gt;A; p.Arg145His</td>
<td>4.46</td>
<td>28.8</td>
<td>D (0)</td>
<td>D (-4.8)</td>
<td>D (0.53)</td>
<td>D (0.99)</td>
<td>No</td>
<td>2.4</td>
</tr>
</tbody>
</table>

B=beneic; D=deleterious/damaging/disease-causing; N=neutral; P=polymorphism; T=tolerated

<sup>a</sup>Positive scores represent a substitution deficit and indicate that a site may be under evolutionary constraint. Negative scores indicate that a site is probably evolving neutrally. Positive scores scale with the level of constraint, such that the greater the score, the greater the level of evolutionary constraint inferred to be acting on that site.

<sup>b</sup>C-scores greater or equal 10 indicates that the variant is predicted to be the among the 10% most deleterious substitutions that you can do to the human genome; a score of greater or equal 20 indicates the 1% most deleterious
Sanger Sequencing of KCTD17 in a Cohort of Phenotypically Similar Myoclonus-Dystonia Cases

The entire coding region of *KCTD17* was subsequently sequenced by Sanger sequencing in a further 87 unrelated probands with familial SGCE-negative M-D. Exon 4 only (containing the c.434 G>A mutation) was sequenced in a further 358 sporadic M-D cases. This analysis revealed the presence of the same *KCTD17* mutation, c.434 G>A p.Arg145His, in the index case of a German family with autosomal dominant M-D (Figure 5-5). No further pathogenic mutations in KCTD17 were identified.

![Figure 5-5 Pedigree of the German families with the KCTD17 c.434 G>A p.Arg145His mutation.](image)

The clinical presentation of this case closely resembled that of III-2 and III-5, the older affected subjects from the British family. He reported arm jerks and difficulty writing, starting in childhood. Right torticollis and a jerky head tremor appeared around age 40, becoming progressively debilitating. There was no response to alcohol or psychiatric comorbidities. He underwent surgery for bilateral pallidal deep brain stimulation at age 58, which resulted in marked improvement of cervical dystonia and myoclonus of the upper limbs. Clinical examination at age 62 showed generalized dystonia, with prominent cranio-cervical involvement, and myoclonic jerks involving the upper limbs.

His father was also affected with a movement disorder, presenting with perioral dyskinesia in his forties. The proband’s brother had M-D, with similar clinical features, including generalized jerks, cervical dystonia and dysarthria. Unfortunately, DNA
samples of the deceased father and brother were not available for segregation analysis. The 25-years old proband’s only son, who had no signs upon examination, refused genetic testing.

To determine whether the KCTD17 c.434 G>A; p.Arg145His occurred on the same haplotype in the British and in the German kindreds, SNP markers located 0.5 Mb up- and down-stream the mutation were compared (see Table 5-5). In the British family, the haplotype of the identified KCTD17 mutation was reconstructed. The German case was genotyped using the same array. In the German case SNP phasing was possible only for homozygous alleles, due to lack of genetic data on other family members. Different alleles were located at markers rs5756477 and rs228924 (45.9 kb upstream and 53 kb downstream of KCTD17, respectively), delimitating a small region of approximately 99 kilobases (Kb) of a possibly shared haplotype. Further analysis with a highly polymorphic microsatellite (19xAG; chromosomal position 22:37,446,300-37,446,339), located only 1.4 Kb upstream of the 5’ end of the KCTD17 open reading frame, revealed that the two pedigrees had different alleles.

<table>
<thead>
<tr>
<th>Marker</th>
<th>Chromosomal position (hg19)</th>
<th>Genotype UK family</th>
<th>Genotype German family</th>
</tr>
</thead>
<tbody>
<tr>
<td>rs5756370</td>
<td>37242476</td>
<td>A</td>
<td>A</td>
</tr>
<tr>
<td>rs6000449</td>
<td>37251377</td>
<td>A</td>
<td>A</td>
</tr>
<tr>
<td>rs4821542</td>
<td>37252918</td>
<td>G</td>
<td>G</td>
</tr>
<tr>
<td>rs909483</td>
<td>37260474</td>
<td>A</td>
<td>A</td>
</tr>
<tr>
<td>rs2413429</td>
<td>37289869</td>
<td>G</td>
<td>A</td>
</tr>
<tr>
<td>rs4821558</td>
<td>37308785</td>
<td>G</td>
<td>A</td>
</tr>
<tr>
<td>rs11705394</td>
<td>37329676</td>
<td>A</td>
<td>A</td>
</tr>
<tr>
<td>rs9622506</td>
<td>37338286</td>
<td>A</td>
<td>G</td>
</tr>
<tr>
<td>rs8137446</td>
<td>37347959</td>
<td>G</td>
<td>G</td>
</tr>
<tr>
<td>rs9622521</td>
<td>37350881</td>
<td>G</td>
<td>G</td>
</tr>
<tr>
<td>rs4821576</td>
<td>37357169</td>
<td>G</td>
<td>G</td>
</tr>
<tr>
<td>-----------------</td>
<td>---------</td>
<td>----</td>
<td>----</td>
</tr>
<tr>
<td>rs8142593</td>
<td>37363121</td>
<td>A</td>
<td>A</td>
</tr>
<tr>
<td>rs877166</td>
<td>37369148</td>
<td>C</td>
<td>C</td>
</tr>
<tr>
<td>rs5756437</td>
<td>37375668</td>
<td>G</td>
<td>G</td>
</tr>
<tr>
<td>rs1157557</td>
<td>37381674</td>
<td>G</td>
<td>G</td>
</tr>
<tr>
<td>rs5756477</td>
<td>37407527</td>
<td>G</td>
<td>A</td>
</tr>
<tr>
<td>rs5756492</td>
<td>37424991</td>
<td>G</td>
<td>G</td>
</tr>
<tr>
<td>Microsatellite 19xAG</td>
<td>37446300</td>
<td>17&quot;</td>
<td>18/14&quot;</td>
</tr>
<tr>
<td>KCTD17 c.434G&gt;A</td>
<td>37453460</td>
<td>A</td>
<td>A</td>
</tr>
<tr>
<td>rs2160906</td>
<td>37493178</td>
<td>G</td>
<td>G</td>
</tr>
<tr>
<td>rs228924</td>
<td>37507250</td>
<td>A</td>
<td>G</td>
</tr>
<tr>
<td>rs11914132</td>
<td>37509087</td>
<td>G</td>
<td>G</td>
</tr>
<tr>
<td>rs228942</td>
<td>37524619</td>
<td>C</td>
<td>C</td>
</tr>
<tr>
<td>rs3218258</td>
<td>37544245</td>
<td>A</td>
<td>G</td>
</tr>
<tr>
<td>rs229483</td>
<td>37553619</td>
<td>G</td>
<td>A</td>
</tr>
<tr>
<td>rs12167757</td>
<td>37567490</td>
<td>G</td>
<td>G</td>
</tr>
<tr>
<td>rs229518</td>
<td>37577872</td>
<td>A</td>
<td>A</td>
</tr>
<tr>
<td>rs11913300</td>
<td>37580627</td>
<td>A</td>
<td>A</td>
</tr>
<tr>
<td>rs5756540</td>
<td>37582205</td>
<td>G</td>
<td>G</td>
</tr>
<tr>
<td>rs5756546</td>
<td>37589805</td>
<td>G</td>
<td>G</td>
</tr>
<tr>
<td>rs64547</td>
<td>37592504</td>
<td>A</td>
<td>A</td>
</tr>
<tr>
<td>rs9610680</td>
<td>37621951</td>
<td>A</td>
<td>G</td>
</tr>
<tr>
<td>rs8137698</td>
<td>37624236</td>
<td>G</td>
<td>A</td>
</tr>
<tr>
<td>rs739042</td>
<td>37625419</td>
<td>G</td>
<td>G</td>
</tr>
<tr>
<td>----------</td>
<td>----------</td>
<td>---</td>
<td>---</td>
</tr>
<tr>
<td>rs2285110</td>
<td>37628145</td>
<td>G</td>
<td>G</td>
</tr>
<tr>
<td>rs9607431</td>
<td>37629938</td>
<td>C</td>
<td>A</td>
</tr>
<tr>
<td>rs5995404</td>
<td>37632938</td>
<td>C</td>
<td>C</td>
</tr>
</tbody>
</table>

All alleles where the haplotype of the UK family differs from that of the German family are highlighted in yellow. *These values indicate the number of AG repeats*

**Brain expression analysis**

*KCTD17* gene expression was evaluated in 10 brain regions from 134 normal individuals using our in-house data from the UK Brain Expression Consortium. *KCTD17* mRNA expression throughout the course of human brain development was assessed using the data available in the Human Brain Transcriptome (HBT) database. *KCTD17* mRNA expression was high across all brain regions, but it was highest in the putamen followed by the thalamus (Figure 5-6A). There was a 3.2-fold difference (p value < 3.5<sup>-71</sup>) between the striatum and the cerebellum, the regions with respectively the highest and lowest expression. These findings are consistent with the data available in the HBT database, showing increasing *KCTD17* brain mRNA levels in the striatum and the thalamus from early midfetal development (between 21 and 25 post-conception weeks of age) to adolescence (between 12 and 20 years of age) in the striatum and the (Figure 5-6B).
Figure 5-6 *KCTD17* brain regional mRNA expression data

(A) Box plot of mRNA expression levels for *KCTD17* in 10 adult brain regions, based on exon array experiments and plotted on a log2 scale (y axis). The plot shows significant variation in *KCTD17* transcript expression across the 10 CNS regions analysed. “N” indicates the number of brain samples analysed to generate the results for each CNS region. KCT17 expression is higher in the putamen, followed by the thalamus. Whiskers extend from the box to 1.53 the interquartile range. (B) Graph to show mRNA expression levels for *KCTD17* in 6 brain regions during the course of human brain development, based on exon array experiments and plotted on a log2 scale (y axis).
Immunocytochemistry

Stably transfected SH-SY5 cells were generated by incorporating either N- or C-terminally HA-tagged wild-type and mutant KCTD17 cDNAs. KCTD17 staining with anti-HA primary monoclonal antibodies showed that the protein is diffusely distributed in the cytosol with fine reticular pattern and does not localize at the plasma membrane (Figure 5-7). We did not observe significant changes in subcellular localization of mutant versus wild-type KCTD17, indicating that the amino acid substitution does not lead to cellular mislocalisation of the protein.

![Wild-type](image1.png) ![Mutant p.Arg145His](image2.png)

**Figure 5-7 Immunocytochemistry in stably transfected SH-SY5 cells**
These experiments show no difference between wild-type and mutant KCTD17 subcellular localization. HA-tagged KCTD17 is shown in green. Cellular nuclei are stained with DAPI and are shown in blue.

Weighted Gene Co-expression Network Analysis

KCTD17, in common with other dystonia genes (e.g. ANO3, GNAL and HPCA), shows the highest expression in the putamen and this brain structure has an established role in the pathogenesis of dystonia. I therefore focused the analysis on the putamen module including KCTD17.

This module contains 179 transcripts (equating to 172 genes; see Figure 5-8 for a graphic representation of the module).
Figure 5-8 Network representation of the putamen KCTD17-containing gene module

Shown are all genes in the putamen KCTD17-containing module connected with a topological overlap measure exceeding 0.03. The dystonia genes in the module (KCTD17 and HPCA) and all the direct connections of KCTD17, based on topological overlap values, are highlighted in red. Larger circles represent the most interconnected genes in the module, including KCTD17.

I first assessed if the module was enriched for genes linked to Mendelian forms of dystonia. The analysis focused on the nine genes known to be associated with dystonia (TOR1A, THAP1, SGCE, TUBB4A, CIZ1, ANO3, GNAL, ATP1A3, HPCA). I did not include in the analysis genes causing DOPA-responsive dystonia (GCH1, TH, and SPR) as their established functional role in nigrostriatal dopamine synthesis, together with the specificity of the clinical presentation, clearly identifies them as a separate entity.

Importantly, the putamen KCTD17-module showed significant clustering of dystonia genes (KCTD17 and HPCA; Fisher’s exact test $P = 5 \times 10^{-3}$), suggesting the relevance of this gene network to the molecular pathogenesis of dystonia. The module was poorly preserved across other brain regions, indicating its specificity to the putamen (Figure 5-9).
Module preservation statistics were calculated (z score) to assess how well modules from one tissue are reproducible (or preserved) in another brain region (Langfelder et al., 2011). Previously proposed thresholds were considered (z score of <2 indicates no evidence of module preservation, z score between 2 and <10 indicates weak to moderate evidence, and z score of ≥10 indicates strong evidence). The module is poorly conserved across other brain regions, indicating its specificity to the putamen.

To infer the biological and functional relevance of the putamen KCTD17 gene network, functional annotation enrichment analysis was then carried out using the online tool g:Profiler (Reimand et al., 2011). This analysis allowed the identification of over-represented genes assigned to specific Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways, namely “Circadian entrainment” (KEGG:04713; g:Profiler’s custom threshold $P = 5.13 \times 10^{-3}$) and “Dopaminergic synapse” (KEGG:04728; g:Profiler’s custom threshold $P = 2.71 \times 10^{-2}$). This suggests the involvement of the genes in the module with these molecular pathways.

**Ca\textsuperscript{2+} imaging in Patient Fibroblasts**

Disruption of calcium (Ca\textsuperscript{2+}) homeostasis has been recently implicated in the pathogenesis of several genetic forms of dystonia (eg. TOR1A, ANO3, HPCA) (Charlesworth et al., 2012, Iwabuchi et al., 2013, Charlesworth et al., 2015). As the putamen KCTD17-module included HPCA, a gene with an established role in intracellular Ca\textsuperscript{2+}-dependent signalling (Palmer et al., 2005), I hypothesized that the
KCTD17 p.Arg145His substitution may have a significant impact on intracellular Ca\(^{2+}\) homeostasis.

For this purpose, fibroblasts were isolated from a skin biopsy taken from a subject with the KCTD17 p.Arg145His mutation and two unrelated age- and passage-matched controls. The expression of KCTD17 in fibroblasts was confirmed by RT-PCR (data not shown). I observed that stimulation with ATP (50μM), which stimulates P2Y receptors and releases Ca\(^{2+}\) from the ER via IP\(_3\) receptors, resulted in significantly reduced and delayed cytosolic Ca\(^{2+}\) signal (one-way ANOVA \(P < 0.01\); Figure 5-10A and B) in cells carrying the p.Arg145His mutation when compared to both control cells, indicating a smaller calcium pool within the ER.

To further prove this finding, a second round of experiments using thapsigargin (1μM) in Ca\(^{2+}\)-free medium (plus 0.5 mM EGTA) was subsequently carried out. Thapsigargin is an inhibitor of the ER calcium ATPase (SERCA) and induces the release of calcium from the ER to the cytosol, allowing an estimation of the ER Ca\(^{2+}\)-pool. Ca\(^{2+}\) was then added at the end of the experiment to stimulate elevation of cytosolic Ca\(^{2+}\) through opening of store operated calcium channels. Thapsigargin stimulation resulted in a significantly smaller Ca\(^{2+}\)signal in fibroblasts bearing the p.Arg145His mutation when compared to controls (one-way ANOVA \(P < 0.01\); Figure 5-10C and D), confirming that the Ca\(^{2+}\) pool in the ER of mutation-carrying fibroblasts is reduced.

Furthermore, stimulation of the store-operated Ca\(^{2+}\) channels induced a smaller Ca\(^{2+}\) influx in mutated fibroblasts (Figure 5-10C), possibly suggesting an insufficient Ca\(^{2+}\) influx across the plasma membrane in response to the fall in Ca\(^{2+}\) concentration within the ER lumen.
Figure 5.10 Functional studies showing abnormalities of endoplasmic reticulum calcium signalling in KCTD17 p.Arg145His substitution bearing fibroblasts

All experiments were carried out in triplicate. Data are represented as the mean ± SEM. “n” indicates the total number of cells analysed. The asterisks indicate P < 0.05 (*), P < 0.01 (**) and P < 0.001 (***) (A) Typical trace of [Ca²⁺]c in control and KCTD17-mutant fibroblasts in response to the application of 50 nM ATP. (B) Histograms showing a significantly decreased [Ca²⁺]c response upon ATP stimulation in mutation-bearing fibroblasts (n=56) versus controls (control 1 n=41; control 2 n=35), as measured by changes in Fura-2 fluorescence intensity. (C) Typical trace of [Ca²⁺]c in control and KCTD17-mutant fibroblasts in response to the application of thapsigargin (1 µM), and subsequent Ca²⁺ challenge (2 mM). (D) Histograms demonstrating a significant reduction in ER calcium pool in response to thapsigargin in mutation-bearing fibroblasts (n=53) versus controls (control 1 n=51; control 2 n=65), as measured by changes in Fura-2 fluorescence intensity.

5.3.5 Discussion

Several lines of evidence support the pathogenic role of the KCTD17 p.Arg145His.

1) In the discovery pedigree, the p.Arg145His variant was located within the only segregating chromosomal locus and Sanger sequencing confirmed perfectly cosegregation with dystonia. This was not the case for any other variant identified by WES that survived the filtering strategy. These latter variants were either not carried by definitely affected subjects (LRRC6) or carried by clinically unaffected subjects (FLG and OBSCN). Given the high penetrance of the disease indicated by the pedigree structure (no obligate unaffected carriers observed over four generations) and the onset of symptoms in the first decade in all affected individuals, we consider extremely
unlikely the possibility of the causative mutation for this pedigree being carried by unaffected adults. Furthermore the identified FLG and OBSCN variants are located in large and polymorphic genes tolerating a high burden of genetic variation (2560 and 4135 rare [MAF < 0.01] non-synonymous variants spread across FLG and OBSCN, respectively, are listed in the ExAC database versus 71 rare variants observed in KCTD17), arguing against a pathogenic role of these changes. Moreover, pathogenic variants in FLG and OBSCN have been previously linked to diseases unrelated to dystonia, respectively skin diseases (Irvine et al., 2011) and cardiomyopathy (Marston et al., 2015).

2) A screening of the KCTD17 gene in a cohort of SGCE-negative M-D cases revealed the same nucleotide change in the index case of an additional unrelated pedigree with an autosomal dominant pattern of inheritance. In this second pedigree I could not prove segregation due to unavailability of the other affected family members. However, given the striking clinical similarity between this case and the affected individuals of the index family, together with the absence of the variant in over 65,000 internal and publicly available controls, it is highly unlikely this variant could merely represent a private benign polymorphism.

Importantly, I demonstrated that the two families with the p.Arg145His mutation did not share a common founder haplotype and therefore two independent mutational events must have occurred. This is relevant for two reasons; firstly, I can exclude that the KCTD17 p.Arg145His variant is in linkage with the actual causative mutation but it is not itself pathogenic. Secondly, were the two families to be linked by an ancestral common founder, given the high penetrance of the mutation, I would expect a much higher prevalence of the p.Arg145His mutation as a cause of familial M-D across the British and German population. This is the case, for instance, for the recurrent GOSR2 G144W mutation, a widespread cause of progressive myoclonus-ataxia syndrome in Nordic populations (Boisse Lomax et al., 2013).

3) KCTD17 is abundantly expressed in all brain regions and the highest expression was observed in putamen and thalamus. In light of the current view about the neuroanatomical bases of dystonia, seen as a network disorder originating from disrupted circuits involving the sensory-motor cortex, the basal ganglia and the cerebellum (Neychev et al., 2011), this pattern of expression is highly relevant and
supports the pathogenic role of *KCTD17* in the pathogenesis of M-D. Furthermore, focal ischemic lesions of the thalamic subnuclei have been causally linked to the development of a movement disorder that closely resembles M-D (Lehericy *et al.*, 2001). This strengthens the relevance of high *KCTD17* expression in the thalamus and supports the hypothesis of a crucial role for disrupted thalamic pathways in the pathogenesis of *KCTD17*-associated M-D.

*KCTD17* encodes for a member of a recently identified family of 26 closely related and highly conserved proteins, the potassium channel tetramerisation domain (KCTD)-containing proteins. KCTD proteins are characterized by the presence of a N-terminal bric-a-brack, tram-track, broad complex/poxvirus zinc finger (BTB/POZ) domain, homologous to the cytoplasmic domain T1 of voltage-gated potassium channels (Liu *et al.*, 2013). The BTB/POZ domain is known to permit protein-protein interactions, either promoting self-oligomerisation or facilitating interaction with other biological partners (Stogios *et al.*, 2005). KCTDs are small soluble proteins, which are not predicted by their structure to form transmembrane domains. Despite the homology reflected in their names, a direct interaction with potassium channels has not been shown for most members of the family and was explicitly excluded for KCTD5, a paralog 85% identical to KCTD17 (Dementieva *et al.*, 2009).

KCTD proteins, despite the high level of sequence similarity, are involved in a surprisingly wide spectrum of cell functions, including regulation of cellular proliferation, gene transcription, cytoskeleton organization, protein degradation targeting via the ubiquitin-proteasome system, and regulation of G protein-coupled receptors (Skoblov *et al.*, 2013). Several members of the KCTD family have a primary role in the central nervous system (Schwenk *et al.*, 2010, Matsui *et al.*, 2013) and a growing number of neurological diseases have been linked to mutations in KCTD genes. *KCTD7* (MIM 611725) mutations cause recessive progressive myoclonic epilepsy (Van Bogaert *et al.*, 2007, Kousi *et al.*, 2012, Krabichler *et al.*, 2012). Copy number variants in *KCTD13* (MIM 60894) have been associated with size of the head, autism disorder and epilepsy (Golzio *et al.*, 2012). More recently a homozygous missense mutation in *KCTD3* (MIM 613272) was identified as the likely cause in a pedigree with severe psychomotor retardation, seizure, and cerebellar hypoplasia (Alazami *et al.*, 2015).
The precise cellular localization and function of the KCTD17 protein are largely unknown. Recent work has shown that KCTD17 contributes to the ubiquitin-proteasome machinery, acting as an adaptor for the CUL3-RING E3 ligase and targeting substrates for degradation through poly-ubiquitinylation (Kasahara et al., 2014). Although most of the KCTD17-CUL3 substrates are currently unknown, CUL3 has been implicated in the elaboration of dendrite branching and neurite terminal morphogenesis in drosophila models (Zhu et al., 2005, Djagaeva and Doronkin, 2009).

A number of cellular pathways, encompassing abnormalities of the cell-cycle and transcriptional regulation, endoplasmic reticulum and nuclear envelope function and control of synaptic function, are thought to underlie molecular pathogenesis of primary genetic dystonias (Ledoux et al., 2013).

How KCTD17 abnormal function could contribute to dystonia disease mechanisms is currently not known. WGCNA analysis indicates that KCTD17 may closely interact with HPCA, another recently identified dystonia-gene that is involved in calcium signalling (Charlesworth et al., 2015). Through live-cell imaging experiments in patients fibroblast I showed that the presence of the KCTD17 p.Arg145His mutation reduces Ca\(^{2+}\) storage in the ER and opening of store-operated calcium channels. Importantly, our group recently showed very similar defects of ER Ca\(^{2+}\) storage in fibroblasts bearing a pathogenic mutation in ANO3, another dystonia gene (Charlesworth et al., 2012). This indicates that defective ER calcium signalling may represent a converging pathogenic mechanism in genetically unrelated forms of dystonia.

WGCNA also showed that the putamen KCTD17 module is enriched with genes involved in dopaminergic signalling, suggests the involvement of the genes in the module with these molecular pathway. Recent work in drosophila strongly reinforces the results of WGCNA and further suggests a relevant contribution of KCTD17 to regulation of dopaminergic transmission in the putamen. Insomniac (the KCTD17 fly homolog) is an essential regulator of sleep homeostasis through the control of the dopaminergic arousal pathways (Stavropoulos and Young, 2011, Pfeiffenberger and Allada, 2012). More specifically, insomniac seems to regulate dopaminergic signalling at the post-synaptic level, possibly controlling the turnover of dopamine receptors or their downstream effectors (Pfeiffenberger and Allada, 2012).
Importantly, abnormal post-synaptic dopaminergic signalling in the basal ganglia is one of the main themes in molecular dystonia pathogenesis, a concept recently strengthened by the identification of mutations in GNAL causing dystonia (Fuchs et al., 2013). Fitting well with this model, all the genes in the putamen KCTD17 module assigned to the KEGG “Dopaminergic synapse” pathway (CACNA1C, PPP1R1B, AKT1, GNAO1, and GNB2) localize and act at the post-synaptic level.

The KCTD17/HPCA putaminal module is not preserved across brain regions, which indicates it is a putamen-specific gene network. Importantly, the brain regional specificity of this module may suggest why mutations in KCTD17 and HPCA manifest purely as a dysfunction of the basal ganglia (i.e. dystonia), in spite of the ubiquitous expression in the human brain.

Finally, the molecular pathogenic mechanism whereby the identified KCTD17 p.Arg145His mutation could lead to M-D is still unclear. Our data shows that the protein stability or subcellular localization is unaltered in presence of the p.Arg145His. The substitution falls outside the BTB/POZ domain. However it lies in an extremely highly conserved amino acid motif, which is not only completely conserved throughout all species down to Drosophila, but is also identical in human paralogous genes (KCTD2 and KCTD5), suggesting an essential role for the protein function.

Notably, recent studies highlighted that tetramerisation is not a prerogative of the BTB/POZ domain but also the C-terminal domain of KCTD proteins may be able to promote tetramers assembly (Correale et al., 2013). Hence, it is possible that p.Arg145His mutation might prevent a correct oligomerisation or interaction with specific protein partners. Alternatively, the KCTD17 p.Arg145His substitution could abrogate the interaction with CUL3, resulting in abnormal degradation of specific substrates via the ubiquitine-proteasome system. For instance, this mechanism has been recently shown for a KCTD7 pathogenic mutations located outside the BTB/POZ domain (Staropoli et al., 2012).

In conclusion, through a combination of genome-wide linkage analysis and WES in large British kindred, I showed that the missense mutation p.Arg145His in KCTD17 represents a novel genetic cause for inherited autosomal dominant M-D. The clinical features of the KCTD17-mutated cases, although fully consistent with a clinical diagnosis of M-D, were distinct in many ways from the usual phenotype of subjects.
with *SGCE* mutations. Dystonia dominated the clinical picture and showed a progressive course, worsening over time and spreading to other sites (including speech involvement), a course unusual for *SGCE*-related M-D. Myoclonus, despite being the presenting symptom in most cases, was overall mild and not as disabling as in *SGCE*-mutated subjects. These phenotypic differences may be explained by the different functions of the two genes, but also by the clearly distinct patterns of brain regional expression. *SGCE* is highly expressed in the cerebellum, whereas its expression is low to moderate in putamen and globus pallidus (Ritz et al., 2011). On the other hand, *KCTD17* expression is high in the putamen and thalamus but relatively low in the cerebellum. Intriguingly, this could be the explanation for the scarce response to alcohol consumption in *KCTD17*-mutated cases, as alcohol probably exerts its beneficial effect in M-D secondary to *SGCE* mutations by modulating cerebellar activity (Ritz et al., 2011).

Preliminary data suggest an involvement of KCTD17 in dopamine synaptic transmission regulation and an effect of the p.Arg145His substitution on ER-derived Ca$^{2+}$ signalling. Further insight into the physiological role of KCTD17 and a better understanding of the pathogenic effect of the p.Arg145His substitution will shed light onto the mechanisms leading to abnormal neuronal activity underlying M-D. Furthermore, the identification of KCTD17 interactors will possibly highlight new potential pharmacological targets for the treatment of dystonia.

Mutational screening of additional cohorts of M-D cases will help to define the frequency and the spectrum of *KCTD17* mutations. *KCTD17* mutations should be considered in cases without mutations in *SGCE* presenting with myoclonus, dystonia or a combination of both, particularly if there is predominant cranio-cervical and laryngeal involvement.
5.4 Tyrosine Hydroxylase Deficiency Causes Autosomal Recessive, DOPA-Responsive, Myoclonus-Dystonia

5.4.1 Statement of Contribution
Dr Stamelou and I examined the family described in this chapter. I performed all the genetic analysis.

5.4.2 Background
I describe here the genetic analysis I conducted in a second, highly informative, family with SGCE-negative M-D. While the pattern of transmission was autosomal-dominant in the pedigree described in the previous chapter, in this family the disease was clearly transmitted in an autosomal recessive fashion (see Figure 5-11); three out of four siblings presented with a severe, early-onset movement disorders, while both parents were undoubtedly unaffected.

An important clinical clue toward the identification of the underlying genetic abnormality in this family was the reported L-DOPA-responsiveness of symptoms in all affected subjects, which raised the suspicion of an atypical form of DRD. As discussed in the Introduction chapter, mutations in four genes have been convincingly demonstrated to cause DRD: GCH1 (GTP Cyclohydrolase 1), TH (Tyrosine Hydroxylase) and SPR (Sepiapterin Reductase) and DDC (DOPA decarboxylase).

Although M-D has only been described once in a single family with DRD due to GTP Cyclohydrolase 1-deficiency (Leuzzi et al., 2002), given the strong clinical suspicion of a disorder of the dopamine synthesis pathway, after the exclusion of pathogenic SGCE mutations, I sequenced the coding regions of the genes mentioned above and identified compound-heterozygous mutations in TH.

5.4.3 Subjects, Materials and Methods

Index family
The clinical characteristics of the patients and the family pedigree are described in the Results. All living family members from the index family were assessed and recruited. Two children of an unaffected individual were not examined. Assessment included a detailed medical interview and a full videotaped neurological examination, with focus on movement disorders. The University College London Hospitals/University College
London ethics committee (University College London Hospitals project ID number 06/N076) approved the study, and all family members gave their informed consent. After obtaining informed consent, a blood sample was collected and DNA was extracted from all examined family members.

![Figure 5-11 Pedigree of a family with recessive L-DOPA responsive Myoclonus-Dystonia](image)

**Genetic analysis**

Sanger sequencing was performed as described in Materials and Methods section. Primers were designed to amplify all coding exons and splicing sites of the *GCH1* (ENST00000491895; 6 coding exons), *TH* (ENST00000381178; 14 coding exons), *SPR* (ENST00000234454; 3 coding exons), and *DDC* (ENST00000444124; 14 coding exons). The primers used in this study are listed in the Appendix.

**5.4.4 Results**

**Clinical description**

Patient II-2, the index case, is a 27-year-old, female, who had normal birth. At the age of 6 months she was floppy with poor head control, and subsequently never properly reached normal motor milestones. She then went on and developed prominent and violent myoclonic jerks in all four limbs, trunk and neck. She also developed dystonia, which was most prominent on the upper limbs and face without diurnal fluctuation. She would also have intermittently painful spasms, with extension of all four limbs and
trunk that occurred several times daily, lasted up to several hours, and were accompanied by oculogyric crises and double incontinence and exacerbated by fever, infections or tiredness. Magnetic resonance imaging (MRI) brain and dopamine transporters imaging were normal. \textit{TOR1A} genetic testing was performed and resulted negative.

At the age of 13, she first tried L-DOPA, initially at a dose of 100 mg and subsequently increased up to 800 mg over about 5 years. This led to a gradual but significant improvement of dystonia and myoclonic jerks. She also started speaking and became able to control the electric wheelchair with her left hand, but she could still not use her right hand due to myoclonic jerks. The spasms improved in frequency and severity (5-7 times a week). After the introduction of L-DOPA, cognition improved and she was able to attend a special school and later college. Given the marked response to L-DOPA, the suspicion of dopamine synthesis defect was raised. A phenyalalanine loading test was normal. Lumbar puncture was performed under treatment with L-DOPA since the patient refused to stop treatment for investigations and consequently; CSF showed normal pterins; homovallinic acid (HVA) was in the lower normal range 71 nmol/L (normal range: 71-565) and 5-hydroxyindoleacetic acid (5-HIAA) was lower than normal 28 nmol/L (normal range: 58-220) and 3-methyl DOPA was markedly elevated, due to L-DOPA treatment.

She was first seen at the NHNN at age 18. On examination eye movements were normal, she had dystonic grimacing and dysarthria; she had generalized dystonia more prominent on the upper limbs with occasionally choreo-athetoid movements; generalized non-stimulus sensitive, spontaneous and action-induced myoclonus was more prominent in the right side, mostly in the right leg. There was slowing in finger-tapping but not true bradykinesia. She could walk some steps with help. The rest of the examination was unrevealing. Follow-up over the next 9 years revealed no progression of the symptoms. She is still taking 800 mg L-DOPA/daily and occasionally up to 1000/mg, when the spasms are more severe. She has no dyskinesias or further side effects from L-DOPA.

Two of her siblings were similarly affected; case II-3 is a 24 year-old female with an identical picture to patient II-2. She was started on L-DOPA at age 8 years with similar improvement to patient II-2, but significantly better improvement of the jerks. On
examination she had similar findings to the index case, but less myoclonic jerks. Patient II-4 is now 19 and is the youngest sibling. He developed similar symptoms at the same age as II-2 and II-3, with generalized dystonia and prominent myoclonus, but he was treated earlier on L-DOPA at the age of 4. The jerks nearly disappeared. Both II-2 and II-3 are treated with 800 mg L-DOPA/daily (occasionally up to 1000 mg), with no dyskinesias or other side effects.

The older sister, aged 29 and who has two healthy children aged 6 and 8, is unaffected. The father (originally from Nigeria) and the mother (British Caucasian) are asymptomatic and did not show any neurological sign upon detailed examination.

Given the results of the phenyalalanine loading test and CSF analysis, a diagnosis of GCH1-related DRD was considered unlikely. The presence of the DYT1 mutations was excluded. However, the analysis of SGCE revealed the presence of the splice-site variant c.391-3T>C in the index case, which was interpreted as the likely cause of the disease (Valente et al., 2005). Therefore, when I first started working on this project, the consensus was that this was a family with an atypical, particularly severe, form of SGCE-related M-D.

Identification of the Causative Mutation

As a detailed analysis of the segregation of the SGCE splice-site variant had not been performed in all available family members, I first checked whether the variant segregated in the remaining subjects (both parents, the affected brother II-4, and the healthy sister II-1). Sequencing of the variant was performed using primers previously designed and used by our own diagnostics laboratory.

Importantly, I showed that the variant did not segregate, being present in the father, the unaffected sister and absent in the affected subject II-4. Furthermore, the variant is reported at a MAF of 0.2786 in African subjects (as per the ExAC database), thus completely ruling out a pathogenic role and the initial diagnosis of SGCE-related M-D was incorrect.

Subsequently, based on the marked L-DOPA responsiveness, we decided to direct the genetic analysis based on the clinical suspicion of a disorder of the monoamine synthesis pathway. Although the CSF analysis was not indicative of any specific neurotransmitter deficiency, this was performed while the patient was on L-DOPA
treatment, thus making unreliable the dosage of dopamine degradation products.

Analysis of GCH1, SPR and DCC genes did not reveal any pathogenic change. However, TH analysis revealed that all three affected siblings carried two mutations: a previously described point mutation (c.1-71 C>T) in the promoter region (Ribases et al., 2007, Verbeek et al., 2007) and a novel missense mutation in exon 12 (c.1282G>A, p.Gly428Arg). The chromatograms of the mutations are shown in Figure 5-12A.

Figure 5-12 Genetic results
(A) Chromatograms of TH mutations identified in the present study. (B) Amino acid alignment of TH protein showing conservation of the p.Gly428Arg residue in different species.
Segregation analysis confirmed that the patients were compound heterozygotes, being the c.1-71 C>T mutation inherited from the mother and the c.1282G>A inherited from the father. The unaffected sister was heterozygote for the latter mutation. The identified mutation in the gene promoter lies in the Cyclic Adenosine Monophosphate Response Element (CRE), a highly conserved octomer (TGACGTCA→TGATGTCA) located at -74 to -68 upstream the transcription initial codon. The novel variant identified in exon12 p.Gly428Arg is not reported in the ExAC database (containing WES data from more than 10,000 subjects of African ethnical background) and it affects an amino acid highly conserved throughout species. It is consistently predicted by both SIFT, Polyphen-2 and MutationTaster to have a deleterious effect.

5.4.5 Discussion

I report here three members of a family Tyrosine hydroxylase deficiency (THD) who presented with severe myoclonus accompanied by dystonia, consistent with a diagnosis of M-D, as the predominant feature.

It is important that this unusual THD phenotype is recognised and considered in the differential diagnosis of early-onset M-D, as early treatment with L-DOPA is crucial for the final motor outcome in THD (Willemsen et al., 2010). This may be reflected here by the fact that although all three patients were similarly affected before the introduction of L-DOPA, and were subsequently treated with the same L-DOPA dose, the motor outcome is significantly different; the patient that was treated earliest (II-4, 4 years old) showed better response with regard to both the myoclonus and the dystonia than the one that was treated at age 9 years (II-3), and both of them showed a better outcome than the index case (II-2), who was first treated at the age of 13.

Due to presence of a splice-site variant in SGCE, this pedigree was initially labelled as atypical SGCE-related M-D. When the variant was first found in this family, it was not observed in 384 control chromosomes and therefore the variant was interpreted as likely pathogenic and responsible for the phenotype (Valente et al., 2005). Furthermore, L-DOPA-responsiveness has been reported before in cases with M-D (Luciano et al., 2009). Hence, this feature did not put off the treating clinicians from the diagnosis.

However, in spite of its rarity in European population controls (MAF in EXAC 0.001), the variant is common in African controls. This element, together with the complete
analysis of the variant segregation in all family members, completely ruled out the pathogenic role for the variant and discarded the initial diagnosis of SGCE-related M-D.

THD is a rare autosomal recessive, DRD secondary to mutations in \(TH\) (Iwata et al., 1992, Brautigam et al., 1998). Tyrosine hydroxylase is the enzyme responsible for catalysing the conversion of the amino acid L-tyrosine to L-DOPA and therefore \(TH\) loss-of-function mutations result in reduced production L-DOPA and dopamine in nigrostriatal neurons.

The phenotype of THD is thought to be a spectrum with overlap of clinical features between two main phenotypes (Willemsen et al., 2010): on the one hand, a progressive hypokinetic-rigid syndrome with onset of symptoms in the first year, usually presenting with a combination of generalized dystonia and parkinsonism and a good response to L-DOPA; on the other a more ‘complexencephalopathy’, with earlier onset and featuring perinatal abnormalities, diurnal fluctuations, autonomic disturbances and less pronounced response to L-DOPA (Willemsen et al., 2010) (Brautigam et al., 1999, De Lonlay et al., 2000, de Rijk-van Andel et al., 2000, Diepold et al., 2005, Zafeiriou et al., 2009, Yeung et al., 2011).

Despite the large phenotypic variability of THD (Willemsen et al., 2010), prominent myoclonus is not considered classically to belong to THD phenotype. There is one case with ‘jerks’ reported, however, those presented in the concept of severe encephalopathy, rigidity, spasticity, no response to L-DOPA and death at the age of 9 (Hoffmann et al., 2003).

A recent review of all genetically confirmed THD cases, revealed no phenotype-genotype correlations (Willemsen et al., 2010). Our family showed compound heterozygosity of a point mutation in the promoter region (c.1-71 C>T) and a novel non-synonymous substitution in exon 12 (c.1282G>A, p.Gly428Arg). Mutations in the promoter region of the \(TH\) gene have been reported previously; mutagenesis studies in rats and cell lines clearly point out that each nucleotides substitution in the CRE of the \(TH\) gene results in a drastic reduction of the basal transcription (Lazaroff et al., 1995, Nagamoto-Combs et al., 1997). The c.1-71 C>T mutation has been found to cause a reduction of 90% of basal transcription of the \(TH\) promoter (Tinti et al., 1997) allowing a residual activity, which explains the good response to L-DOPA. This variant has been reported previously in two cases, one in the homozygous and one in the compound
heterozygous state (Verbeek et al., 2007). The novel variant detected in exon 12 p.Gly428Arg is highly unlikely to be benign: it is absent in all publicly available databases (over 130,000 chromosomes studied), it affects an amino acid completely conserved throughout species down to invertebrates, and it is predicted by all in silico tools to have a deleterious effect.

TH pathogenic variants reported up to now are often private mutations confined to single or small number of pedigrees and, apart from a common mutation (c.698G>A, p.Arg233His) in the Dutch population (van den Heuvel et al., 1998), and one reported in three Greek families (c.707T>C, p.L236P) (Pons et al., 2010), no other founder effects have been described. Consequently, despite most of THD having been reported in Western Europe, the disease is expected to occur worldwide (Willemsen et al., 2010) and the novel mutation described here is the first pathogenic TH mutation in patients of African ancestry.

In conclusion, THD should be included in the differential diagnosis of early-onset M-D. This is crucial for an early initiation of L-DOPA treatment, which could influence the motor and cognitive outcome of patients.
5.5 Assessment of The Role Of The CACNA1B Arg1389His Variant In a large Myoclonus-Dystonia Cohort

5.5.1 Statement of Contribution
I performed all the genetic analysis described in this chapter.

5.5.2 Background
Recently Groen and colleagues identified the missense variant c.4166G>A; p.Arg1389His (rs184841813) in CACNA1B as the likely causative mutation in a Dutch pedigree with five subjects affected by autosomal dominant M-D lacking mutations in SGCE (Groen et al., 2015). Unique features in the pedigree were lower limb orthostatic high-frequency myoclonus, attacks of limb painful cramps and cardiac arrhythmias in 3 of the affected subjects (Groen et al., 2011). As CACNA1B mutations have not been identified in other unrelated pedigrees, the implication of mutations in this gene as a cause for M-D is not confirmed.

In this study, we assessed the frequency of the CACNA1B c.4166G>A; p.Arg1389His variant in a large multicentric cohort of M-D cases without mutations in SGCE. The results of this analysis do not support a causal association between the variant and M-D.

5.5.3 Subjects, Materials and methods

Patients
A total of 520 M-D cases of British, German and Italian origin were recruited for this study in four tertiary movement disorders centres (London, Lübeck, Tübingen and Milan). This cohort is the same described in the previous chapter, with some additional samples recruited by the Tubingen centre. 146 cases (28%) had a positive family history of M-D. 489 white healthy controls of UK and US origin were provided by the International Parkinson’s Disease Genomic Consortium (IPDGC). All participants provided written informed consent. Additionally, I assessed the frequency of the variant in European cases listed in publicly available datasets of genetic variation (1000 Genomes, Exome Variant Server and Exome Aggregation Consortium).

Genetic analysis
Given the large size of the entire CACNA1B open-reading frame (47 coding exons) and that only one variant has been described to date in M-D, I decided to perform Sanger
sequencing only of the coding exon (exon 28; RefSeq NM_000718.3), which contains the c.4166G>A; p.Arg1389His variant. Primers used for this study are listed in the Appendix.

### 5.5.4 Results

None of the 146 probands with familial M-D carried the CACNA1B c.4166G>A; p.Arg1389His variant. However, the variant was detected only in a single female case of UK origin with sporadic M-D (Figure 5-13). The clinical features of this case are an onset in her mid 30s of tremulous cervical dystonia and myoclonic jerks in the upper limbs. She had no family history for M-D or any other movement disorder. No other family members were available for segregation analysis of the variant.

The total carrier frequency in our M-D cohort, including familial and sporadic cases, is 0.19% (1/520 cases). Surprisingly, the variant was found at a similar frequency in our healthy controls (0.2%; 1/489 individuals). The control carrier of the variant is a 38-year old male without any neurological disease and with no relevant family history of movement disorders.

The variant was found at similar frequencies in the 1000 genome project (0.26%; 1/379 individuals) and Exome Variant Server (0.28%; 12/4,203 individuals) databases. In the Exome Aggregation Consortium database, c.4166G>A; p.Arg1389His is present in 0.11% (38/33,367) of the European subjects (difference to M-D cases not significant; Fisher’s exact test $p = 0.4$).
Figure 5-13 Chromatograms showing the CACNA1B c.4166G>A; p.Arg1389His variant (A) and a control sequence (B).
5.5.5 Discussion

The c.4166G>A; p.Arg1389His variant was identified by Groen and colleagues combining WES and linkage analysis (13 chromosomal regions identified, with a maximum LOD score of 1.2) in a single dominant M-D pedigree. Sanger sequencing of the CACNA1B exons coding for the protein portion spanning from III-S5 to III-S6 failed to reveal other mutations in a further 47 M-D cases.

CACNA1B encodes neuronal voltage-gated calcium channels CaV2.2, which have a key role in controlling synaptic neurotransmitter release (Beuckmann et al., 2003). Furthermore CACNA1A mutations in the homologous region of the gene cause familial hemiplegic migraine (Carrera et al., 1999) and episodic ataxia type 2 (Jen et al., 2001).

The CACNA1B p.Arg1389His substitution represents therefore an excellent candidate as a disease-causing mutation for M-D.

While assessment of co-segregation with disease of the candidate variant represents the first requirement to prove causality in family-based studies, co-segregation in a single pedigree does not establish with certainty its pathogenic role, especially if other co-segregating coding variants and the possibility of a separate undetected pathogenic variant in linkage disequilibrium cannot be convincingly ruled out. This is the case in the pedigree reported by Groen et al., where two other rare missense changes, c.10355A>G; p.Gln3452Arg in VPS13D and c.5308C>T; p.Arg1770Cys in SPTAN1, perfectly co-segregated with the disease. However, given the clinical presentation pointing towards a possible channelopathy, the authors assumed that the causative variant was the one in CACNA1B.

De novo mutations in SPTAN1 have been shown to cause a neurological phenotype (West syndrome with severe cerebral hypomyelination, spastic quadriplegia, and developmental delay) (Saitsu et al., 2010) and more recently a microdeletion encompassing SPTAN1 was detected in a child with epileptic encephalopathy and severe dystonia (Matsumoto et al., 2014). Furthermore, its frequency in the ExAC database is significantly lower (2/33,360 European subjects) than the CACNA1B p.Arg1389His. Therefore, the SPTAN1 could also represent a good candidate gene responsible for M-D.

A candidate variant responsible for a rare disease should be found at a low frequency in
population controls, consistent with the proposed model of inheritance and disease prevalence. As previously mentioned, M-D is a very rare disorder with a suggested prevalence of around 2 per million in Europe (Asmus and Gasser, 2010). I would therefore anticipate highly penetrant mutations causing dominant forms of M-D to be absent or extremely rare in the general population. Yet, this is not the case for p.Arg1389His, which is present at a considerable frequency in our healthy controls and all publicly available databases (~0.1-0.3%). According to the Exome Aggregation Consortium database, the carrier frequency of this variant in Europeans is ~4 times higher than the TOR1A c.904_906delGAG deletion (0.026%), which is by far the most common single mutation responsible for dystonia described to date (Valente et al., 1998). Given this frequency, if c.4166G>A; p.Arg1389His were a pathogenic variant, it would be expected to be responsible for a large proportion of familial M-D cases. However, in our cohort not only was the variant not identified in any of the probands with familial M-D, but the overall frequency of the variant did not differ between M-D cases and healthy controls. This does not support a pathogenic effect of the variant even assuming a reduced penetrance.

These results strongly question the pathogenic role of the CACNA1B variant p.Arg1389His as a cause for M-D. Despite its compelling biological role, fitting very well with the current themes in dystonia pathogenesis (i.e. abnormal calcium signalling), further genetic evidence is needed before designating CACNA1B mutations as a cause for dominant M-D. In the era of NGS, the rate of identification of rare genetic variants has hugely increased. Hence, to avoid false assignment of pathogenicity, a close scrutiny is necessary before causally linking a candidate variant to a disease.

Groen and colleagues performed functional work (whole-cell and single-channel patch recording studies) showing that the identified CACNA1B variant affected the normal activity of the channel (i.e. the mutant channel carried less current when open). However, the frequency of this variant in controls is simply incompatible with a dominant variant causing a disease with an aggregate frequency of 1 per 500,000, even if incomplete penetrance is posited. As such there is no remaining support for CACNA1B as a causal gene for M-D.
Chapter 6. Exploring the Genetics of Benign Hereditary Chorea

6.1 Outline of chapter

This chapter describes the work I performed to characterise the genetic aetiology, and where possible to identify novel genetic causes, in a cohort of patients with benign hereditary chorea (BHC). As discussed below, *NKK2-1* represents the only gene that has been incontrovertibly associated with familial BHC.

The chapter begins with an overview on BHC. The specific aims of the study were:

1) To define the frequency of *NKX2-1* mutations in the BHC Queen Square cohort

2) To evaluate whether mutations in *ADCY5*, the gene responsible for the condition Familial Dyskinesias with Facial Myokymia, are a cause of BHC

3) Identify novel genetic causes of BHC

6.2 Benign Hereditary Chorea

The term BHC was first introduced in the medical literature in 1967. The authors described a “non-progressive syndrome of inherited childhood-onset chorea with a good outcome in the absence of an underlying progressive disease” affecting 14 individuals, members of a large autosomal dominant 5-generation African American pedigree (Haerer et al., 1967). A similar presentation, with different patterns of inheritance, was subsequently observed in a number of other families, (Nutting et al., 1969, Chun et al., 1973, Bird et al., 1976, Burns et al., 1976).

A single epidemiological study in the Welsh population has tried to establish the frequency of BHC (Harper, 1978). BHC appears to be a very rare syndrome with an estimated prevalence of 1 in 500,000, although this figure is likely to be an underestimate due to poor recognition of the syndrome.

Based on the analysis of 42 individuals with BHC reported in the medical literature since 1967, Kleiner-Fisman and Lang outlined the phenotype of BHC. BHC is mainly
characterized by the onset during infancy or early childhood of a choreic syndrome, often preceded by hypotonia and delay in reaching motor milestones. Affected children are often described as clumsy and report several falls. BHC cases, differently for neurodegenerative forms of choreas, present relatively scarce progression of symptoms and the absence of other major neurological deficits, in particular severe mental retardation or cognitive decline. In some families improvement of chorea can be observed over time (Kleiner-Fisman and Lang, 2007).

After mapping the disease locus to chromosome 14q (de Vries et al., 2000, Fernandez et al., 2001), mutations in the NKX2-1 gene, encoding the thyroid transcription factor 1 (TTF-1), were recognized in 2002 as a major cause for autosomal dominant BHC (Breedveld et al., 2002). At the same time, a second independent group also identified NKX2-1 point mutations in cases with congenital hypothyroidism and who also suffered from choreoathetosis, muscular hypotonia, and pulmonary problems (Krude et al., 2002). This discovery represented a major breakthrough in the recognition and diagnosis of this condition. To date ~190 cases and ~100 NKX2-1 mutations (both point mutations and large deletions; Figure 6-1) have been reported, allowing a better definition and an expansion of the phenotype associated with mutations in this gene (Inzelberg et al., 2011, Gras et al., 2012, Thorwarth et al., 2014). It is now well recognized that NKX2-1 mutations often lead to a complex multisystem disease, defined brain-lung-thyroid syndrome, which features hypothyroidism, pulmonary defects and a variety of neurological symptoms (including not only chorea, but also hypotonia, neurodevelopmental delay, learning disabilities, behavioural problems, dystonia, myoclonus, tics and ataxia; Figure 6-2) (Gras et al., 2012). In particular, a series of clinical observations have consistently pointed out that dystonic features and EMG-confirmed myoclonic jerks are often present in carriers of NKX2.1 mutations (Asmus et al., 2007, Armstrong et al., 2011, Gras et al., 2012). Importantly, as chorea tends to improve during adulthood, myoclonus-dystonia (M-D) symptoms may become the most prominent features in adult individuals with NKX2.1 mutations, leading to significant diagnostic difficulties in differentiating BHC from M-D secondary to SGCE mutations.

Around 97% of cases reported in the literature present neurological symptoms, with only a minority of cases presenting with isolated lung involvement (Devriendt et al., 1998, Iwatani et al., 2000, Hamvas et al., 2013). However, it is important to acknowledge that this figure may reflect an ascertainment bias due to selection of cases
to be screened for *NKX2.1* mutations mainly based on the presence of neurological involvement. The two larger case series reported to date have demonstrated the presence of thyroid dysfunction in ~70% of mutation positive cases and slightly lower frequency of pulmonary involvement (Gras *et al.*, 2012, Thorwarth *et al.*, 2014). The full triad of brain, thyroid and lung involvement is observed in ~30% of patients, whereas isolated chorea in observed in only ~20% of mutation carriers. Pathogenic *NKX2.1* are often loss-of-function and lead to haploinsufficiency. Missense mutations affect the ability of the protein to bind to DNA (Provenzano *et al.*, 2016).

---

![Figure 6-1 Schematic representation of *NKX2-1* point mutations reported to date. Reproduced from (Peall and Kurian, 2015)](image-url)
NKX2.1 is a transcription factor gene essential for a correct early development of lungs, thyroid, and forebrain regions, in particular the basal ganglia. NKX2.1 is expressed in brain, especially in striatal cholinergic neurons. Pathological analysis in one case with genetically confirmed NKX2-1 associated BHC showed normal gross macroscopic and microscopic appearances, though with reduced count of striatal interneurons (Kleiner-Fisman et al., 2003, Kleiner-Fisman et al., 2005). In mice, homozygous Nkx2-1 gene deletion results in a lethal phenotype, featuring pulmonary hypoplasia, absence of functional thyroid tissue and complex malformations of brain structures, including the basal ganglia (Kimura et al., 1996).

Figure 6-2 Clinical features described in cases with NKX2-1 mutations
Reproduced from (Peall and Kurian, 2015)
Importantly, a significant number of BHC families do not carry mutations affecting the coding sequence of \( NKX2-I \). Breedveld et al. failed to demonstrate linkage to chromosome 14 in four dominant BHC families (Breedveld et al., 2002). Bauer and colleagues did not identify \( NKX2-I \) mutations in 18 patients with chorea of unknown origin, including the index cases of three families with dominant benign chorea (Bauer et al., 2006). Genetic linkage analysis mapped an alternative disease associated locus on chromosome 8 (8q21.3-q23.3; locus named benign hereditary chorea 2 - BHC2) in two unrelated Japanese families with autosomal dominant non-progressive chorea (Shimohata et al., 2007). However, the clinical phenotype in these pedigrees differed from that of classic BHC, as the disease onset was in adult years and not in childhood.

Finally, Thorwarth and colleagues recently completed an extensive clinical and genetic study in a large cohort of 101 cases with suspected BHC cases, clinically ascertained based on the presence of a variable combination of chorea, thyroid and lung disturbances. They identified \( NKX2-I \) point mutations in 17 cases and large deletions in 10 cases (Thorwarth et al., 2014). Of 74 cases without \( NKX2-I \) mutations, 15 presented the full triad of brain-thyroid-lung involvement, 34 had brain-thyroid involvement and 8 brain-lung involvement (Figure 6-3), suggesting the existence of variants in \( NKX2-I \) enhancer and/or promoter region or mutations in other yet unknown genes.

![Figure 6-3 Brain, thyroid and lung involvement in index patients and family members, with and without \( NKX2-I \) mutations, in the largest cohort of Benign Hereditary Chorea cases screened to date](image)

Reproduced from (Thorwarth et al., 2014).
Intriguingly, 2 of the deletions detected in BHC cases were neighboring but not including the \( \textit{NKX2-1} \) coding region. One deletion was de novo, while the other segregated in a large pedigree with chorea of dominant inheritance. This finding confirms previous isolated observations of chromosomal deletions neighboring but not including \( \textit{NKX2-1} \) in 2 other BHC cases (Barnett et al., 2012, Dale et al., 2012). The pathogenic mechanism of such deletions is currently not clear. The deletions may remove essential regulatory elements downstream of \( \textit{NKX2-1} \) and critically affect \( \textit{NKX2-1} \) gene expression. The only coding gene encompassed by all deletions not \( \textit{NKX2-1} \) reported to date is \( \textit{MBIP} \), which encodes MAP3K12 binding inhibitory protein, a histone acetyltransferase complex involved into chromatin modifications (Suganuma et al., 2010). In situ hybridisation of \( \textit{MBIP} \) in early mice embryos revealed the same expression pattern of \( \textit{NKX2-1} \), including the basal ganglia, forebrain, thyroid and lung, making \( \textit{MBIP} \) a plausible cause of a \( \textit{NKX2-1} \) deficiency like phenotype (Thorwarth et al., 2014).
6.3 NKX2-1 Mutational Analysis in the Queen Square Benign Hereditary Chorea Cohort

6.3.1 Statement of Contribution
All participants to this study were clinically characterized and diagnosed by Professor Niall Quinn and Professor Kailash P. Bhatia. Dr Liana Veneziano and I performed the NKX2.1 mutational analysis.

6.3.2 Background
The Queen square cohort of BHC, consisting of ten families (7 dominant and 3 recessive) and four sporadic cases, was first characterised and published in 2000 (Schrag et al., 2000), before the identification of NKX2-1 mutations as the underlying genetic cause of BHC. Surprisingly, a systematic mutational analysis of NKX2-1 was never performed. In this section, I describe the NKX2-1 analysis performed in the BHC Queen Square cohort.

6.3.3 Subjects, Materials and Methods

Selection of cases for the project
Of the 14 cases reported by Schrag and colleagues, DNA was available for only seven subjects (five familial and two sporadic cases).

A search of the NHNN Neurogenetics database revealed an additional 16 cases, recruited after the publication of the initial series, with a clinical description consistent with BHC. These cases were clinically diagnosed as BHC based on the presence of a movement disorder, with onset prior to the age of 20, predominantly characterized by chorea in the absence of other major neurological features. In particular, none of the selected cases presented significant cognitive decline. Amongst these six had a positive family history for a similar disorder consistent with an autosomal dominant transmission, whereas ten cases were sporadic. In total, 23 cases (11 familial cases and 12 sporadic) were included in the study. Amongst these, four familial cases and one sporadic case had evidence of extra-neural involvement, namely thyroid and lung defects.

All these cases had been extensively investigated for acquired causes of chorea, without reaching a conclusive diagnosis. All cases had previously tested negative for the HD
triplet repeat expansion. Furthermore, given the reported clinical overlap with Myoclonus-Dystonia, all cases had been screened for SGCE mutations and were negative. All cases included in this study were recruited at the Movement disorder center of the National Hospital for Neurology and Neurosurgery, Queen Square, London and had given written consent and formal ethical approval by the relevant research ethics committee.

**Genetic Analysis**

I performed direct Sanger sequencing of the NKX2-I coding exons (3 coding exons; transcript ENST00000354822) in all 23 BHC cases, as described in the Materials and Methods section. MLPA analysis to detect NKX2-I copy-number variants was performed by Dr Liana Veneziano at the Institute of Translational Pharmacology in Rome.

### 6.3.4 Results

**NKX2-I mutational analysis**

This analysis revealed a total of five different heterozygous NKX2-I mutations in five out of 23 BHC cases (21.7%). All mutations were found in cases with familial autosomal dominant BHC (5/11; 45.4%). The identified mutations were a missense mutation (c.605A>G; p.Gln202Arg), two STOP-gain mutations (c.622C>T; p.Arg208* and c.817A>T; Lys241*), one splice-site mutation (c.463+1G>C; p?), and one intronic variant (c.464-24A>C). Two point mutations (c.463+1G>C and p.Arg208*) have been previously described in BHC cases and are therefore established pathogenic variants, whereas the other variants are reported here for the first time. All variants are absent in all publicly available datasets (including ExAC). Segregation analysis supported a pathogenic role for all the identified variants as they were all found in other affected family members (see below). The missense change Gln202Arg is predicted pathogenic by Polyphen-2, SIFT and MutationTaster, and affects an amino acid completely conserved down to invertebrates. The effect of the two intronic mutations on splicing was evaluated *in silico* using Human Splicing finder (http://www.umd.be/HSF3/). The software predicted a highly detrimental effect for the c.463+1G>C (i.e. broken wild-type donor site, most probably affecting splicing). Conversely, the *in silico* prediction tool did not highlight any significant splicing motif alteration determined by the c.464-24A>C variant, suggesting that this mutation may not have any impact on splicing.
However, the variant segregated in other two affected family members and is completely absent in all publicly available datasets. As we did not have access to RNA to test experimentally the effect of the mutation on splicing, the pathogenic role of this variant remains uncertain.

Transmitted and *de novo* large deletions neighbouring, but not encompassing, the *NKX2-1* coding region have been recently described in cases with familial and sporadic BHC (Thorwarth *et al.*, 2014). Thus, all cases without *NKX2-1* point mutations or large deletion were genotyped using a dense array HumanOmni 2.5 (Illumina). However, no deletions were detected (Figure 6-4).
Figure 6-4. Array data of patient with BHC showing no deletions either encompassing or neighbouring the *NKX2-1* locus
Clinical presentations of cases with NKX2-1 mutations

Family-1

This pedigree has been reported before (Robinson and Thornett, 1985) and represents family 7 in the paper of Schrag and colleagues.

The proband, aged 46 when last examined, is the product of a normal pregnancy and his birth was uncomplicated. His motor milestones were delayed and as a child he suffered with several falls per day. He developed involuntary movements and upper limb jerks early in childhood, which progressed until age 11. After that, the movement disorder began to improve. When last assessed, the involuntary movements would still represent a source of social embarrassment, but they were not functionally invalidating. He had no cognitive difficulties. On examination he had generalised dystonia affecting the face, hands and feet, with super-imposed myoclonic jerks, mainly distributed in the upper body. He had a dystonic gait. Dystonia dramatically improved with a small dose of trihexyphenidyl (4 mg/day) combined with a small dose of L-DOPA (600 mg/day), with an almost complete resolution of symptoms.

His son, who carried the same mutation, reported a similar medical history, with delayed motor milestones, frequent falls in childhood and an onset of the movement disorder very early in life. On examination, at age 22, he displayed generalised chorea, frequent stutter and gasps. He was not on any treatment when last assessed. Nobody else in the family was affected. The available medical notes did not mention any lung or thyroid involvement. Both cases carried the pathogenic splice site mutation c.463+1G>C.

Family-2

This family represents family 8 in the paper of Schrag et al. This 38-year-old woman was the product of a normal pregnancy and delivery. She had delayed reaching of motor milestones, with walking independently only at 3 and ½ years. As a child she would have multiple falls per day and she was clumsy and poorly coordinated. She was unable to run or ride a bicycle. Involuntary movements appeared at a very early age and these would affect mainly lower limbs, making her balance worse. Over the years the condition worsened, forcing her to use a wheelchair outdoors to prevent falls. Due to the presence of significant learning difficulties she had to attend special schools.

Her examination showed mild jerky choreic movements affecting the face, the trunk and
the four limbs. Her speech was slurred, but the rest of the cranial nerve examination was unremarkable, including eye movements. There was mild posturing of hands and feet. Gait was choreic with superimposed elements of ataxia. A trial of L-DOPA up to 600mg/day was not beneficial and therefore trihexyphenidyl was prescribed, though no information about the response to this treatment is available in the patient’s notes.

She carried the pathogenic mutation p.Arg208*, which was also carried by her similarly affected mother and brother.

Family-3

This family represents family 10 in the paper of Schrag and colleagues

Unfortunately, I was not able to retrieve any relevant clinical information for this pedigree. The pedigree consists of two siblings and their mother, all affected. They all had onset of symptoms in the first two years of life and all had delayed motor milestones and experienced many falls per day during childhood. Intellectual functioning was reported to be low average. The disease symptoms slightly improved with age in all family members. They all carried the intronic variant c.464-24A>C, of uncertain clinical significance.

Family-4

The proband is a 49-year old lady born via an uncomplicated delivery after a normal pregnancy. She was a ‘floppy’ baby. Her motor milestones were delayed, with delayed independent walking. At age 2 she was diagnosed with cerebellar ataxia. She experienced lifelong balance problems and, at around age 16, the balance worsened due to the appearance of choreic movements. Benztropine was then started resulting in improved symptoms. Hypothyroidism was diagnosed around age 40 and treated successfully with levothyroxine. She took no medications for the movement disorder.

On examination, her gait was mildly ataxic and she had clear chorea and dystonic movements of her head and shoulders and dystonic posturing of the hands while walking. Eye movement examination revealed ocular apraxia with difficulty initiating saccades, broken pursuit movements, hypometric saccades and gaze impersistence. She had mild dysarthria. Her daughter was also affected and both cases carried the p.Lys248* mutation, which was not found in her parents, brothers and sisters (all
unaffected). After identifying the *NKX2-1* variant in this case, I was informed that a genetic diagnosis had actually already been reached and the clinical description of this case had already been published (Veneziano *et al.*, 2014).

**Family-5**

This case, aged 35 when last assessed, was born 3 weeks premature and initially fed with difficulty. She had a chest infection as a new-born and her initial motor milestones were delayed. Twitching choreic movements in her limbs were noticed since birth and, as soon as she started walking, she began to experience several falls per day. Her symptoms gradually improved with age. On examination she had a mix of chorea and dystonic posturing affecting the four limbs. She was not taking any medication for the movement disorder. Her daughter was similarly affected and both carried the novel missense change *p.Gln202Arg*.

### 6.3.5 Discussion

While several studies are available in the literature describing ~200 BHC cases caused by *NKX2-1* mutations (recently reviewed in [Peall and Kurian, 2015]), surprisingly few studies have systematically assessed the contribution of *NKX2-1* mutations in cohorts of clinically defined BHC cases.

In this chapter I report the results of a detailed screening of *NKX2-1*, by means of Sanger sequencing and MLPA analysis, of the BHC Queen Square cohort. All these cases fully met the current clinical definition of BHC, as they all presented in childhood with a non-acute onset, predominantly choreic, movement disorder, which was scarcely progressive and did not include prominent additional neurological features, in particular cognitive decline. As dystonia, in combination with chorea, has been observed in genetically proven *NKX2-1*–mutation cases (Fung *et al.*, 2006, Armstrong *et al.*, 2011), this feature was not considered as an exclusion criterion for the study.

The study included both familial and sporadic cases as it has been recently shown that a high proportion of mutated cases carry de novo mutations and therefore the absence of a positive family history should not excluded a diagnosis of *NKX2-1*–related disease (Gras *et al.*, 2012). All sporadic cases had all been extensively investigated for acquired causes of chorea.

Importantly, I show that only five out of 23 BHC cases carried variants in the gene. All
cases had a positive family history for the condition and mutations were shown to segregate in additional affected family members, strongly supporting their pathogenic role. Four of the five identified mutations are either established pathogenic changes or are clear loss-of-function variants, and therefore pathogenic by definition. More controversial is the role of the intronic variant found in family 3, as in silico tools did not predict an obvious effect on splicing and I did not have access to patients’ RNA to establish experimentally the effect on splicing. However, in support of its pathogenic effect, the mutation is completely absent in all publicly available datasets and segregated in the three available family members. Furthermore, the role of deep intronic mutations in Mendelian diseases is increasingly recognised (Akman et al., 2015, Stepensky et al., 2016).

Although the yield of mutation positive cases (21.7%) of this study could look surprisingly low, it is actually in line with what has been reported by other groups. In particular, Thorwarth and colleagues have recently completed an extensive clinical and genetic study in a large cohort of 101 cases with clinically suspected BHC cases and found that 74 cases (73% of the cohort), of which 15 presented with the full triad of brain-thyroid-lung involvement, actually did not carry mutations in NKX2-1 (Thorwarth et al., 2014). Similarly, in our cohort, three out five cases with brain-thyroid-lung disease, two of which with a clear autosomal dominant family history, did not carry NKX2-1 point mutations or deletions, or the recently described deletions of the chromosomal region neighbouring NKX2-1. These results clearly indicate the existence of either undetected variants in NKX2-1 (e.g. in the enhancer and/or promoter region or deep intronic mutations) or mutations in other yet unknown genes that critically interact with NKX2-1 for the correct development of striatal neurons and of thyroidal and pulmonary tissues.

Interestingly, part of this cohort (and three of NKX2-1-mutated cases; family 1, 2, and 3) had been reported back in 2000, before the discovery of NKX2-1 mutations (Schrag et al., 2000). While all three families had at some point been labelled as affected with typical BHC, several cases from all three families, when clinically reappraised for the paper of Schrag and colleagues, were re-assigned to other diagnostic categories, such as M-D or generalised dystonia. This element clearly suggests the difficulty of recognising genetic entities on clinical grounds alone. Furthermore, it highlights the broad clinical heterogeneity associated with mutations in NKX2-1.
In conclusion, while the term BHC is often used to imply the presence of *NKX2-1* mutations, a significant number of sporadic cases and families with BHC do not carry mutations in this gene. Furthermore, *NKX2.1*-mutated cases can frequently present with movement disorders other than chorea.
6.4 ADCY5 Mutational Analysis in a Cohort of Benign Hereditary Chorea Cases

6.4.1 Statement of Contribution
I performed all the genetic analysis described in this section.

6.4.2 Introduction
In 2001 Fernandez and colleagues described a North American 5-generation pedigree of German origin with 18 members affected with a childhood-onset autosomal dominant movement disorder. The clinical presentation in this family was that of a hyperkinetic movement disorder featuring predominantly chorea and dystonia involving the limbs, neck, and face. Perioral and periorbital twitching was also observed in some family members and needle electromyogram (EMG) in two individuals revealed what was initially thought to be myokymia (Fernandez et al., 2001). The disease was consequently named Familial Dyskinesia with Facial Myokymia (FDFM).

In 2009 the disease locus was subsequently mapped to chromosome 3p21-3q21 (Raskind et al., 2009), but only recently Chen and colleagues identified using WES a missense variant in ADCY5, encoding adenyl cyclase 5 (AC5), as the causative mutation (Chen et al., 2012). Prior to the identification of facial myokymia as one of the core features, individuals from the original FDFM kindred had been described as affected with familial essential (“benign”) chorea (Bird et al., 1976). Hence, given the large phenotypic overlap between FDFM and BHC, I hypothesised that ADCY5 mutations could be responsible for BHC in cases lacking mutations in NKX2-1.

I therefore performed a mutational analysis of ADCY5 in the Queen Square cohort of NKX2-1-negative BHC cases and in a second cohort of NKX2-1-negative BHC cases of Italian origin.

6.4.3 Subjects, Materials and Methods

Samples
The eighteen BHC cases that did not carry NKX2-1 mutations (see previous section), including six with an autosomal dominant family history and 12 sporadic, were initially selected for this study. An additional cohort of 35 cases of Italian origin, clinically diagnosed by Prof Nardo Nardocci’s and Dr Carlo Fusco’s groups, was subsequently
studied. Dr Liana Veneziano and Dr Barbara Garavaglia’s laboratories had previously performed the NKX2-1 analysis in the Italian cohort, including Sanger sequencing and MLPA analysis.

**Genetic Analysis**

WES was performed in ten of the BHC cases. Cases were prioritised for WES if they had an autosomal dominant family history (five cases) and/or an age at onset < 10 years (five cases), assuming that a genetic aetiology would be more likely for these cases.

In the Queen Square BHC cohort, the search for ADCY5 mutations was performed either by interrogating WES data or by Sanger sequencing of the ADCY5 coding region (21 exons; transcript ENST00000462833). As WES poorly covered exon 1 due to its high GC content, Sanger was used to sequence exon 1 in all cases. Visual inspection revealed that all other exons were adequately covered by WES.

In the Italian cohort I screened by Sanger sequencing only ADCY5 exons 2 and 10. This strategy was undertaken as these two exons are likely to represent mutational hotspots.

WES and Sanger sequencing were performed as described in the Materials and Methods chapter. The primers used to amplify the coding exons and splice sites of ADCY5 are listed in the appendix.

Segregation analysis was performed in all available relatives of mutation positive-cases.

**Expression Profiling of ADCY5 and NKX2-1 in Brain Tissue**

Brain ADCY5 and NKX2-1 mRNA regional distribution and expression in the normal adult human brain and longitudinal expression changes during the course of human brain development were determined using microarray analysis of human post-mortem brain tissue from the UK Human Brain Expression Consortium and the Human Brain Transcriptome database, as previously detailed in the Materials and Methods chapter.

**6.4.4 Results**

**Genetic Analysis**

The search for mutations in the entire ADCY5 coding region in the Queen square BHC cohort revealed the presence of five different heterozygous mutations.

The c.1252C>T; p.Arg418Trp variant was identified in two unrelated cases. This amino
acid change has been previously reported in association with FDFM (Chen et al., 2014) and is absent in all publicly available datasets of population controls (dbSNP, 1000genome project, NHLBI exome variant server and Exome Aggregation Consortium). All \textit{in silico} prediction tools (Polyphen-2, SIFT and Mutation Taster) consistently predict a deleterious effect of the mutation.

Another missense variant (c.2117C>T; p.A706V) was detected in the index case of a pedigree with autosomal dominant BHC (brain-lung-thyroid syndrome). This variant is novel, as it is not reported in any of the databases of genetic variations.

The variant was predicted benign and tolerated respectively by Polyphen-2 and SIFT, but disease causing by MutationTaster. Segregation analysis in the family showed that the mutation was absent in the affected mother and was inherited by the unaffected father, thus ruling out any pathogenic relevance. A non-coding variant in the 5’ UTR (c.1-5G>C; rs566725675) together with the missense variant c.29C>T; p.P10L (rs143905423) (phase of the mutation not know) were detected in a sporadic case of African ancestry. The phase of the two mutations is not known. Both variants are reported with a MAF ≥ 1% in control individuals of African ancestry as reported by the ExAC, thus excluding a pathogenic role.

In total 2/18 probands (11.1%) from the Queen Square cohort carried pathogenic \textit{ADCY5} variants, both located in exon 2.

The analysis of exons 2 and 10 in the Italian BHC cohort revealed that three probands (3/35; 8.5%) carried pathogenic \textit{ADCY5} missense mutations, all involving the same amino acid residue, Arg418. The variant c.1252C>T; p.Arg418Trp, the same change identified in the two Queen Square cases, was also found in one Italian sporadic case. Two novel variants, c.1253G>A; p.Arg418Gln and c.1252C>G; p.Arg418Gly were found in a sporadic and in a familial case, respectively. Both variants are absent in all publicly available datasets and are predicted pathogenic by Polyphen-2, SIFT and MutationTaster.

\textit{Clinical Presentations of Cases with Pathogenic ADCY5 Mutations}

Clinical information of the patients with pathogenic \textit{ADCY5} mutations identified in this study is summarized in Table 6-1.
Family-1

The clinical description of this pedigree has been previously published in 1981, as part of a report describing a series of families with benign (non-paroxysmal) familial chorea (Sleigh and Lindenbaum, 1981). The index case is a British 36-year old man. His birth and early development throughout infancy were normal. In the first year of life he progressively developed brief choreic movements at rest, involving the face and the four limbs. Movements were markedly worsened by excitement, stress or tiredness. Around the age of 18 he also developed painful spasms of the four limbs, particularly frequent and severe upon awakening. Symptoms progressed over the years and neurological examination at the age of 28 revealed dysarthric speech and severe abnormal involuntary movements comprising generalized chorea with facial grimacing and marked dystonic features. Both chorea and dystonia were present at rest, but action significantly exacerbated movements in the limbs. Eye movements were abnormal with gaze impersistence and use of head thrust to initiate saccades. Gait was unsteady with both choreic and dystonic features, but cerebellar testing was otherwise normal. He never displayed facial myokymia and EMG of periorbital and perioral regions failed to show either myokymia or other signs of motor neuron hyperexcitability. Over the years different medications were tried without major improvements, including trihexyphenidyl, tetrabenazine, baclofen, L-DOPA and clonazepam. He is currently on a combination of trihexyphenidyl (3 mg/day) and tetrabenazine (75 mg/day), which he finds helpful in reducing the intensity and frequency of the spasms. His 64-year old father also developed hyperkinetic movements in the first year of life. His neurological examination mainly consisted of generalized chorea, but no dystonic posturing. Overall, his clinical presentation was much milder than his son’s one. Besides chorea, on examination he had marked ocular and motor impersistence and tandem walking difficulties. Cerebellar testing was otherwise normal. At the last follow-up, when he was 61, occasional twitches were observed in the periorbital areas, but an EMG performed on the same day did not show myokymia. A low dose of tetrabenazine was tried but was not tolerated due to the developed of depression. He is currently taking trihexyphenidyl (7.5 mg/day), with satisfactory control of the dyskinesias. Higher doses were not tolerated due to the occurrence of cognitive difficulties. There was no history of movement disorders in other family members, including both paternal grandparents of the index case.
The heterozygous ADCY5 c.1252C>T; p.Arg418Trp mutation was found by WES and confirmed by Sanger sequencing in subject III-1. However, Sanger sequencing surprisingly failed to detect the mutation in the affected father (Figure 6-5). Mosaicism has been previously reported in patients with mild presentation of Mendelian disorders (Beicht et al., 2013). We therefore suspected that Sanger sequencing might have missed the mutation in the father due to low-grade somatic mosaicism.

As next-generation sequencing has been shown to be a more sensitive method to detect this type of mutations (Beicht et al., 2013; Miyatake et al., 2014), WES was subsequently performed also in subject II-1. Interestingly, this disclosed the presence of the mutated allele in 9/110 of the sequence reads (~8% versus ~40% in his son; Figure 6-5), confirming the presence of low-level somatic mosaicism. This suggests that the mutation arose in the father during early stages of embryogenesis. Sanger sequencing analysis showed that the unaffected mother and sister of the index case did not carry the mutation.

Family-2

This is a 22-year-old man of Pakistani ethnicity. He is the son of a consanguineous marriage (parents are first degree cousins) and he is the first of four siblings. He is the only affected member in the family. He had delayed motor milestones with walking at approximately 20 months. Since then his gait has always been abnormal, being mainly characterized by arching of the trunk, walking on tiptoes and inversion of both feet. Around the age of 2 he developed constant involuntary movements involving all four limbs and trunk, which became progressively more evident over the years. Movements were largely exacerbated by stress and anxiety. Around the age of 10, he developed “fits” of violent hyperkinetic movements and painful spasms of all four limbs, mainly occurring at night. A polysomnographic study performed to investigate their nature excluded any epileptic activity and showed that the attacks occurred when the patient was awake. Neurological examination showed he had moderate generalized chorea at rest, and intermittent distal dystonic posturing in the four limbs. The hyperkinetic movements were dramatically increased by action. Gait was unsteady, but cerebellar testing was otherwise normal. Cranial nerve examination showed ocular impersistence and frequent facial choreic movements were also observed. The patient declined an EMG of the facial muscles.
He is currently taking trihexyphenidyl (30 mg/day) and tetrabenzamine (50 mg/day), which partially reduced the intensity of the dyskinesias and the frequency of the spasms.

WES and subsequent Sanger sequencing analysis showed the presence of the ADCY5 c.1252C>T; p.Arg418Trp change in the heterozygous state in the affected. The mutation was de novo, as Sanger sequencing showed its absence in both healthy parents and in the three unaffected siblings.

Family-3

The proband is a 14-year-old Italian girl born pre-term (38th week) from healthy parents. She presented with psychomotor delay and could walk independently at around age 3. At age 4, she developed a generalized hyperkinetic movement disorder featuring a combination of chorea and dystonia, which made ambulation difficult and caused frequent falls. Pyramidal signs were present in the lower limbs together with a bilateral foot dystonia, requiring splints and regular botulinum toxin injections. Since the first months of life, this patient presented severe paroxysmal worsening of chorea upon falling asleep, which could happen in clusters of 2-10 attacks lasting up to 30 seconds each with spontaneous remission. In addition, she later presented two different kinds of paroxysms, one occurring mainly during sleep (generalized dystonia and dyskinesias with facial grimacing and no sleep interruption, with onset around age 4), and one during the day (multiple sudden falls due to give-way of legs without premonitory sensation, with onset around age 11). All the types of attacks were favored by tiredness and emotions, but no clear triggering factors were reported.

Examination at age 14 showed generalized chorea involving also the perioral muscles, upper and lower limbs dystonic posturing and fixed dystonia (in turning and plantar flexion) of the left foot. Scissoring gait with pyramidal signs in the lower limbs were present and speech was markedly dysarthric.
Figure 6-5 Pedigree and genetic results of family 1

On the left the visual output of the WES data and Sanger sequencing results in the affected individuals carrying the p.Arg418Trp mutation. In the top section of each box the read depth of the exonic portion of ADCY5 involved by the mutation is shown. In the bottom part of each box, samples of the reads carrying the mismatching allele are displayed. The mutant T replacing a C is highlighted in red. Sanger sequencing failed to show the mutation in individual II-1.
Overall, the severity of movement disorder was reduced as compared to infancy, and so was the frequency of nocturnal paroxysms; however, diurnal paroxysms were very disabling and could occur several times per day. She was on Trihexyphenidyl 32 mg/day. Sanger sequencing showed she carried the p.Arg418Trp change, which was demonstrated to be de novo.

Family-4

The proband is an 18-year-old boy of Italian origin born from healthy parents. He presented with delayed motor milestones and a tendency to tiptoe walking around the age of 18 months. Since his first months of life, nocturnal paroxysms characterized by generalized dystonia with inconsolable crying, lasting for hours, disrupted his sleep. During infancy he developed generalized chorea and myoclonic jerks also involving facial muscles. Episodic worsening of dyskinesias without specific triggers accompanied by hyperventilation was noticed during childhood with a weekly frequency. Examination at age 17 showed generalized asynchronous, non-stimulus sensitive myoclonic jerks at rest, mild dysarthria, scissoring gait with pyramidal signs in the lower limbs and dystonia of upper limbs and the trunk. Diurnal and nocturnal paroxysms were still present, albeit at a decreased frequency (one episode per year). Overall no progression in the severity of his movement disorder was noticed as compared to infancy, though choreic movements decreased and myoclonic and dystonic aspects became more prominent. He carried the novel p.Arg418Gln substitution, which was not found in both his parents and therefore confirmed to be a de novo mutation.

Family-5

The proband is a 2-year-old boy of Italian origin. He was born normally after a normal pregnancy. Generalised choreic and myoclonic movements appeared in the first months of life. Motor and language development were severely delayed. He is currently able to sit, but not to stand without support. The neurological examination showed axial hypotonia, diffuse chorea and myoclonic jerks and dystonic posturing in the four limbs. A tendency to tiptoe walking was noted. A trial with trihexyphenidyl was unsuccessful.

His father, currently aged 47, is affected with generalised dystonia mixed with choreic elements, started around age 3. Differently from his child, his motor and language development were entirely normal. He presented frequent facial twitches and severe
dysarthria. The dyskinesias are markedly worsened by action, emotions and stress. Paroxysmal episodes of severe and painful dyskinesias are triggered by strong emotions and sleep. Several pharmacological trials (including trihexyphenidyl, L-DOPA and tetrabenazine) were ineffective. Both father and son have recently started a therapeutical trial with acetazolamide.

Both cases were found to carry the novel change p.Arg418Gly. Interestingly visual inspection of the chromatograms shows an imbalanced ratio between the wild-type and the mutated allele, with the latter significantly less represented (Figure 6-6). This was not observed in the son, suggesting that the father could be mosaic for the mutation.

![Family-5](image.png)

**Figure 6-6. Pedigree of family 5 with the novel Arg418Gly mutation**

The electropherogram of the affected father clearly shows an unbalance ratio between the wild-type and the mutant allele (top panel), with the wild-type significantly more represented and suggesting somatic mosaicism. This was not observed in the proband where the chromatograms of the normal and mutated allele are equivalently represented (bottom panel).
<table>
<thead>
<tr>
<th>Family</th>
<th>Ethnic origin</th>
<th>Age at onset (y)</th>
<th>Age at examination(y)</th>
<th>Mutation</th>
<th>Symptoms at onset</th>
<th>Developmental milestones</th>
<th>Dysarthria</th>
<th>Facial dyskinesias</th>
<th>Chorea</th>
<th>Dystonia</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>UK-Caucasian</td>
<td>1</td>
<td>64</td>
<td>c.1252C&gt;T; p.Arg418Trp (low-level mosaicism)</td>
<td>Chorea</td>
<td>Normal</td>
<td>No</td>
<td>Yes</td>
<td>Yes, mild</td>
<td>No</td>
</tr>
<tr>
<td>2</td>
<td>UK-Pakistani</td>
<td>2</td>
<td>36</td>
<td>c.1252C&gt;T; p.Arg418Trp</td>
<td>Chorea</td>
<td>Delayed</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>3</td>
<td>Italian-Caucasian</td>
<td>4</td>
<td>22</td>
<td>c.1252C&gt;T; p.Arg418Trp</td>
<td>Chorea and lower limb dystonia</td>
<td>Delayed</td>
<td>Yes</td>
<td>Yes, worsened by action and anxiety</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>4</td>
<td>Italian-Caucasian</td>
<td>1</td>
<td>14</td>
<td>c.1252C&gt;T; p.Arg418Trp</td>
<td>Violent dyskinesias upon falling asleep</td>
<td>Delayed</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes, mild</td>
<td>Yes</td>
</tr>
<tr>
<td>5</td>
<td>Italian-Caucasian</td>
<td>3</td>
<td>18</td>
<td>c.1253G&gt;A; p.Arg418Gln</td>
<td>Chorea, dystonia</td>
<td>Normal</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes, mild</td>
<td>Yes</td>
</tr>
<tr>
<td>5</td>
<td>Italian-Caucasian</td>
<td>1</td>
<td>47</td>
<td>c.1252C&gt;G; p.Arg418Gly (mosaicism)</td>
<td>Chorea, dystonia</td>
<td>Delayed</td>
<td>Yes</td>
<td>Yes, worsened by action</td>
<td>Yes</td>
<td>Yes</td>
</tr>
</tbody>
</table>

Table 6-1 Clinical features of cases with pathogenic ADCY5 mutations
<table>
<thead>
<tr>
<th>Other features</th>
<th>Gaze impersistence, difficulty initiating saccades</th>
<th>Gaze impersistence, difficulty initiating saccades</th>
<th>Gaze impersistence</th>
<th>Pyramidal signs in the lower limbs</th>
<th>Frequent myoclonic jerks and pyramidal signs in the lower limbs</th>
<th>No</th>
<th>Severe language development</th>
</tr>
</thead>
<tbody>
<tr>
<td>Paroxysmal events</td>
<td>Ballistic episodes and spasms upon awakening</td>
<td>Ballistic episodes and spasms upon awakening</td>
<td>Ballistic episodes and spasms at night</td>
<td>Triggered by sleep, favoured by emotions</td>
<td>Triggered by stress and sleep</td>
<td>Triggered by stress and sleep</td>
<td>No</td>
</tr>
<tr>
<td>Gait</td>
<td>Mild difficulty on tandem walking</td>
<td>Abnormal due to lower limb choreo-dystonia</td>
<td>Abnormal due to lower limb choreo-dystonia</td>
<td>Dystonic and pyramidal features, frequent falls</td>
<td>Mildly affected with dystonic features</td>
<td>Dystonic</td>
<td>Not applicable</td>
</tr>
</tbody>
</table>
Brain expression data
In order to investigate the differences in the phenotype and clinical course between BHC due to ADCY5 and NKX2-1 mutations, the expression of the two genes across multiple brain regions and during the course of development was compared.

Consistent with the observed phenotype, mainly characterized by chorea and dystonia with preserved cognition and no other major neurological features, both ADCY5 and NKX2-1 mRNA expression profiling in 10 brain regions showed the highest regional expression in the putamen (Figure 6-7 and Figure 6-8). When comparing ADCY5 to NKX2-1 expression, ADCY5 had a much higher expression in adult striatum. Furthermore, longitudinal analysis of mRNA expression during brain development showed opposite trends for ADCY5 and NKX2-1, with ADCY5 expression progressively increasing in the striatum from 50 to 500 days post-conceptualization and NKX2-1 progressively decreasing (Figure 6-7 and Figure 6-8).

6.4.5 Discussion
I report here the results of the ADCY5 mutational analysis in a cohort of NKX2-1 negative BHC cases and show that transmitted and de novo heterozygous ADCY5 mutations are an important cause of familial and sporadic BHC. Overall, in this study pathogenic mutations were identified in 5/53 (9.4%) of cases.

The heterozygous mutation p.Alal726Thr in ADCY5 was identified as the genetic cause in the original FDFM pedigree (Chen et al., 2012). The presentation of the affected members of this family was that of a hyperkinetic choreo-dystonic movement disorder starting from early childhood to late adolescence. A distinctive feature of this disorder was initially thought to be the presence of prominent periorbital and perioral facial myokymia (Fernandez et al., 2001). After the identification of ADCY5 as the responsible gene in the original FDFM pedigree, the same group identified a second pathogenic missense variant in ADCY5 (c.1252C>T; p.Arg418Trp), which occurred de novo in two unrelated individuals with a sporadic complex hyperkinetic movement disorder (Chen et al., 2014). The phenotype of these cases shared some features with the original FDFM pedigree, but it was more severe and further encompassed delayed motor milestones, axial hypotonia and progressive gait difficulties in one case.
Figure 6-7 Graphical Summary of ADCY5 Brain Expression Data

On the left, box plots representing ADCY5 mRNA expression levels. The expression levels are based on exon array experiments and are plotted on a log2 scale (y axis). ADCY5 mRNA expression is higher in the putamen than in all other brain regions.

On the right, graphs to show ADCY5 longitudinal mRNA expression in six brain regions during the course of human brain development. The expression levels are based on exon array experiments and are plotted on a log2 scale (y axis). The plots show increasing expression of ADCY5 mRNA during human brain development, particularly in the striatum, from 50 to 500 days post-conceptualization.
Figure 6-8 Graphical Summary of NKX2-1 Brain Expression Data

On the left box plots representing NKX2-1 mRNA expression levels. The expression levels are based on exon array experiments and are plotted on a log2 scale (y axis). NKX2-1 mRNA expression is higher in the putamen than in all other brain regions. However NKX2-1 expression in the putamen is significantly lower than ADCY5. On the right, graphs to show NKX2-1 longitudinal mRNA expression in six brain regions during the course of human brain development. The expression levels are based on exon array experiments and are plotted on a log2 scale (y axis). The plots show decreasing expression of NKX2-1 mRNA during human brain development, particularly in the striatum, from 50 to 500 days post-conceptualization, an opposite trend compared to ADCY5.
In this chapter, I report seven cases from five unrelated pedigrees bearing pathogenic ADCY5 mutations. Importantly, all the mutations that I identified involved the same amino acid p.Arg418, highlighting a particular relevance of this residue in the disease mechanisms. Three unrelated cases carried the previously reported Arg418Trp mutation, whereas two others carried the novel mutations Arg418Gln and Arg418Gly.

The clinical phenotype of these cases predominantly featured generalized chorea, though with significant intra- and inter-familial phenotypic variability.

As previously described, three of the four mutated subjects with the ADCY5 p.Arg418Trp mutation were more severely affected than subjects from the original FDFM pedigree and carriers of the p.Arg418Gln and p.Arg418Gly. The phenotype of these cases with the p.Arg418Trp included axial hypotonia, delayed motor milestones, and disabling choreo-dystonic dyskinesias.

Conversely, subject II-1 from family 1 had a much milder presentation, with very little functional impact from the dyskinesias. Furthermore, he had isolated chorea and no dystonic features. Interestingly, low-level somatic mosaicism is the likely mechanism for the observed milder phenotype. Indeed, WES data showed a significantly reduced number of reads carrying the mutated allele in the latter case, indicating the presence low-level somatic mosaicism. Somatic mosaicism was likely present also in cases II-1 from family 5 (carrier of the p.Arg418Gly), who, differently from his son, had a predominantly dystonic phenotype and normal motor and language development.

Mosaicism has been recently described in two further cases with ADCY5 mutations and a milder phenotype, suggesting that somatic mosaicism may represent a common explanation for inter- and intra-familial phenotypic variability in ADCY5 mutation carriers (Chen et al., 2015).

The clinical presentation of the subject with the Arg418Gln mutation, who initially presented in childhood with BHC, is now consistent with a diagnosis of myoclonus-dystonia. A similar evolution was also observed in subject III-1 from family I (with the p.Arg418Trp), who predominantly showed chorea during childhood, while dystonic elements became more prominent over the years.

In some of the previously described cases with ADCY5 mutations, dyskinesias are
reported as paroxysmal in the initial phase of the disease (Fernandez et al., 2001) (Chen et al., 2014). In contrast, in all the cases described herein, dyskinesias were constantly present from the onset of symptoms, albeit markedly influenced by emotional states and action. However, two subjects developed superimposed diurnal episodic worsening of dyskinesias, triggered by intense emotional stress. Furthermore, six subjects (excluding individual III-2 from family 5, who is currently only 18 months old) presented exacerbations of hyperkinesias at night, sometimes evolving into episodes of violent ballistic movements and painful spasms. The episodes were particularly frequent at night, either upon falling asleep or awakening. Similar episodes have been described in several other cases with ADCY5 mutations (Chen et al., 2015), indicating that this could be a specific feature of ADCY5-related movement disorders.

Facial myokymia was initially highlighted as a distinctive feature of ADCY5-related disorder. However, an EMG study of the facial muscles performed in both affected subjects from family 1 (both presenting facial involuntary movements) excluded the presence of myokymia or other signs of motoneuron hyperexcitability. Although EMG was not performed in the other cases, who all reported facial dyskinesias, these were clinically deemed to be choreic and/or myoclonic in nature. This indicates that patients with ADCY5 mutations present prominent facial hyperkinetic movements in the absence of myokymia. Consistent with this, neither facial myokymia was described in all affected members from the original FDFM kindred nor was recognized as a prominent feature in the other ADCY5 mutation carriers that were subsequently reported. Importantly, a subsequent EMG study, performed in the same members of the FDFM family in whom myokymia was initially described, failed to confirm the presence of myokymia (Chen et al., 2015).

Importantly, even though the clinical presentation of patients with ADCY5 mutations is consistent with the clinical definition of BHC, there are several differences that differentiate ADCY5-mutation carriers from BHC secondary to variants in NKX2-1.

Firstly, although both conditions can have extra-neural involvement, NKX2-1 carriers often present pulmonary symptoms or evidence of endocrine defects (Inzelberg et al., 2011), whereas cardiac heart failure has been observed in some affected individuals from the original FDFM pedigree (Chen et al., 2012). ADCY5 overexpression has also been shown to lead to cardiomyopathy in a mouse model (Ho et al., 2010).
Secondly, the clinical progression between the two disorders seems to be different; BHC cases secondary to NKX2-1 mutations tend to remain relatively mildly affected and often improve after childhood, while we show here that ADCY5 mutation carriers are often severely affected by the dyskinesias, and may present significant progression of symptoms, at least until the adult age. This difference in the clinical progression is nicely mirrored by the brain transcriptome data. The analysis and comparison of ADCY5 and NKX2-1 mRNA regional brain expression showed that, although both genes are most highly expressed in the adult striatum, ADCY5 levels progressively increase during brain development whereas NKX2-1 levels show an opposite trend. Thirdly, prominent dystonic and myoclonic features are present in several affected cases with ADCY5 mutations and may dominate the clinical picture as the disease progresses. Dystonic features can be also observed in some NKX2-1 cases (Armstrong et al., 2011, Peall et al., 2014), but they are generally not as severe as we observed in ADCY5 mutation carriers. Lastly, ADCY5-related dyskinesias are dramatically worsened by action, excitement or stress. Furthermore, some patients may present frank episodic worsening of the dyskinesias, closely mimicking paroxysmal movement disorders. This feature likely reflects the biological role of AC5. This enzyme catalyzes the formation of cyclic adenosine monophosphate (cAMP) upon β-adrenergic receptor stimulation via G-proteins in striatal cells and pathogenic mutations seem to increase the cyclase activity (Chen et al., 2014).

In conclusion, this study contributes to expand and delineate the phenotype associated with ADCY5 mutations. I suggest the term FDMD should not be used to refer to ADCY5-related movement disorder, as facial myokymias are not part of the phenotype.

A diagnosis of ADCY5-related disorder should be always suspected when assessing patients with a choreic disorder with onset in childhood, in particular if neuroimaging is normal and a fluctuating course with severe dyskinetic attacks triggered by drowsiness are described. Differences in the clinical presentation between ADCY5 mutation carriers may be attributed to distinct molecular effects of different mutations, though somatic mosaicism may explain intra-familial phenotypic variability, as shown in family 1 and 5. Given the possible reported association of ADCY5 mutations with cardiac involvement, it will be fundamental to closely follow-up ADCY5 mutation carriers in order to prevent or adequately treat potential cardiac complications.
Importantly, the majority of the BHC cases included in this project did not carry mutations in either *NKX2-1* or *ADCY5*, including four cases with clear autosomal dominant family history. This indicates that more genes responsible for BHC have yet to be identified.
6.5 *De novo* Mutations in *PDE10A* Cause Childhood-Onset Chorea with Bilateral Striatal Lesions

6.5.1 Statement of Contribution
I performed all the genetic analysis performed in this chapter. I performed the *in silico* modelling of the mutations. Dr Haruhide Kimura’s team (Takeda Pharmaceutical Company Limited) performed the *in vitro* functional characterisation of the identified PDE10A substitution.

6.5.2 Background
WES can be used to identify the disease-causing gene in series of unrelated cases with a very rare and homogenous phenotype, under the assumption that only few causative genes are involved. For example, this strategy has been successful for the identification of *TUBB4A* (MIM 602662) mutations in cases with hypomyelination with atrophy of the basal ganglia and cerebellum (Simons et al., 2013) or *WDR45* (MIM 300526) mutations in an X-linked form of neurodegeneration with brain iron accumulation (Haack et al., 2012).

In this study I identified *de novo* dominant missense mutations in *PDE10A* as a novel disease-associated additional genetic cause in three subjects without a family history who presented with a similar childhood-onset movement disorder predominantly characterized by chorea and striking bilateral striatal abnormalities on cerebral magnetic resonance imaging (MRI).

6.5.3 Subjects, Materials and Methods

**Subjects**
I initially analysed WES data from the eight unrelated BHC cases without pathogenic variants in *NKX2-1* and *ADCY5*, including four with autosomal dominant family history and four sporadic cases. The ascertainment criteria of this cohort have been described in the previous sections of this chapter. For three of the familial cases, DNA and WES from at least another affected relative (1\textsuperscript{st} degree relative in one case and 2\textsuperscript{nd} degree in the other two) were available for segregation analysis.

An additional cohort of 60 cases with BHC of Italian, Dutch and UK origin, negative for mutations in *NKX2-1* and *ADCY5*, was included to replicate the finding made in the
initial cohort. The cases of Italian origin are the same subjects described in the previous section of the chapter. Additional cases were recruited by Dr Manju Kurian at UCL Great Ormond Street Hospital, London (United Kingdom) and Dr Lisenka Visser and Dr Michel Willemsen at Radboud University Medical Center, Nijmegen (Netherlands).

**Genetic analysis**

WES, alignment, variant calling and annotation were performed as previously described in the Materials and Methods chapter. The average read depth of targeted regions was 84.29 and all samples had >95% of the target bases were covered at a read depth of 2x and >90% at a depth of 10x, ensuring high-confidence variant detection. The variant filtration strategy that led to the identification of the PDE10A variants is described in the results section.

**Expression profiling in brain tissue**

Brain mRNA regional distribution and expression in the normal adult human brain were determined using microarray analysis of human post-mortem brain tissue from the UK Human Brain Expression Consortium, as previously detailed in the Materials and Methods chapter.

**In silico modelling of the identified PDE10A mutations**

*In silico* modelling of the 3-D structure of the PDE10A GAF-B domain binding pocket and its interaction with the cAMP, was generated with the software PyMol (Schrodinger) using the PDB-file 2ZMF (downloaded from http://www.rcsb.org/pdb/home/home.do).

**In vitro functional analysis of the identified PDE10A mutations**

**Materials**

1-NO-cAMP was purchased from Biolog Life Science Institute (Bremen, Germany). cAMP and cyclic guanosine monophosphate (cGMP) were purchased from Sigma–Aldrich (St. Louis, MO). The [³H]-labeled nucleotides, [³H]cAMP (25.9 Ci/mmol) and [³H]cGMP (7.1 Ci/mmol), were purchased from PerkinElmer (Waltham, MA).

**Cloning and Expression of Constructs**

cDNA for human PDE10A2 (GenBank: NM_001130690) was used as a template and the mutants (c.898T>C; p.Phe300Leu and c.1000T>C; p.Phe334Leu) were constructed by site-directed point mutation. All constructs were cloned into the pcDNA3.1(+)neo
vector (Thermo Fisher Scientific, Inc., Waltham, MA) and transfected into COS-7 cells (ECACC, Salisbury, UK). The membrane fractions were used for the enzyme assay.

**In Vitro Phosphodiesterase Enzyme Assay**

Phosphodiesterase (PDE) activities were measured using scintillation proximity assay (SPA)-based method (Matthiesen and Nielsen, 2009). In this assay, the product of the PDE reaction, either \[^3\text{H}\]AMP or \[^3\text{H}\]GMP, can bind directly to yttrium silicate PDE SPA beads (GE Healthcare Ltd. UK), leading to light emission from the scintillant in the beads. The enzyme assays were conducted in a buffer (50 mM HEPES-NaOH, 8.3 mM MgCl\(_2\), 1.7 mM EGTA, 0.1% bovine serum albumin (pH 7.4)) in 96-well half-area plates (Corning Inc., Corning NY). For kinetic studies, reaction was conducted at presence of the indicated concentrations of substrate using mixture (20 \(\mu\)L) of \[^3\text{H}\]-labelled and unlabelled cAMP or \[^3\text{H}\]-labelled and unlabelled cGMP with the 20 \(\mu\)L of membrane fractions of PDE10A-expressing COS-7 cells at room temperature, followed by reaction termination by SPA beads addition (20 \(\mu\)L of 20 mg/mL). The plates were counted for 1 min in a scintillation counter (PerkinElmer). For the purpose of determination of baseline counts, every tested concentration of labelled and unlabelled cyclic nucleotides was run with matching controls having the identical reagent concentrations except for the addition of buffer alone in place of enzyme. Specific enzyme activity was then calculated by subtracting the baseline counts.

To obtain the Michaelis–Menten constants (Km), the initial rates of the reaction were fitted to the following equations using GraphPad Prism (GraphPad Software, Inc., La Jolla, CA): 

\[ V = \frac{V_{\text{max}} [S]}{(K_m + [S])} \]

where \(V\) is the initial velocity of the enzyme-catalyzed reaction, \([S]\) is the substrate concentration, \(V_{\text{max}}\) is the limiting reaction velocity at saturating substrate concentrations, and \(K_m\) is the Michaelis–Menten constant (concentration of substrate at 1/2 of \(V_{\text{max}}\)).

The effects of cAMP or its analogue, 1-NO-cAMP, on PDE10A enzyme activity were assessed using \[^3\text{H}\]cGMP as a substrate. The indicated concentrations of unlabelled nucleotide (cAMP or 1-NO-cAMP, 10 \(\mu\)L) and 70 nM \[^3\text{H}\]cGMP (10 \(\mu\)L) were added to the 20 \(\mu\)L buffer containing membrane fractions of PDE10A-expressing COS-7 cells at 37°C. After 40-min incubation, the reaction was terminated by addition of SPA beads (20 \(\mu\)L of 20 mg/mL) and the plates were counted in a scintillation counter. Activity rates were calculated on the basis of 100% control wells without unlabelled nucleotides.
and 0% control wells without enzyme.

## 6.5.4 Results

### Identification of PDE10A mutations in the discovery cohort

A total of 26,387 unique variants were detected in eight exomes of unrelated BHC cases. I excluded from the analysis (i) UTR and intergenic variants outside of canonical splice sites as well as variants that are synonymous, non-frameshift insertions and deletions; (ii) variants that lie in segmental duplication regions, as these are prone to produce false-positive variant calls due to mapping errors; (iii) poor quality variants (quality scores ≤ 50 or read depth ≤ 5).

Given a dominant family history in four out of eight cases and the fact that mutations in the two currently known BHC genes, *NKX2-1* and *ADCY5*, exert a pathogenic effect in the dominant state, I decided to initially focus the analysis only on single heterozygous changes. Furthermore, based on the extreme rarity of BHC in the general population, I assumed that causative variants would be absent in dbSNP version 137 (provided they were not know disease-associated mutations) and in the publicly available datasets of genetic variations (1000 genome project, EVS, ExAC and CG69). I also excluded novel variants that were observed more than five times in our in-house exome database (~1000) from individuals with unrelated diseases, as these are likely to represent sequencing or alignment artefacts. Variants present in the familial cases that did not segregate in the affected relatives were filtered out.

After applying this filtering strategy, I was left with a list of individual variants that included 18 frameshift indels, 11 nonsense mutations, four mutations affecting canonical splice-sites, and 212 missense variants.

I subsequently focused only the genes that contained at least a variant in more than one sample. This filtering strategy left us with only three candidate genes (see Table 6-2), each containing a variant in two unrelated samples, *PDE10A* (phosphodiesterase 10A) on chromosome 6, *THSD1* (thrombospondin, type I, domain containing 1) located on chromosome 13, and *PRMT7* (protein arginine n-methyltransferase 7) on chromosome 16.
I then explored the regional expression of these three genes in the normal adult human brain. This analysis showed that PDE10A expression is exceptionally high in the striatum (Figure 6-9A), whereas this is not the case for THSD1 and PRMT7 (see Figure 6-10). These data are consistent with the data available on the Allen Mouse Brain Atlas (Lein et al., 2007) (see Figure 6-9B and C) and previous work in the literature, demonstrating high and selective expression of PDE10A in the striatum (Fujishige et al., 1999), particularly in MSNs (Seeger et al., 2003), which makes PDE10A an excellent candidate gene for chorea.

The two cases with PDE10A mutations carriers were both sporadic and carried the exact same heterozygous change, c.898T>C; p.Phe300Leu.

Sanger sequencing confirmed the presence of the mutation in both cases. The mutation was demonstrated to be de novo in the first case, as it was absent in both parents and in the unaffected 19-year old sister. To confirm that no other variants could be responsible for chorea in this subject, WES was performed in all family members, including the two parents and the healthy sister. Importantly, I could not identify other putatively damaging mutations (defined as nonsense, frameshift, canonical splice site, predicted damaging missense mutations based on CADD scores >20; with a MAF <1% in ExAC), assuming either a de novo or a recessive pattern of transmission.

Parents of the second individual were both deceased and therefore not available for DNA analysis. However, DNA analysis in six siblings, all unaffected with chorea, demonstrated that they were all homozygous for the reference allele. Furthermore, haplotype analysis using three microsatellites (Marker 1 – chr6:166069747-166069785; Marker 2 – chr6:165862198-165862227; Marker 3 - chr6:165839259-165839288) spanning the PDE10A locus identified the four parental haplotypes and revealed that the individual carrying the mutation shared one of the haplotypes with two siblings and the other with three other siblings, strongly indicative of the de novo occurrence of the mutation also in this case (Figure 6-11). Analysis of the same three microsatellites in the family of case 1, who carries the same de novo PDE10A change, indicates the mutation arose on a different background haplotype.
Table 6-2 Summary of variants detected by whole-exome sequencing that survived the filtering strategy.

<table>
<thead>
<tr>
<th>Gene</th>
<th>Chr</th>
<th>Position (hg19)</th>
<th>Variant</th>
<th>GERP score(^a)</th>
<th>CADD C-score(^b)</th>
<th>SIFT</th>
<th>PolyPhen-2</th>
<th>Mutation Taster</th>
<th>ExAC constraint metrics</th>
<th>Number of unrelated probands with the variant</th>
</tr>
</thead>
<tbody>
<tr>
<td>PDE10A</td>
<td>6</td>
<td>165832223</td>
<td>c.898T&gt;C; p.(Phe300Leu)</td>
<td>4.47</td>
<td>28.7</td>
<td>D (0)</td>
<td>D (0.98)</td>
<td>D (0.99)</td>
<td>Missense: z(^c) 3.78 LoF: pLi(^d) 1.00</td>
<td>2 (both sporadic)</td>
</tr>
<tr>
<td>THSD1</td>
<td>13</td>
<td>52952435</td>
<td>c.1669_1670del; p.(Ser557fs)(^d)</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>Missense: z(^c) 0.75 LoF: pLi(^d) 0.00</td>
<td>1 (familial)(^f)</td>
</tr>
<tr>
<td></td>
<td>13</td>
<td>52952567</td>
<td>c.1379G&gt;A; p.Ser460Asn</td>
<td>5.62</td>
<td>12.16</td>
<td>T (0.57)</td>
<td>B (0.02)</td>
<td>D (0.99)</td>
<td></td>
<td>1 sporadic</td>
</tr>
<tr>
<td>PRMT7</td>
<td>16</td>
<td>68373391</td>
<td>c.522dupT; p.(Ser174fs)(^d)</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>Missense: z(^c) 0.23 LoF: pLi(^d) 0.00</td>
<td>1 (familial)(^f)</td>
</tr>
<tr>
<td></td>
<td>16</td>
<td>68382285</td>
<td>c.1214T&gt;C; p.(Ile405Thr)</td>
<td>5.6</td>
<td>24.2</td>
<td>D (0.01)</td>
<td>B (0.03)</td>
<td>D (0.99)</td>
<td></td>
<td>1 (familial)</td>
</tr>
</tbody>
</table>

Chr=chromosome; B=benign; D=deleterious/damaging/disease-causing; T=tolerated; NA=not applicable; fs=frameshift; LoF=loss-of-function; pLI= probability of loss-of-function intolerance

\(^a\)Positive scores represent a substitution deficit and indicate that a site may be under evolutionary constraint. Negative scores indicate that a site is probably evolving neutrally. Positive scores scale with the level of constraint, such that the greater the score, the greater the level of evolutionary constraint inferred to be acting on that site.

\(^b\)C-scores greater or equal 10 indicates that the variant is predicted to be the among the 10% most deleterious substitutions that you can do to the human genome; a score of greater or equal 20 indicates the 1% most deleterious.

\(^c\)z-missense scores reflect expected mutation probabilities and are based on deviation of observed standing variation patterns (higher values = more constrained [fewer variants observed than expected]).

\(^d\)The closer pLI is to 1, the more LoF intolerant the gene appears to be.

\(^e\)The two frameshift variants in THSD1 and PRMT7 were found in the same individual.
De novo mutations in PDE10A have not been observed in control individuals (Rauch et al., 2012, Xu et al., 2012, Gulsuner et al., 2013, Genome of the Netherlands Consortium, 2014, Iossifov et al., 2014) and the p.Phe300Leu is neither listed in ExAC, nor in our in-house exomes with unrelated diseases (~1,000), and in 500 exomes from ethnically matched healthy controls. All used in silico tools consistently predicted a highly deleterious effect of the substitution. Interspecies alignment of protein sequences revealed that the amino acid substitution lies in an acid motif that is completely conserved down to invertebrate species (see Figure 6-12). Constraint metrics reported in ExAC indicate that PDE10A is highly intolerant to both loss-of-function and missense mutations (Samocha et al., 2014).

The frameshift mutations in THSD1 (p.Ser557fs) and PRMT7 (p.Ser174fs) were both found in the same subject with a dominant family history of BHC and were present also in the similarly affected sister (the DNA of the affected father was not available). The missense heterozygous changes in THSD1 and PRMT7 were carried by a sporadic case and a familial case, respectively. No family members of the carrier of the THSD1 missense mutation were available for segregation analysis, whereas the PRMT7 missense change were inherited from the affected mother and grandmother of the proband. Prediction tools displayed contrasting results for both missense mutations. Importantly, although a pathogenic role for this variant cannot be excluded, the ExAC constraint metrics indicate that both THSD1 and PRMT7 are tolerant to missense and loss-of-function variation (see Table 6-2), suggesting it is unlikely that heterozygous mutations in these two genes could be responsible for the phenotype.

Finally, bi-allelic variant variants in PRMT7 have been recently detected in three pedigree with a recessive condition called Albright hereditary osteodystrophy, which feature mild intellectual disability with obesity and symmetrical shortening of the digits and posterior metacarpals and metatarsals (Akawi et al., 2015). This element further argues against any pathogenic role for the heterozygous PRMT7 variants I identified in patients with BHC.
Figure 6.9 Summary of human and mouse brain PDE10A mRNA expression data

(A) Box plots of PDE10A mRNA expression levels in 10 adult brain regions. The expression levels are based on exon array experiments and are plotted on a log2 scale (y axis). This plot shows significant variation in PDE10A expression across the 10 brain regions analysed, with expression higher in the putamen than in any other region. PDE10A expression in mouse brain in (B) sagittal and (C) coronal sections. PDE10A is very highly and selectively expressed in the striata and in the olfactory tubercula. Images were obtained from the Allen Mouse Brain Atlas website (© 2015 Allen Institute for Brain Science). Expression intensity is color-coded, ranging from blue (low intensity) through green and yellow to red (high intensity).
Box plots of mRNA expression levels in 10 adult brain regions showing that the expression in the putamen of *PRMT7* (A) and *THSD1* (B) is significantly lower than *PDE10A*.

**Figure 6-10 Summary of human brain expression of PRMT7 and THDS1**

Fold change between CRBL and WHMT = 1.3 ($p=2\times10^{-36}$)

Fold change between HIPP and MEDU = 1.3 ($p=1.9\times10^{-30}$)
Figure 6-11 Pedigree and haplotype analysis in the families of cases 1 and 2
The four parental haplotypes were reconstructed in both families (each haplotype defined by a different colour). Cases 1 and 2 did not share the haplotype encompassing the c.898C>T variant, suggesting the mutations arose on different haplotype backgrounds. Furthermore, haplotype analysis indicates that case 2, who carries the PDE10A c.898T>C variant, shares one of the allele (marked in yellow) with two unaffected siblings (II-2, II-7) whereas the other allele (marked in red) is shared with three unaffected siblings (II-1, II-4 and II-7).
**Figure 6-12 Interspecies alignment of the identified pathogenic PDE10A mutations**

Interspecies alignment showing the complete conservation down to invertebrates of the amino acid residues involved by the substitutions. Asterisks indicate invariant residues (full conservation), whereas a colon (:) and period (.) represent strong and moderate similarities, respectively.
Sanger Sequencing of PDE10A in a Cohort of Childhood-Onset Chorea Cases

As these results outlined in the previous section strongly support the pathogenic role of the PDE10A c.898T>C; p.Phe300Leu mutation, I then Sanger sequenced the entire coding sequence and the exon-intron boundaries of PDE10A in a further 60 cases with BHC without mutations in NKX2-1 and ADCY5.

Importantly, this analysis revealed the presence of two additional variants: the unreported c.1000T>C; p.Phe334Leu was found in a Dutch case, and the c.1459C>G; the p.Pro487Ala (rs61733392) was found in an Italian case.

Sequencing in the parents of the carrier of the c.1000T>C; p.Phe334Leu variant showed that it was absent in both subjects, demonstrating the de novo occurrence of the mutation and highly supporting its pathogenic role. WES was performed also in this family and no other possibly pathogenic mutations, de novo or bi-allelic, were found.

This mutation is absent in all publicly available datasets, in all our exomes from healthy controls and in all exomes from cases with unrelated diseases. Furthermore, the change is predicted pathogenic by all in silico tools (Polyphen-2, SIFT and MutationTaster) and has a CADD score of 31. Interspecies alignment showed complete interspecies conservation (see Figure 6-12).

The p.Pro487Ala is a rare benign coding polymorphism, present in 478 European alleles reported in ExAC (MAF 0.007; two subjects are homozygotes) and is predicted benign by both Polyphen-2 and SIFT.

Clinical and Neuroradiological Features of Carriers of Pathogenic PDE10A mutations

Case 1, carrier of the Phe300Leu substitution, is a 22-year-old woman who developed generalized chorea around the age of 8. Although developmental motor milestones were reached normally, she had mild difficulties with speech and coordination during childhood. She now suffers from anxiety with some obsessive features. On examination she showed selective upgaze hypometric saccades and a mixed movement disorder consisting mainly of generalised chorea associated with some myoclonic jerks.

Case 2, carrier of the Phe300Leu substitution, is a 60-year-old woman who developed chorea around the age of 5. She had a normal birth and normal developmental
milestones. The chorea remained stable over the years, with parkinsonian symptoms becoming apparent in her fifth decade with freezing and falls. Striatal dopamine reuptake transporter density imaging was bilaterally abnormal, consistent with nigrostriatal dopaminergic denervation. On her last examination, there was generalised chorea, most prominently in the orofacial area, with motor and gaze impersistence and severe dysarthria. She also had generalised bradykinesia, more prominent in the lower than in upper limbs, and marked axial parkinsonian symptoms (freezing and postural instability), which responded well to a L-DOPA treatment (600 mg/day), without worsening of the choreic movements.

Case 3, carrier of the p.Phe334Leu substitution, is an 11-year-old, who developed an isolated non-progressive choreic disorder at around age 5. He had normal motor and language development.

Very importantly, brain MRI consistently showed striking bilateral T2 hyperintensity within the striatum in all three cases (Figure 6-13). It is noteworthy that the MRI images of case 3 (aged 11 when scanned) showed slight swelling of the striata (Figure 6-13A), together with restricted diffusion (Figure 6-13B and C), suggesting an active disease process. Conversely, MRI of case 1 (aged 22 when scanned) demonstrated modest atrophy of the putamina (Figure Figure 6-13D) and normal diffusion (Figure Figure 6-13 E and 1F), suggesting a more advanced stage of disease. The MRI of case 3 (aged 53 when scanned) was markedly degraded by movement artefacts, but also showed T2 bilateral hyperintensities within the posterolateral putamina, albeit less dramatic than in the two younger cases.
Figure 6-13 MRI features associated with de novo dominant *PDE10A* mutations

Axial MR images of case 3 (A-C) and 1 (D-F). There is increased signal intensity within the striatum on T2-weighted images (A, D) and diffusion-weighted images (DWI) (B, E). In case 3, the putamen and caudate nucleus appear slightly swollen (A) and high signal on DWI (B) is confirmed to represent abnormal restricted diffusion on the ADC map (C). In case 1, the abnormal signal is principally located in the postero-lateral putamina, which also appear atrophic (D). There is no corresponding restriction of diffusion on the ADC map (F), and appearances suggest a more chronic disease stage.

Overall, the three individuals presented with a very similar and unique clinical presentation of childhood-onset chorea and characteristic brain MRI (Table 6-3), strongly supporting *PDE10A* mutations as the unitary aetiology underlying all three cases.
<table>
<thead>
<tr>
<th></th>
<th>Case 3</th>
<th>Case 1</th>
<th>Case 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age at most recent clinical examination (years)</td>
<td>11</td>
<td>22</td>
<td>60</td>
</tr>
<tr>
<td>Gender</td>
<td>Male</td>
<td>Female</td>
<td>Female</td>
</tr>
<tr>
<td>Descent</td>
<td>European (Dutch)</td>
<td>European (British)</td>
<td>European (British)</td>
</tr>
<tr>
<td><strong>PDE10A mutation</strong></td>
<td>Genomic&lt;sup&gt;a&lt;/sup&gt;</td>
<td>Chr 6:165829768 A&gt;G</td>
<td>Chr 6:165832223 A&gt;G</td>
</tr>
<tr>
<td>cDNA&lt;sup&gt;b&lt;/sup&gt;</td>
<td>c.1000T&gt;C</td>
<td>c.898T&gt;C</td>
<td>c.898T&gt;C</td>
</tr>
<tr>
<td>Protein</td>
<td>p.Phe334Leu</td>
<td>p.Phe300Leu</td>
<td>p.Phe300Leu</td>
</tr>
<tr>
<td>Inheritance</td>
<td>de novo</td>
<td>de novo</td>
<td>de novo&lt;sup&gt;d&lt;/sup&gt;</td>
</tr>
<tr>
<td>CADD score&lt;sup&gt;c&lt;/sup&gt;</td>
<td>31.0</td>
<td>28.7</td>
<td>28.7</td>
</tr>
<tr>
<td><strong>Neurology</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Developmental milestones</td>
<td>Normal</td>
<td>Normal</td>
<td>Normal</td>
</tr>
<tr>
<td>Cognition</td>
<td>Normal</td>
<td>Normal</td>
<td>Normal</td>
</tr>
<tr>
<td>Chorea (age of onset, yrs)</td>
<td>+ (5)</td>
<td>+ (8)</td>
<td>+ (5)</td>
</tr>
<tr>
<td>Other</td>
<td>No</td>
<td>Anxiety</td>
<td>Adult-onset parkinsonism</td>
</tr>
<tr>
<td><strong>MRI</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bilateral striatal hyperintensities</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Bilateral striatal swelling</td>
<td>+</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Restriction of diffusion</td>
<td>+</td>
<td>-</td>
<td>N.A.</td>
</tr>
<tr>
<td>Bilateral striatal atrophy</td>
<td>-</td>
<td>+</td>
<td>+</td>
</tr>
</tbody>
</table>

+=present; N.A.=not available
<sup>a</sup>(GRCh37)
<sup>b</sup>Ref Seq ID: NM_001130690.2
<sup>c</sup>Haaplotype analysis in unaffected siblings suggests the de novo occurrence of the mutation in case 3
In silico Modelling of the Identified PDE10A Mutations

PDE10A encodes a member of the cyclic nucleotide (cNMP) phosphodiesterase (PDE) family, consisting of 21 different genes, grouped into 11 sub-families based on their affinity for the type of cNMP (cAMP and/or cGMP), cellular regulation, expression and tissue distribution (Lakics et al., 2010). cNMPs are ubiquitously expressed intracellular second messengers, which modulate a broad range of cellular functions and pathways (Beavo and Brunton, 2002). The intracellular concentration of cNMPs is tightly regulated through a fine balance between their synthesis, controlled by the activity of adenylyl/guanylyl cyclases (Koesling et al., 1991, Steegborn, 2014), and degradation, mediated by PDEs which hydrolyse the cNMPs into their corresponding monophosphate nucleoside (Soderling and Beavo, 2000).

PDEs function as homodimers, with the dimer interface extending over the entire length of the molecule, and all share a highly similar catalytic domain located in the C-terminal portion of the protein. Conversely, the N-terminal portion, which contains the regulatory domains, is variable and differs between different PDE families (Bender and Beavo, 2006).

PDE10A contains two N-terminal domains, GAF-A and GAF-B, of which the latter binds to cAMP (Figure 6-14A) (Gross-Langenhoff et al., 2006, Heikaus et al., 2009). cAMP binding increases the enzyme activity of the PDE10A catalytic domain (Jäger et al., 2012). Although details of the GAF-B dependent modulation of PDE10A enzyme activity are currently unclear, a general mechanism for the regulation of all PDEs has been postulated. In the non-activated state the dimerised catalytic domains are packed against each other at the dimer interface, occluding the catalytic pockets. The binding of cAMP to the GAF-B domain induces a rotating movement of the catalytic domains, enabling substrate access to the catalytic pockets and a consequent increase of cNMP hydrolysis (Pandit et al., 2009).

The crystal structure of the PDE10A-GAF-B domain and its interaction with cAMP has been elucidated and consists of six stranded anti-parallel β-sheet (β3, β2, β1, β6, β5, β4), sandwiched between a three-helix bundle (α1, α2, and α5) on one side and three short helices (α3, α4, 3₁₀) on the other side (Handa et al., 2008). The cAMP molecule is almost completely buried deep into a tight binding pocket, the floor of which is formed by the β-sheets and the roof by two α-helices (α3 and α4). Importantly, the amino acids
p.Phe300 and p.Phe334 are located in the β1 and β3 sheets, positioned deep into the cAMP binding pocket of GAF-B and in very close proximity to the cAMP molecule (Figure 6-14B). It is therefore postulated that the substitutions severely affect the morphology of the GAF-B binding pocket and/or alter its affinity for cAMP.

**Figure 6-14 Schematic representation of the PDE10A protein and in silico modelling of the 3D structure of the PDE10A GAF-B domain**

(A) A schematic representation of the PDE10A protein showing its organization in three domains, the regulatory GAF-A and GAF-B domains located in the N-terminal portion of the protein and the catalytic domain located in the C-terminus. (B) In silico modelling of the 3-D structure of the GAF-B domain binding pocket and its interaction with the cyclic adenosine monophosphate (cAMP; shown in blue). The mutated residues p.Phe300 and p.Phe334 and their aromatic side chains, located in the β1 and β3 sheets forming the floor of the cAMP binding pocket, are shown in red and cyan respectively. Both mutated residues (in magenta) are located in very close proximity to the cAMP molecule and are therefore likely to play an essential role in nucleotide binding.
**In vitro Functional Analysis of the Identified PDE10A Mutations**

To assess *in vitro* the functional effect of the identified PDE10A substitutions, we investigated whether they affect (i) PDE basal enzyme activity and/or (ii) the stimulatory effect on PDE catalytic activity mediated by cAMP binding to the GAF-B domain.

Reactions for kinetic studies were performed using a mixture of $[^3]$H-labelled and unlabelled cAMP or cGMP together with either WT or mutant PDE10A-expressing COS-7 cell membrane fractions. Km values for cAMP and cGMP in wild-type PDE10A were $152.5 \pm 17.1 \text{ nM}$ and $1.7 \pm 0.3 \text{ \mu M}$, respectively. There were no statistically significant differences in Km values among the wild-type and the mutant PDE10As (Figure 6-15), suggesting that both p.Phe300Leu and p.Phe334Leu do not affect substantially basal PDE10A enzyme activity.

We then explored whether the identified substitutions affect the stimulatory properties of cAMP binding to the GAF-B domain. Experiments were conducted using only $[^3]$HcGMP as a substrate to avoid the binding of $[^3]$HcAMP substrate to the GAF-B domain and the cAMP analogue 1-NO-cAMP, which has a higher selectivity for the GAF-B domain over the catalytic site compared to cAMP (247-fold for 1-NO-cAMP vs. 8.7-fold for cAMP). These experimental conditions were chosen as, on the one hand, cAMP activates PDE10A enzyme activity via its binding to GAF-B and, on the other hand, competes at the catalytic domain with radio-labelled substrates and thus inhibits their degradation.

Wild-type and mutant PDE10As were incubated in the presence of various concentrations of 1-NO-cAMP with 70 nM $[^3]$HcGMP, and then catalytic activities were measured. 1-NO-cAMP markedly increased (approximately 2.7-fold over the basal levels) the enzyme activity of WT PDE10A, whereas this effect was almost completely abolished for both mutant PDE10As (Figure 6-16). These experiments demonstrate that p.Phe300Leu and p.Phe334Leu severely affect the positive regulatory mechanism of cAMP binding to the GAF-B domain on PDE catalytic activity.
Figure 6-15 Enzyme Kinetics of Wild-Type and Mutant PDE10As.
PDE10A enzymes was incubated at presence of mixture of unlabelled cAMP and [3H]cAMP (A) or unlabelled cGMP and [3H]cGMP (B) with the total concentration as indicated. The velocity was divided by each Vmax value. Km values for cAMP (A) and cGMP (B) were determined by fitting to the Michaelis-Menten equation using GraphPad Prism. Each data point represented as mean ± SEM of five (for cAMP) or four (for cGMP) independent experiments. There was no statistically significant difference in Km values among the wildtype and the mutant PDE10As (p > 0.05 by Dunnett's test compared with wild-type).
Effect of 1-NO-cAMP on the PDE activity of a wild-type and mutants of PDE10A. Various concentrations of a cAMP analogue, 1-NO-cAMP, was added and catalytic activity of wild-type and mutant PDE10A, p.(Phe300Leu) and p.(Phe334Leu), at presence of 70 nM [3H]cGMP was determined. Each data point represents the means ± SEM for three independent experiments.
6.5.5 Discussion

Strong genetic evidence, fulfilling the recently proposed criteria for inferring causality of genetic variants (MacArthur et al., 2014), supports the pathogenic role of the identified PDE10A mutations: (i) segregation analysis confirmed that both variants arose de novo in the three mutation carriers; (ii) the changes are otherwise absent in all interrogated public databases and in-house exomes from healthy controls or subjects with unrelated conditions (including more than 70,000 exomes); (iii) the mutations are consistently predicted pathogenic by all in silico prediction tools and affect amino acid residues completely conserved in all species homologs; (iv) the substitutions are located in a functionally important regulatory domain of the protein and in vitro functional data show that they disrupt the regulatory effect of the GAF-B domain; (v) PDE10A is very highly (and almost selectively) expressed in the striatum, the brain area most closely involved in the pathogenesis of choreic movements.

PDEs have previously been implicated in the pathogenesis of neurodegenerative disorders, such as PD and HD (Bollen and Prickaerts, 2012). Mutations in PDE8B, a gene highly expressed in the brain and especially in the putamen, causes autosomal dominant striatal degeneration (ADSD, MIM 609161), a disease that clinically presents with adult-onset parkinsonism (Appenzeller et al., 2010, Barsottini et al., 2015). Although the reported MRI abnormalities observed in subjects with ADSD are slightly different from those observed in our cases, it is striking that both diseases are caused by mutations in PDEs leading to clearly visible, largely symmetric, striatal MRI signal abnormalities. Furthermore, the fact that two PDEs are now directly linked to a basal ganglia disease may point towards a crucial role of PDEs in these types of disorders. The latter is of great interest given the pharmacological potential to manipulate PDE activity. Given its high and selective expression in striatal MSNs, PDE10A is a primary target in pharmacological research for diseases where dysregulation of striatal circuits is believed to be crucial (e.g. psychosis, HD, substance abuse and PD) (Chappie et al., 2012).

According to the classic model of basal ganglia motor circuits, chorea may result from dysregulation of the activity of the striatal MSNs (Marsden, 1984). Based on different anatomical projections and biochemical features, MSNs are divided into two classes that constitute two parallel and segregated pathways exiting the striatum; (i) the ‘direct pathway’, formed by MSNs expressing high levels of the D1 dopamine receptors
(D1DR) and projecting directly to the basal ganglia output nuclei (the substantia nigra pars reticulata and the globus pallidus internal segment); (ii) the ‘indirect pathway’, mainly formed by MSNs expressing both D2 dopamine receptors (D2DR) and adenosine 2 receptors (A2AR) and projecting indirectly to the globus pallidus internal segment via the globus pallidus external segment and the subthalamic nucleus (Gerfen, 1992).

The two pathways control striatal output in an antagonistic way, the ‘direct pathway’ being involved in action production and the ‘indirect pathway’ in action suppression (DeLong and Wichmann, 2007). According to the classic model of basal ganglia motor circuits, chorea mainly results from disruption of the indirect pathway with consequent overactivity of the direct pathway (Marsden, 1984).

Importantly, modulation of MSN activity is largely dependent on cAMP signalling (Threlfell and West, 2013). cAMP synthesis, and thus indirectly its signalling, is promoted by the stimulation of the G protein-coupled receptors D1DR and A2AR, whereas synthesis is inhibited by dopamine stimulation of D2DR (Herve, 2011). The G protein $G_{\alpha_{olf}}$ positively couples D1DR and A2AR to the activation of AC5, the main molecule responsible for cAMP production in MSNs (Lee et al., 2002). Interestingly, the genes encoding $G_{\alpha_{olf}}$ (GNAL) and AC5 (ADCY5) have both been identified as cause of primary dystonia (Fuchs et al., 2013) and chorea (Chen et al., 2012) respectively.

Mechanistically, $ADCY5$ mutations seem to increase the AC5 activity with consequent raised intracellular cAMP levels in cellular models (Chen et al., 2014). As both $PDE10A$ and $ADCY5$ pathogenic mutations cause chorea, but with PDE10A exerting an opposite effect to AC5 on cAMP levels, one would expect that the p.Phe300Leu and p.Phe334Leu variants exert a deleterious effect on the PDE enzyme activity. Recent studies suggest that PDE10A has two functional states; ‘active’ and ‘super-active’ (Jäger et al., 2012, Russwurm et al., 2015). In presence of high intracellular levels of cAMP, its binding to the GAF-B domain would stimulate the PDE catalytic activity, switching PDE10A from the ‘active’ to the ‘super-active’ state. In light of this, PDE10A may function as a ‘brake’ for MSN activation. Our functional studies show that pathogenic $PDE10A$ mutations located in the GAF-B domain severely disrupt this positive regulatory mechanism without affecting the basal PDE enzyme activity. These mutations may therefore have a strong impact on the in vivo regulation of MSN activity,
especially when MSNs are activated by high levels of cAMP. Given the homodimerised structure of PDE10A, the mutant proteins could exert a dominant negative effect on the activity of the WT protein. Intriguingly, the PDE10A knockout mouse does not show a movement disorder reminiscent of chorea, but conversely presents reduced exploratory activity, more consistent with a parkinsonian phenotype (Siuciak et al., 2006, Siuciak et al., 2008).

In conclusion, I demonstrate here that de novo dominant mutations in PDE10A are a novel cause of a unique movement disorder characterized by benign childhood-onset chorea and typical MRI abnormalities of the striatum. Of note, PDE10A mutations were not found in cases with a BHC syndrome but with normal brain MRI. This suggests that PDE10A-related chorea may represent a distinct genetic clinico-radiological entity. Mutational screening of additional cohorts of cases with such MRI abnormalities is warranted to further define the clinical spectrum associated with PDE10A mutations.
Chapter 7. Conclusions

This thesis describes the work I conducted during my PhD to investigate the genetic bases of a range of movement disorders, including PD, dystonia, and chorea, and gain novel insights into the biological mechanisms of these conditions.

I used various combinations of genetic technologies (i.e. Sanger sequencing, WES and linkage analysis) and strategies (i.e. linkage mapping in combination with WES, interrogation of WES data searching for mutations in candidate genes, and analysis of WES from a series of unrelated cases with a specific clinical and radiological presentation).

My research has resulted in the identification of several novel genetic causes for these conditions, including genes harbouring high-risk fully penetrant mutations responsible for Mendelian forms of disease (e.g. \textit{KCTD17} mutations in M-D and \textit{PDE10A} mutations in chorea) and rare variants associated with intermediate risks (e.g. \textit{GCH1} mutations in PD). I also expanded the phenotype associated with known Mendelian genes. I recognised that deficiency of tyrosine hydroxylase is a novel, treatable cause of M-D and mutations in \textit{ADCY5} are not simply responsible for the unique condition Familial Dyskinesias with Facial Myokyimia, but actually represent a frequent cause of chorea with onset in childhood.

The main findings of this thesis are summarised below, together with the implications deriving from this work, both concerning clinical practice and future research into disease mechanisms.

7.1 Genetics of Parkinson Disease

\textit{GBA Mutations Are Common in Early-Onset Parkinson Disease and Influence the disease course}

I showed that \textit{GBA} mutations are the most common genetic abnormality, not only in patients with the classic late-onset form of the disease, but also in patients with early-onset PD. The frequency of \textit{GBA} mutations in the latter group is actually significantly higher than in patients with classic PD.
This finding possesses an important clinical corollary. GBA should be systematically assessed in patients with an early onset of symptoms, as information regarding the mutational status may have crucial implications in terms of prognosis and response to pharmacological or surgical treatments (i.e. DBS). It is thought that patients with early-onset PD, and in particular those carrying mutations in the recessive genes PARK2 or PINK1, tend to have a more indolent and benign disease course, in particular with regard to the risk of developing cognitive decline (Schrag and Schott, 2006).

However, this seems not to be true for GBA carriers, both from a motor and cognitive point of view. In our study, PD patients carrying GBA mutations, compared to patients with mutations in PARK2, LRRK2 and no mutations, needed DBS significantly earlier in their disease, suggesting a more aggressive form of the disease in PD patients with GBA mutations. Furthermore, although the motor response to DBS treatment was comparable to that of the other patients, patients with GBA mutations presented a more rapid cognitive impairment following DBS implantation. Although this aspect should be carefully assessed in larger cohorts, these data overall indicate the genetics may be a powerful tool in recognizing and predicting clinical heterogeneity in PD.

The GBA E326K variant predisposes to Parkinson disease

I established incontrovertibly the pathogenic role of the GBA E326K variant, which, to date, represents the single most common genetic risk factor for PD with significant clinical relevance (7.5% of the PD cases screened in our study were carriers of the variant; odds ratio ~3). Notably, these results approximately double the proportion of PD patients with GBA-related disease, a fact that will have tremendous implications as soon as treatments for addressing GBA dysfunction in PD become available.

Confirming my finding, Nalls and colleagues recently demonstrated the association of the E326K with PD-dementia complex and Dementia with LBs (Nalls et al., 2013). Moreover, it has been shown that the presence of the E326K negatively influences both motor disease progression (Winder-Rhodes et al., 2013) and the risk of progression to dementia (Mata et al., 2015).

Finally, the E326K predisposes to PD but does not cause GD (even when inherited in the homozygous state), clearly indicating that other molecular mechanisms, in addition to reduced GBA activity, may contribute to the development of PD disease in the presence of GBA dysfunction.
Beyond DOPA-responsive dystonia; GCH1 mutations: a novel risk for Parkinson disease

Through the analysis of WES data in a large cohort of PD cases and controls, I have identified rare coding *GCH1* as a novel risk locus for PD. Interestingly, as it was the case for the discovery of *GBA* mutations, this association was first suspected on clinical grounds, thanks to the observation in pedigrees with DRD of striatal dopaminergic denervation in *GCH1* pathogenic variant carriers with a clinical diagnosis of PD.

Interestingly, to date, no genes containing rare variants conferring intermediate risk for PD have been identified through unbiased large NGS studies. Therefore, astute clinical observations remain crucial for driving and inspiring genetic studies.

Since the appearance of this study in the literature, several other reports of DRD families with members affected by PD have been published (Lewthwaite *et al.*, 2015, Terbeek *et al.*, 2015, Someko *et al.*, 2016) and many others have been presented at scientific conferences (personal communication).

Furthermore, a recent study performed within the Canadian population confirmed that rare *GCH1* coding mutations are enriched in PD. The authors identified two variants, the known pathogenic p.Lys224Arg (x2) and the novel variant p.Ala99Asp (likely pathogenic according to in silico prediction tools and interspecies conservation) in three unrelated cases with PD, while no pathogenic changes were found in controls (Guella *et al.*, 2015). The mutational frequency was 0.56% (3/528) in cases versus 0% (0/290) in controls, consistent with the frequency we observed in our study (0.75% in cases versus 0.1% in controls).

The finding that *GCH1* loss-of-function variants are not only responsible for DRD, but are also associated with adult-onset neurodegenerative PD, is strengthened by the recent identification that *GCH1* has also been identified as a low-risk susceptibility locus for PD in the latest GWAS meta-analysis (Nalls *et al.*, 2014). This finding potentially extends the role of GTPCH deficiency in the pathogenesis of PD beyond the rare carriers of deleterious coding mutations.

An outstanding and crucial question remains as to whether patients with DRD eventually develop nigral neurodegeneration, or whether neurodegeneration can be avoided by dopaminergic replacement therapy. This could indeed support a central role
of dopamine deficiency in favouring nigral degeneration, with important therapeutic implications for asymptomatic carriers of pathogenic variants.

To date, no DRD cases treated with L-DOPA since childhood have been shown to develop frank nigral cell loss (Snow et al., 1993, Turjanski et al., 1993, Jeon et al., 1998). This is consistent with post mortem analysis of few cases showing normal nigral cell count and absence of LBs (Furukawa et al., 1999, Grotzsch et al., 2002, Segawa et al., 2013). However, Sawle and colleagues report in six cases with DOPA-responsive dystonia modest but significant reduction in the uptake of $^{18}$F-fluorodopa into both caudates and putamen (Sawle et al., 1991). Tadic and colleagues describe that parkinsonian signs are present in DOPA-responsive dystonia cases as a relatively common residual motor sign following treatment, possibly hinting at underlying neurodegeneration (Tadic et al., 2012). Terbeek et al. recently described a 41-year-old patient carrying a known pathogenic $GCH1$ variant (p.Tyr75Ser) with onset of classic DRD at age 9 (Terbeek et al., 2015). He was treated with L-DOPA from the age of 11 with good and sustained response. At age 41 dopaminergic imaging ($^{123}$I-FP-CIT SPECT) was performed showing severe bilateral and asymmetric reduction of putaminal tracer uptake, a pattern typical of idiopathic PD, arguing against a neuroprotective role of dopamine replacement.

I anticipate that post mortem analysis and longitudinal clinical, neuroimaging, and metabolic studies of $GCH1$ mutation carriers—including asymptomatic carriers, individuals with classic DRD and cases with adult-onset parkinsonism - will facilitate the understanding of the pathogenesis of $GCH1$–associated PD.

Finally, more work is needed to understand the biological relevance of the GWAS signal in the $GCH1$ locus. A functional haplotype at the $GCH1$ genomic locus associated with reduced plasma GTPCH activity and BH$_4$ levels has been previously recognised (Antoniades et al., 2008). Interestingly, the most-associated $GCH1$ SNP (rs11158026) detected in the GWAS is in moderate linkage disequilibrium ($r^2$ 0.457; D’ 0.932) with the SNPs constituting the functional haplotype, suggesting a potential functional basis for the association of this SNP with increased PD risk.
7.2 Genetics of Myoclonus-Dystonia

*Mutations in KCTD17 and TH are Novel Genetic Causes for SGCE-negative Myoclonus-Dystonia.*

Through a combination of linkage analysis and WES, I identified that a dominant missense mutation in *KCTD17* causes autosomal dominant M-D in families without *SGCE*.

The clinical presentation of the *KCTD17*-mutated cases differs from the phenotype usually observed in *SGCE*-related M-D. All cases initially presented with mild myoclonus affecting the upper limbs. Dystonia showed a progressive course, with increasing severity of symptoms and spreading from the cranio-cervical region to other sites.

Like many other dystonia genes (e.g. *ANO3, GNAL, HPCA*), *KCTD17* is most highly expressed in the putamen, confirming an essential role of this brain structure in the pathogenesis of primary dystonias. Furthermore, a system biology analysis indicates that *KCTD17* is part of a striatal gene network significantly enriched for other dystonia-associated genes and genes involved in post-synaptic dopaminergic signalling in striatal cells. Intriguingly, knock-out of the *KCTD17* homologue in *Drosophila* leads to hyper-arousability and lack of sleep due to increased signaling downstream of the dopaminergic receptors (Stavropoulos and Young, 2011, Pfeiffenberger and Allada, 2012), suggesting that *KCTD17* may interact mechanistically with molecules within the dopaminergic signaling pathway in striatal cells. *KCTD17* is known to contribute to the ubiquitin-proteasome machinery, acting as an adaptor for the CUL3-RING-E3 ligase and regulating turnover of substrates (Kasahara *et al.*, 2014). Functional work is warranted to identify the precise substrates of *KCTD17* in striatal neurons and the mechanism whereby *KCTD17* may regulate post-synaptic dopaminergic signalling. Despite its rarity, the discovery of *KCTD17* mutations as a cause of dystonia may help to elucidate pivotal molecular pathways and novel pharmacological targets to improve treatment of dystonia and other movement disorders.

Finally, the identification of family with tyrosine hydroxylase deficiency presenting with predominant myoclonus-dystonia highlight the importance of consider this treatable condition in the differential diagnosis of M-D, as early treatment with L-DOPA...
is crucial for the motor and cognitive outcome of these patients.

**The CACNA1B p.Arg1389His Variant Does Not Cause Myoclonus-Dystonia**

Through sequencing in a large multi-centric cohort of M-D cases and controls, I show that the CACNA1B p.Arg1389His does not a disease-causing mutation for M-D.

The advent of next generation sequencing has led to an extraordinary acceleration in the discovery rate of rare genetic variants, the majority of which are of uncertain clinical significance. Hence, a close scrutiny is necessary before causally linking a candidate variant to a disease. To avoid false assignment of pathogenicity, MacArthur and colleagues have recently proposed guidelines for implicating causality of rare variants in human disease (MacArthur et al., 2014). However, co-segregation of a variant with disease in a single pedigree does not establish with certainty its pathogenic role, especially if other co-segregating coding variants and the possibility of a separate undetected pathogenic variant in linkage disequilibrium cannot be convincingly ruled out (as it was the case in the pedigree where the p.Arg1389His was identified). In addition, a candidate variant responsible for a rare disease should be found at a low frequency in population controls, consistent with the proposed model of inheritance and disease prevalence.

M-D is an exceedingly rare condition. Although CACNA1B, by a biological point of view, could be a very plausible candidate gene for M-D and functional work performed by the authors shows that the mutation might affect the properties of encoded channel, the frequency of the CACNA1B p.Arg1389His in controls is simply too high to support its role in the pathogenesis of M-D.

### 7.3 Genetics of Chorea

**Benign Hereditary Chorea is Not a Genetic Entity**

Surprisingly, a thorough genetic screen of NKX2-1 in the BHC Queen Square cohort revealed that only five probands (~20% of cohort) carried pathogenic mutations in this gene. A search for the recently described chromosomal deletions neighbouring but not encompassing NKX2-1 (Thorwarth et al., 2014), did not identify any further mutation carrier.

Importantly, I did not detect mutations in three cases with a clinical presentation fully
consistent with *NKX2-1*-related disease (i.e. presenting the full picture of brain-lung-thyroid disease), of which two had a clear autosomal dominant family history.

These results clearly indicate the existence of other undetected pathogenic variants, either in the *NKX2-1* non-coding regions (e.g. in the promoter, in intronic regulatory regions or in yet uncharacterised deep-intronic splice-site) or in other closely functionally related genes. In the future, WGS analysis in these cases will be potentially revealing for the identification of deep intronic mutations in the *NKX2-1* locus. Currently, WES analysis of the familial cases is ongoing in order to identify coding variants in novel genes responsible for the phenotype. Identification of such genes may shed light on the biological mechanisms essential for the correct development of striatal neurons, lungs and thyroid tissue.

A last remark regards the use of the term BHC. Importantly, the work presented in this thesis, in agreement with previous publications (Thorwarth *et al*., 2014), demonstrates that ~70% of the cases with a clinical diagnosis of BHC do not carry mutations in this gene. Furthermore, *NKX2-1* mutations have been identified in patients with a range of movement disorders other than chorea, encouraging, overall, to abandon the use of the term BHC to label patients with *NKX2-1* mutations.

*The identification of ADCY5 and PDE10A mutations crucially implicates cAMP metabolism in striatal neurons in the pathogenesis of chorea*

I describe the identification of pathogenic dominant mutations (both de novo and transmitted) in *ADCY5* and *PDE10A* as important novel genetic causes of chorea with onset in childhood.

From a clinical point of view, *ADCY5* mutations appear to be a relatively common cause of chorea with onset in childhood (~9.4% of the cases I studied here). Importantly, while preparing this thesis, several other papers have appeared in the literature confirming the important clinical role of *ADCY5* mutations and further broadening the phenotype associated with *ADCY5* mutations to also include variable combinations of dystonia, myoclonus and paroxysmal dyskinesias (Carapito *et al*., 2015, Chen *et al*., 2015, Chang *et al*., 2016, Friedman *et al*., 2016).

Furthermore, a manuscript detailing the identification of recessive *PDE10A* mutations in patients with childhood-onset chorea has been recently published (Diggle *et al*.,
Therein, the authors describe the identification of two homozygous mutations (p.Tyr107Cys and p.Ala116Pro), located in the GAF-A domain, in two consanguineous pedigrees. The phenotype of these cases was more severe than that of the patients I describe. Patients with recessive PDE10A mutations showed a much earlier age at onset of chorea (< 1 year), severe dysarthria, axial hypotonia, cognitive and language development delay. The authors also showed that both mutations lead to a reduction in PDE10A levels both in patients and in cellular models. Furthermore, a knock-in mouse model had decreased striatal PDE10A. Interestingly, despite a more severe neurological involvement, the MRI of the cases with recessive mutations did not show the same abnormal signal we observed in the cases with dominant mutations.

ADCY5 and PDE10A are almost exclusively expressed in the striatum, the brain region most closely involved in the pathogenesis of chorea and dystonia, and encode the two main enzymes that regulate the synthesis (ADCY5) and degradation (PDE10A) of cAMP in striatal neurons.

Additionally, mutations in GNAL (Fuchs et al., 2013) and GPR88 (Alkufri et al., 2016), coding for G proteins almost exclusively expressed in striatal neurons that couple dopamine receptors to AC5 activation, have been recently linked to dystonia and chorea. Overall, these results strongly suggest that intracellular cAMP signalling in striatal neurons is crucial for normal activity of basal ganglia circuitry, and that disruptions thereof play an important role in the pathophysiology of hyperkinetic movement disorders. Modulation of the striatal cAMP levels may therefore represent a very attractive target for the development of novel treatments targeting striatal post-receptorial dopaminergic dysregulation in a range of hyperkinetic disorders.

Importantly, modulation of PDE10A activity is already a primary target in pharmacological research of basal ganglia disorders, including HD and PD (Chappie et al., 2012). More specifically, a phase II clinical study (the Amaryllis study) of a PDE10A inhibitor is currently ongoing in HD. Importantly, the identification of loss-of-function PDE10A mutations as a cause of chorea suggests that pharmacological inhibition of PDE10A may not be the best option for the treatment of chorea and other hyperkinetic movement disorders. In support of this notion, it has been recently shown that asymptomatic carriers of the HD triplet expansion have reduced striatal levels of PDE10A (Niccolini et al., 2015).
Finally, it will be important to establish whether the observation of EOPD in one of the cases I identified with a *de novo* PDE10A mutation is coincidental or whether individuals with *de novo* PDE10A mutations are also at an increased risk of developing degeneration of nigral neurons. In this regard, recent work has demonstrated that striatal loss of PDE10A expression is associated with PD duration and severity (Niccolini *et al.*, 2015).
References


Beuckmann CT, Sinton CM, Miyamoto N, Ino M, Yanagisawa M. N-type calcium channel alpha1B subunit (Cav2.2) knock-out mice display hyperactivity and vigilance state differences. J Neurosci. 2003;23(17):6793-7.


Bonifati V, Rohe CF, Breedveld GJ, Fabrizio E, De Mari M, Tassorelli C, et al. Early-onset parkinsonism associated with PINK1 mutations: frequency, genotypes, and


Charlesworth G, Angelova PR, Bartolome-Robledo F, Ryten M, Trabzuni D, Stamelou


Clark LN, Wang Y, Karlins E, Saito L, Mejia-Santana H, Harris J, et al. Frequency of


Datla KP, Blunt SB, Dexter DT. Chronic L-DOPA administration is not toxic to the remaining dopaminergic nigrostriatal neurons, but instead may promote their functional recovery, in rats with partial 6-OHDA or FeCl(3) nigrostriatal lesions. Mov Disord. 2001;16(3):424-34.


Donaldson I, Marsden CD, Schneider SA, Bhatia KB. Clinical approach to movement


Gerfen CR. The neostriatal mosaic: multiple levels of compartmental organization. J Neural Transm Suppl. 1992;36:43-59.


Hypomyelination with atrophy of the basal ganglia and cerebellum: further delineation of the phenotype and genotype-phenotype correlation. Brain. 2014;137(Pt 7):1921-30.


Iwata N, Kobayashi K, Sasaoka T, Hidaka H, Nagatsu T. Structure of the mouse


Koroglu C, Baysal L, Cetinkaya M, Karasoy H, Tolun A. DNAJC6 is responsible for juvenile parkinsonism with phenotypic variability. Parkinsonism Relat Disord.


Kramer PL, Heiman GA, Gasser T, Ozelius LJ, de Leon D, Brin MF, et al. The DYT1


Lesage S, Drouet V, Majounie E, Deramecourt V, Jacoupy M, Nicolas A, et al. Loss of


Mead S, Poulter M, Beck J, Webb TE, Campbell TA, Linehan JM, et al. Inherited prion disease with six octapeptide repeat insertional mutation--molecular analysis of


Nair VD, Olanow CW, Sealfon SC. Activation of phosphoinositide 3-kinase by D2 receptor prevents apoptosis in dopaminergic cell lines. Biochem J. 2003;373(Pt 1):25-32.


Ozelius LJ, Hewett JW, Page CE, Bressman SB, Kramer PL, Shalish C, et al. The early-


Peall KJ, Kurian MA. Benign Hereditary Chorea: An Update. Tremor Other Hyperkinet


Sheerin UM, Charlesworth G, Bras J, Guerreiro R, Bhatia K, Foltynie T, et al. Screening...


Vaarmann A, Kovac S, Holmstrom KM, Gandhi S, Abramov AY. Dopamine protects


Verbeek MM, Steenbergen-Spanjers GC, Willemsen MA, Hol FA, Smeitink J, Seeger J, et al. Mutations in the cyclic adenosine monophosphate response element of the tyrosine


Appendix- Sequences of primers used in this thesis

<table>
<thead>
<tr>
<th>Primer Name</th>
<th>Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>ADCY5_exon1_1_FOR</td>
<td>GCGGATCCTCACGCTACTTCT</td>
</tr>
<tr>
<td>Gene</td>
<td>Forward Primer</td>
</tr>
<tr>
<td>---------------</td>
<td>---------------------------------</td>
</tr>
<tr>
<td>ADCY5_exon7</td>
<td>AGGAACCGCAACTGCATAAAG</td>
</tr>
<tr>
<td>ADCY5_exon8</td>
<td>TGGCATTAGGGAGTTTGAGG</td>
</tr>
<tr>
<td>ADCY5_exon9</td>
<td>TAGTCCAAAGTGGGGGCAA</td>
</tr>
<tr>
<td>ADCY5_exon9</td>
<td>CACGTCTTTTGGGAGACTGT</td>
</tr>
<tr>
<td>ADCY5_exon9</td>
<td>GTGCTGAGGTCTGCTTGTCC</td>
</tr>
<tr>
<td>CACNA1B_exon28</td>
<td>TGGCATTAGGGAGTTTGAGG</td>
</tr>
<tr>
<td>CACNA1B_exon28</td>
<td>CAAGACCAGAAACAGTCCACAG</td>
</tr>
<tr>
<td>DDC_exon 10</td>
<td>GAGGTGGTCCAGTCTAGGGTAC</td>
</tr>
<tr>
<td>DDC_exon 10</td>
<td>CAAGACCAGAAACAGTCCACAG</td>
</tr>
<tr>
<td>DDC_exon 11</td>
<td>GGGAGGAGATGTTGTGACAGC</td>
</tr>
<tr>
<td>DDC_exon 12</td>
<td>AAATAGGCGAGCCGGTGAAG</td>
</tr>
<tr>
<td>DDC_exon 13</td>
<td>TCTAAAGTCCCGAAAGACCTCC</td>
</tr>
<tr>
<td>DDC_exon 13</td>
<td>ATTTGTGTGTCAGGTGGACTG</td>
</tr>
<tr>
<td>DDC_exon 14</td>
<td>GCATGGAGGTAAATGAACAACAC</td>
</tr>
<tr>
<td>DDC_exon 14</td>
<td>GCACGTGCTTCTTCTACATGACC</td>
</tr>
<tr>
<td>DDC_exon 15</td>
<td>CTCGAGGAGGACACTTTTGC</td>
</tr>
<tr>
<td>DDC_exon 16</td>
<td>ACGGCTTGCTTGTTGATCC</td>
</tr>
<tr>
<td>DDC_exon 16</td>
<td>AAATTTCTCCTCACAACCTGACTG</td>
</tr>
<tr>
<td>DDC_exon 17</td>
<td>CTCTTGACTCCTCCGCTATT</td>
</tr>
<tr>
<td>DDC_exon 17</td>
<td>TGGCATTAGGGAGTTTGAGG</td>
</tr>
<tr>
<td>DDC_exon 18</td>
<td>CAAATTCTGCAATGCCTG</td>
</tr>
<tr>
<td>DDC_exon 18</td>
<td>GCATGGAGGTAAATGAACAACAC</td>
</tr>
<tr>
<td>DDC_exon 19</td>
<td>GGTACAATGCGCTTCGCTATG</td>
</tr>
<tr>
<td>DDC_exon 20</td>
<td>GCACGTGCTTCTTCTATTGCC</td>
</tr>
<tr>
<td>DDC_exon 21</td>
<td>ACGAGGAAGGTGTTGCTAATTCC</td>
</tr>
<tr>
<td>DDC_exon 22</td>
<td>AAGAAATTCAGCCTCCGAGGGGGGGG</td>
</tr>
<tr>
<td>DDC_exon 23</td>
<td>TTGAGGGAGGTCGACACAG</td>
</tr>
<tr>
<td>DDC_exon 24</td>
<td>ACCACTGTGGATTATTGTGGCC</td>
</tr>
<tr>
<td>GBA_PCR_fragment 1(ex.1-4)</td>
<td>CTTAGTGGGTCCCCCTG</td>
</tr>
<tr>
<td>GBA_PCR_fragment 1(ex.1-4)</td>
<td>AGCAGACCTACCCTACAGT</td>
</tr>
<tr>
<td>GBA_PCR_fragment 2(ex.5-7)</td>
<td>GCCCTCCTTGATATACCTG</td>
</tr>
<tr>
<td>GBA_PCR_fragment 2(ex.5-7)</td>
<td>AGTTTGGGAGCCAGTCTT</td>
</tr>
<tr>
<td>GBA_PCR_fragment 3(ex.8-11)</td>
<td>ACCACTTACAGGAGGAGAAGT</td>
</tr>
<tr>
<td>GBA_PCR_fragment 3(ex.8-11)</td>
<td>TGTGTCGCAAGGTCCAGGATCAG</td>
</tr>
<tr>
<td>GBA_sequencing_exon1</td>
<td>CTGTTGTGCTGTGCTGCTTCT</td>
</tr>
<tr>
<td>GBA_sequencing_exon1</td>
<td>CAGTGCCAGGATTCAGAGAAG</td>
</tr>
<tr>
<td>GBA_sequencing_exon10-11</td>
<td>5'-GGGTCCCGTGGGTGAG</td>
</tr>
<tr>
<td>Exon</td>
<td>Sequence</td>
</tr>
<tr>
<td>-------------</td>
<td>------------------------------------</td>
</tr>
<tr>
<td>GBA_sequencing_exon10-11_REV</td>
<td>5-TGCTGTGCCCTCTTTTAGTCA</td>
</tr>
<tr>
<td>GBA_sequencing_exon2_FOR</td>
<td>5-CCTGCCCAGGAGAGTAGTTG</td>
</tr>
<tr>
<td>GBA_sequencing_exon2_REV</td>
<td>5-CTCTGTGCTACCTCCCCACT</td>
</tr>
<tr>
<td>GBA_sequencing_exon3_FOR</td>
<td>5-CAAGGGGTGAGAATTTTGA</td>
</tr>
<tr>
<td>GBA_sequencing_exon3_REV</td>
<td>5-TATCAGTACCCAGCGGGAAA</td>
</tr>
<tr>
<td>GBA_sequencing_exon4_FOR</td>
<td>5-GACAGAATGGGCAGAGTGAGA</td>
</tr>
<tr>
<td>GBA_sequencing_exon4_REV</td>
<td>5-CAGGAGCCCAAGTTCCC</td>
</tr>
<tr>
<td>GBA_sequencing_exon5_FOR</td>
<td>5-TGTCTGTACAAGCAGACCTACCC</td>
</tr>
<tr>
<td>GBA_sequencing_exon5_REV</td>
<td>5-GCTGAACCGGGATGCACTG</td>
</tr>
<tr>
<td>GBA_sequencing_exon6_FOR</td>
<td>5-GCTAAATGGGAGGCCAGTC</td>
</tr>
<tr>
<td>GBA_sequencing_exon6_REV</td>
<td>5-ACACCCAGCTGTGCTGTC</td>
</tr>
<tr>
<td>GBA_sequencing_exon7_FOR</td>
<td>5-TGGATGCTGAGATTTGAAGGT</td>
</tr>
<tr>
<td>GBA_sequencing_exon7_REV</td>
<td>5-ACATCATGCTGCAATGACCCTACCT</td>
</tr>
<tr>
<td>GBA_sequencing_exon8_FOR</td>
<td>5-TTCTGTGCTAGTCTTTGAGGAAA</td>
</tr>
<tr>
<td>GCH1_exon 1_FOR</td>
<td>CCGGCCATAAAAGAGGAG</td>
</tr>
<tr>
<td>GCH1_exon 1_REV</td>
<td>CGCCAAAAGTGAGGCAACTC</td>
</tr>
<tr>
<td>GCH1_exon 2_FOR</td>
<td>CTTTCTCTCCTCCGTCTCC</td>
</tr>
<tr>
<td>GCH1_exon 2_REV</td>
<td>ACCTGAGATACGCAATGTCAGGC</td>
</tr>
<tr>
<td>GCH1_exon 3_FOR</td>
<td>AGATGTTTTCAAGGTAATACATTGTCG</td>
</tr>
<tr>
<td>GCH1_exon 3_REV</td>
<td>TAGATTTCAAGGTAATACATTGTCG</td>
</tr>
<tr>
<td>GCH1_exon 4_FOR</td>
<td>GTCTTTTTGTGTTTATGAGGAAGGC</td>
</tr>
<tr>
<td>GCH1_exon 4_REV</td>
<td>GTGATGCTGACCTTTATATATTCAGC</td>
</tr>
<tr>
<td>GCH1_exon 5_FOR</td>
<td>CTCTGCTGCCAGTCAGC</td>
</tr>
<tr>
<td>GCH1_exon 5_REV</td>
<td>AGGCTCAGGGAGGAAATCT</td>
</tr>
<tr>
<td>GCH1_exon 6_FOR</td>
<td>CATTATGTGTAAGAAGGGATATTTCG</td>
</tr>
<tr>
<td>GCH1_exon 6_REV</td>
<td>AGTGACAAGGAATAAAGTTCACATC</td>
</tr>
<tr>
<td>KCTD17_exon 1_FOR</td>
<td>AGGCGCGGACTACGCT</td>
</tr>
<tr>
<td>KCTD17_exon 1_REV</td>
<td>CCACGGCAATGGGTACATC</td>
</tr>
<tr>
<td>KCTD17_exon 2_FOR</td>
<td>TCTCCCCTCCTCTCC</td>
</tr>
<tr>
<td>KCTD17_exon 2_REV</td>
<td>TCCTGTTTGTCCAATGGA</td>
</tr>
<tr>
<td>KCTD17_exon 3_FOR</td>
<td>GAGGGAGAAGGAGGAGAAT</td>
</tr>
<tr>
<td>KCTD17_exon 3_REV</td>
<td>TCCCAACCTCCTCTGCTC</td>
</tr>
<tr>
<td>KCTD17_exon 4_FOR</td>
<td>TCTCTTTTGGTATGTTGCG</td>
</tr>
<tr>
<td>KCTD17_exon 4_REV</td>
<td>TGTCAGAGGCTAGGAGGTC</td>
</tr>
<tr>
<td>KCTD17_exon 5_FOR</td>
<td>GAGGTCGTGATGCTACTGTC</td>
</tr>
<tr>
<td>KCTD17_exon 5_REV</td>
<td>AGAGGTGAGGAGGATGGT</td>
</tr>
<tr>
<td>KCTD17_exon 6_FOR</td>
<td>CTTCACCCTTGCTGAGGACCA</td>
</tr>
<tr>
<td>KCTD17_exon 6_REV</td>
<td>AGGCAAGTGGCTGAGCTAAC</td>
</tr>
<tr>
<td>KCTD17_exon 7_FOR</td>
<td>CAGGGTTAGCTCAGCCACTT</td>
</tr>
<tr>
<td>Exon</td>
<td>Sequence</td>
</tr>
<tr>
<td>------</td>
<td>----------</td>
</tr>
<tr>
<td>KCTD17_exon 7_REV</td>
<td>AGGCAGGGTGAGATGAGAT</td>
</tr>
<tr>
<td>KCTD17_exon 8_FOR</td>
<td>TCTGTGCCACACTAACCCTG</td>
</tr>
<tr>
<td>KCTD17_exon 8_REV</td>
<td>TCAAGAGATGAGCACCCTCC</td>
</tr>
<tr>
<td>KCTD17_exon 9_FOR</td>
<td>CACCCGTCATCTCTCCTC</td>
</tr>
<tr>
<td>KCTD17_exon 9_REV</td>
<td>AGGCAGAGTAAGTCACAGC</td>
</tr>
<tr>
<td>LRRK2_exon41_FOR</td>
<td>AGGGGACAAAGTGAGCACAGAG</td>
</tr>
<tr>
<td>LRRK2_exon41_REV</td>
<td>CACATCTGAGGTCAGTGGTTATC</td>
</tr>
<tr>
<td>NKX2-1_exon 1_FOR</td>
<td>GGCAGAAGAGAGGCAGAG</td>
</tr>
<tr>
<td>NKX2-1_exon 1_REV</td>
<td>CGGACAGGTCTTTAGGAG</td>
</tr>
<tr>
<td>NKX2-1_exon 2_FOR</td>
<td>CGGCCGTTAAGGTACACC</td>
</tr>
<tr>
<td>NKX2-1_exon 2_REV</td>
<td>AGTGCCCCTGTTTCTTGACC</td>
</tr>
<tr>
<td>NKX2-1_exon 3_1_FOR</td>
<td>GCTAGGCTGCTGTGGTCA</td>
</tr>
<tr>
<td>NKX2-1_exon 3_1_REV</td>
<td>CGCTGTGCTGTGCCAGTT</td>
</tr>
<tr>
<td>NKX2-1_exon 3_2_FOR</td>
<td>GCTCCAGAACCACCGTACA</td>
</tr>
<tr>
<td>NKX2-1_exon 3_2_REV</td>
<td>GTTGGCCTGGTTTACCAG</td>
</tr>
<tr>
<td>NKX2-1_exon 3_3_FOR</td>
<td>CAACACGCTCAGACAGT</td>
</tr>
<tr>
<td>NKX2-1_exon 3_3_REV</td>
<td>GGTGGAATGCTGCTGTG</td>
</tr>
<tr>
<td>PARK2_exon1_FOR</td>
<td>GAGGCCCTGAGGAGTTAAC</td>
</tr>
<tr>
<td>PARK2_exon1_REV</td>
<td>GCCCCGTCATTGACAGTT</td>
</tr>
<tr>
<td>PARK2_exon10_FOR</td>
<td>TGTTATTTGCCAAATGCAACC</td>
</tr>
<tr>
<td>PARK2_exon10_REV</td>
<td>GGAACCTCCTGTCAGGACCCAG</td>
</tr>
<tr>
<td>PARK2_exon11_FOR</td>
<td>ACAGGGAACATAACACTCGATCC</td>
</tr>
<tr>
<td>PARK2_exon11_REV</td>
<td>CAACACACCGACGTACAG</td>
</tr>
<tr>
<td>PARK2_exon12_FOR</td>
<td>GTTTGGGAATGCTGCTGTTT</td>
</tr>
<tr>
<td>PARK2_exon12_REV</td>
<td>AGGTAATGAAATGAGGTAGACA</td>
</tr>
<tr>
<td>PARK2_exon2_FOR</td>
<td>ATGTTGCTATCAACCATTTAGGG</td>
</tr>
<tr>
<td>PARK2_exon2_REV</td>
<td>AGATTTGCAAGCGACGGCGCATG</td>
</tr>
<tr>
<td>PARK2_exon3_FOR</td>
<td>TCTCGCATTTTACATGTTTGC</td>
</tr>
<tr>
<td>PARK2_exon3_REV</td>
<td>TAAATATGACCCGGTACGG</td>
</tr>
<tr>
<td>PARK2_exon4_FOR</td>
<td>ACAAGCCTTTTTAAGAGTTTCTTGT</td>
</tr>
<tr>
<td>PARK2_exon4_REV</td>
<td>AGGCCAATGTGTAGTACACA</td>
</tr>
<tr>
<td>PARK2_exon5_FOR</td>
<td>TGGAAACATATGCTTAAAGGTACAG</td>
</tr>
<tr>
<td>PARK2_exon5_REV</td>
<td>TTCTGCGCACAACAGTGAAG</td>
</tr>
<tr>
<td>PARK2_exon6_FOR</td>
<td>AGGAGTTGTTTACTGTTGGAACA</td>
</tr>
<tr>
<td>PARK2_exon6_REV</td>
<td>GTCCGAGAGGAGATGAC</td>
</tr>
<tr>
<td>PARK2_exon7_FOR</td>
<td>TGCTTTTCCACACTGACAGTACT</td>
</tr>
<tr>
<td>PARK2_exon7_REV</td>
<td>TCTGTTCGCTTGGTACAGTAC</td>
</tr>
<tr>
<td>PARK2_exon8_FOR</td>
<td>CAACACTGGCAGTGTAGTACAT</td>
</tr>
<tr>
<td>PARK2_exon8_REV</td>
<td>AGAGGCCCACACTGTCAGC</td>
</tr>
<tr>
<td>PARK2_exon9_FOR</td>
<td>TGCACTGCTTTGAAGGCTC</td>
</tr>
<tr>
<td>PARK2_exon9_REV</td>
<td>AATATAATCCAGCCCATGTCG</td>
</tr>
<tr>
<td>PDE10A_exon1_FOR</td>
<td>GTCCCCACACGGCTAGTCAG</td>
</tr>
<tr>
<td>PDE10A_exon1_REV</td>
<td>GTCCCCACACGGCTAGTCAG</td>
</tr>
</tbody>
</table>
PDE10A_exon1_REV
CTCCTGCCCTCCAGAGAA
PDE10A_exon10_FOR
TGCCAAAGCAGAAAAATCAA
PDE10A_exon10_REV
GGTGAATTCCCTAAGCCACCA
PDE10A_exon11_FOR
CCAACCAATTAGAGATAGGACAGC
PDE10A_exon11_REV
TGCAATGTCCATCCCAGAAG
PDE10A_exon12_FOR
CACGCATTAAAGTAAAACAGTGC
PDE10A_exon12_REV
AACTGAATTGGGACGTGGAG
PDE10A_exon13_FOR
TGCAAAGATTGAGGGAGAT
PDE10A_exon13_REV
CCTTCCAGGGCATTTTATGA
PDE10A_exon14_FOR
TCATGTCTGCACATCCAGT
PDE10A_exon14_REV
AAGCAGACAATGGCAACACA
PDE10A_exon15_FOR
TGGCTAAGAAGGCGAGAAAA
PDE10A_exon15_REV
CAAACAAAGATCGTGAGGAG
PDE10A_exon16_FOR
GATTCCAGTCAGGGTTGAG
PDE10A_exon16_REV
CCATGTCATCAATGTGCTCCT
PDE10A_exon17_FOR
GCCATAATGATCTTCTTTCCT
PDE10A_exon17_REV
CCAAAACCGAATTTGTTGTA
PDE10A_exon18_FOR
TGCCAGCAATCTTGTCAAAC
PDE10A_exon18_REV
TTGGGTTAACAACCACAACA
PDE10A_exon19_FOR
TGTCCCTGCTGCTACTCTGG
PDE10A_exon19REV
TGACTCACCACAGTTCTCTGAAA
PDE10A_exon20_FOR
TGTCCTGGCTGTCACTTCTG
PDE10A_exon20_REV
TGACTCACCACAGTTCTCTGAAA
PDE10A_exon21_FOR
CGCCTCAGAAACAGAACAGTC
PDE10A_exon21_REV
GCCCTCAGAAACAGAACAGTC
PDE10A_exon22_FOR
AAACAAAGGAGAAGAGCTCA
PDE10A_exon22_REV
TGGAAAAGTACCCCTTCTGG
PDE10A_exon3_FOR
AAGAGGAAGGCAAAACAGT
PDE10A_exon3_REV
CTTTCACATTGCACCCCTCTC
PDE10A_exon4_FOR
CAATGTGCTATAGAAGCAATGAGAA
PDE10A_exon4_REV
TTTAGGTTTGAATCTCACCCTCCTG
PDE10A_exon5_FOR
GACTTACATCTGCTTCTACC
PDE10A_exon5_REV
ATGCCTAAATGTGACATGT
PDE10A_exon6_FOR
ATGCCTAAATGACACCTTTCC
PDE10A_exon6_REV
TGGAATAGGGAATTTTCTG
PDE10A_exon7_FOR
CAATGTGCTATAGAAGCAATGAGAA
PDE10A_exon7_REV
TTTACCTCTCAACAGAAAAAGGT
PDE10A_exon8_FOR
TCCCTGGTTGAGATGACTCAAGT
PDE10A_exon8_REV
GGCATTTGGGTACCTAACCT
PDE10A_exon9_FOR
CCCAAGTGATTTGTGAGAGC
<table>
<thead>
<tr>
<th>Gene</th>
<th>Forward Primer</th>
<th>Reverse Primer</th>
</tr>
</thead>
<tbody>
<tr>
<td>PDE10A_exon9_REV</td>
<td>TGGTCTATTTCTGCCACAGG</td>
<td></td>
</tr>
<tr>
<td>SPR_exon 1_FOR</td>
<td>ATCCCAAGGGAACCGAC</td>
<td></td>
</tr>
<tr>
<td>SPR_exon 1_REV</td>
<td>TCTAGGCGTCTATCTGACCC</td>
<td></td>
</tr>
<tr>
<td>SPR_exon 2_FOR</td>
<td>GATATCTGGAGAATCTGGAG</td>
<td></td>
</tr>
<tr>
<td>SPR_exon 2_REV</td>
<td>TGTCGGTTTCTCCCTGGAG</td>
<td></td>
</tr>
<tr>
<td>SPR_exon 3_FOR</td>
<td>CAGGGACCTGAAACCTCTG</td>
<td></td>
</tr>
<tr>
<td>SPR_exon 3_REV</td>
<td>TCTATGTGTAAGGAGGACTGC</td>
<td></td>
</tr>
<tr>
<td>TH_exon 1_FOR</td>
<td>TTCCACGGACGAGCCCTTC</td>
<td></td>
</tr>
<tr>
<td>TH_exon 1_REV</td>
<td>TGGCATGGACCTGAGCCTG</td>
<td></td>
</tr>
<tr>
<td>TH_exon 11_12_FOR</td>
<td>CTGCTCTCCCATCCTT</td>
<td></td>
</tr>
<tr>
<td>TH_exon 13_FOR</td>
<td>TGCTACCAAGGACGAC</td>
<td></td>
</tr>
<tr>
<td>TH_exon 13_REV</td>
<td>CTCAAGGGCAGAAGGAAG</td>
<td></td>
</tr>
<tr>
<td>TH_exon 14_FOR</td>
<td>GAGCTCTGACCCATGTAAG</td>
<td></td>
</tr>
<tr>
<td>TH_exon 14_REV</td>
<td>TACCAACAGGAGCACAAC</td>
<td></td>
</tr>
<tr>
<td>TH_exon 2_FOR</td>
<td>CCTTGATTGAAATGTGGTG</td>
<td></td>
</tr>
<tr>
<td>TH_exon 2_REV</td>
<td>GCTGAACCTCCAAAGG</td>
<td></td>
</tr>
<tr>
<td>TH_exon 3_FOR</td>
<td>TGCTCTGTGTCATCAG</td>
<td></td>
</tr>
<tr>
<td>TH_exon 3_REV</td>
<td>AGGAACCTCAGCCCACACAG</td>
<td></td>
</tr>
<tr>
<td>TH_exon 4_FOR</td>
<td>GAACGGGATCTGTGCTG</td>
<td></td>
</tr>
<tr>
<td>TH_exon 4_REV</td>
<td>ACCTGGTCACTGTAGGGTCC</td>
<td></td>
</tr>
<tr>
<td>TH_exon 5_6_FOR</td>
<td>TCAAGGTGAGTGACACAAG</td>
<td></td>
</tr>
<tr>
<td>TH_exon 5_6_REV</td>
<td>TGCAAGAGATGGTCTCCT</td>
<td></td>
</tr>
<tr>
<td>TH_exon 7_FOR</td>
<td>ACCTGTGCCAGTGAAGCTC</td>
<td></td>
</tr>
<tr>
<td>TH_exon 7_REV</td>
<td>TCACACGACCAGATTCACAG</td>
<td></td>
</tr>
<tr>
<td>TH_exon 8_9_10_FOR</td>
<td>AGGCACCCACAGTTCAGGAG</td>
<td></td>
</tr>
<tr>
<td>TH_exon 8_9_10_REV</td>
<td>GAGCAGGAGCACACTTGC</td>
<td></td>
</tr>
</tbody>
</table>