





**Identifikation und Charakterisierung genetischer Risiko-  
faktoren der Parkinson Erkrankung in der isländischen  
Bevölkerung**

**Dissertation**

der Mathematisch-Naturwissenschaftlichen Fakultät

der Eberhard Karls Universität Tübingen

zur Erlangung des Grades eines

Doktors der Naturwissenschaften

(Dr. rer. nat)

vorgelegt von

Johanna Marlen Huttenlocher-Moser geb. Huttenlocher

Stuttgart-Bad Cannstatt

Tübingen

2016



Gedruckt mit Genehmigung der Mathematisch-Naturwissenschaftlichen Fakultät der Eberhard Karls Universität Tübingen.

Tag der mündlichen Qualifikation:

27.06.2016

Dekan:

Prof. Dr. Wolfgang Rosenstiel

1. Berichterstatter

Prof. Dr. Olaf Rieß

2. Berichterstatter

Prof. Dr. Peter Bauer



---

*Meinen Eltern*





## Inhaltsverzeichnis

<b>1</b>	<b>Promotionsthema.....</b>	<b>1</b>
<b>2</b>	<b>Einleitung .....</b>	<b>3</b>
<b>2.1</b>	<b>Genetische Analyse komplexer oder multifaktorieller Krankheiten.....</b>	<b>3</b>
2.1.1	Assoziationsstudien.....	5
<b>2.2</b>	<b>Das Problem der „fehlenden Erblichkeit“.....</b>	<b>7</b>
2.2.1	„Common Disease, Common Variant“ (CDCV) - Hypothese häufige Erkrankungen, häufige Varianten.....	7
2.2.2	„Common Disease, Multiple Rare Variant“ (MRV) - Hypothese häufige Krankheiten, seltene Varianten.....	8
2.2.3	Pleomorphe Risikobereiche (pleomorphic risk loci - PRL).....	8
<b>2.3</b>	<b>Island: Forschung in einer Isolatpopulation .....</b>	<b>9</b>
2.3.1	Rekombination in kleinen Populationen .....	9
2.3.2	DeCODE Genetics und das Gesamt-Genom-Sequenzierungs-Projekt .....	10
<b>2.4</b>	<b>Statistisches Testen .....</b>	<b>17</b>
2.4.1	p-Wert.....	17
2.4.2	Multiple Testen - Bonferroni-Korrektur.....	17
2.4.3	Relatives Risiko und Odds Ratio .....	18
2.4.4	Fisher-Test und logistische Regression .....	18
2.4.5	Meta-Analysen.....	19
<b>2.5</b>	<b>Parkinson Erkrankung .....</b>	<b>20</b>
2.5.1	Definition und Prävalenz des Parkinson-Syndroms .....	20
2.5.2	Klinisches Bild des idiopathischen Parkinson-Syndroms .....	20
2.5.3	Pathologische Veränderungen beim Parkinson-Syndrom .....	21
2.5.4	Theorien der Pathogenese .....	22
<b>2.6</b>	<b>Genetik der Parkinson Erkrankung .....</b>	<b>23</b>
2.6.1	Monogene Formen der Parkinson Erkrankung .....	23

2.6.2	Genetische Risikofaktoren der sporadische Parkinson Erkrankung .....	27
2.7	Ziel der Arbeit.....	29
<b>3</b>	<b>Ergebnisse .....</b>	<b>31</b>
3.1	Publikation 1: „ <i>EIF4G1</i> is neither a strong nor a common risk factor for Parkinson's disease: evidence from large European cohorts” .....	31
3.1.1	Ergebnisse.....	31
3.1.2	Diskussion .....	33
3.2	Publikation 2: „Heterozygote carriers for CNVs in <i>PARK2</i> are at increased risk of Parkinson’s disease” .....	35
3.2.1	Ergebnisse.....	35
3.2.2	Diskussion .....	38
3.3	Studie 3: “Common genetic PD-risk variants do not affect the penetrance of the <i>LRRK2</i> -gene G2019S mutation” .....	42
3.3.1	Ergebnisse.....	42
3.3.2	Diskussion .....	43
<b>4</b>	<b>Diskussion.....</b>	<b>47</b>
4.1	Fehlende Erbllichkeit - <i>EIF4G1</i> und <i>PARK2</i> .....	49
4.2	Diagnostik - <i>EIF4G1</i> und <i>PARK2</i> .....	51
<b>5</b>	<b>Ausblick .....</b>	<b>55</b>
<b>6</b>	<b>Zusammenfassung.....</b>	<b>57</b>
<b>7</b>	<b>Summary .....</b>	<b>59</b>
<b>8</b>	<b>Literatur .....</b>	<b>61</b>
<b>9</b>	<b>Verzeichnisse .....</b>	<b>83</b>
9.1	Abkürzungsverzeichnis.....	83
9.2	Verwendete Anglizismen .....	89

9.3	Abbildungsverzeichnis .....	91
9.4	Tabellenverzeichnis .....	91
10	Lebenslauf.....	93
11	Publikationen .....	95
12	Anhang.....	97
12.1	Bekannte PD Gene.....	97
12.2	Kumulatives Risikomodell.....	104

❖ **Publikationen**

- *EIF4G1* is neither a strong nor a common risk factor for Parkinson's disease:  
evidence from large European cohorts
- Heterozygote carriers for CNVs in *PARK2* are at increased risk of Parkinson's  
disease

❖ **Erklärung zum Anteil an gemeinschaftlichen Veröffentlichungen**

❖ **Danksagungen**



## 1 Promotionsthema

Identifikation und Charakterisierung genetischer Risikofaktoren der Parkinson Erkrankung in der isländischen Bevölkerung

Genetisch komplexe Erkrankungen, wie z.B. die Parkinson Erkrankung (*Parkinson's disease*, PD), zeichnen sich durch die Beteiligung mehrerer Gene und Umweltfaktoren aus. Die genaue Zahl der beteiligten Faktoren ist jedoch unbekannt. Mit der Erforschung der genetischen Grundlage dieser Erkrankungen beschäftigt sich die Genetische Epidemiologie (Bieckeböller & Fische, 2007). Das Hauptziel ist dabei die Identifikation und Bestimmung der Effektstärke von Desoxyribonukleinsäure (*deoxyribonucleic acid*, DNA)-Sequenzvariationen im menschlichen Erbgut, welche einen Einfluss auf die Entstehung und auf den Verlauf einer solchen Erkrankung haben. Klassische Strategien der Genetischen Epidemiologie sind zum einen die Kopplungsanalyse und zum anderen die Assoziationsstudie. Beide Methoden basieren auf der vergleichenden Frequenzbestimmung genetischer Marker in gesunden und erkrankten Individuen innerhalb einer Familie bzw. einer Population. In den letzten Jahren haben diese Methoden zur Identifikation zahlreicher kausaler Varianten und Risikofaktoren beigetragen. Trotzdem bleibt ein großer Teil der Erblichkeit dieser Erkrankungen nach wie vor unaufgeklärt.

Geographisch isolierte Populationen, wie die isländische Bevölkerung, bieten aufgrund größerer genetischer Übereinstimmung, weniger Haplotypvarianten und höherem Kopplungsungleichgewicht (*linkage disequilibrium*, LD) einzigartige Bedingungen, weitere Assoziationen aufzudecken. Varianten mit mittlerer Effektstärke können dort im Vergleich zu Studien in heterogeneren Bevölkerungsgruppen signifikant nachgewiesen werden.

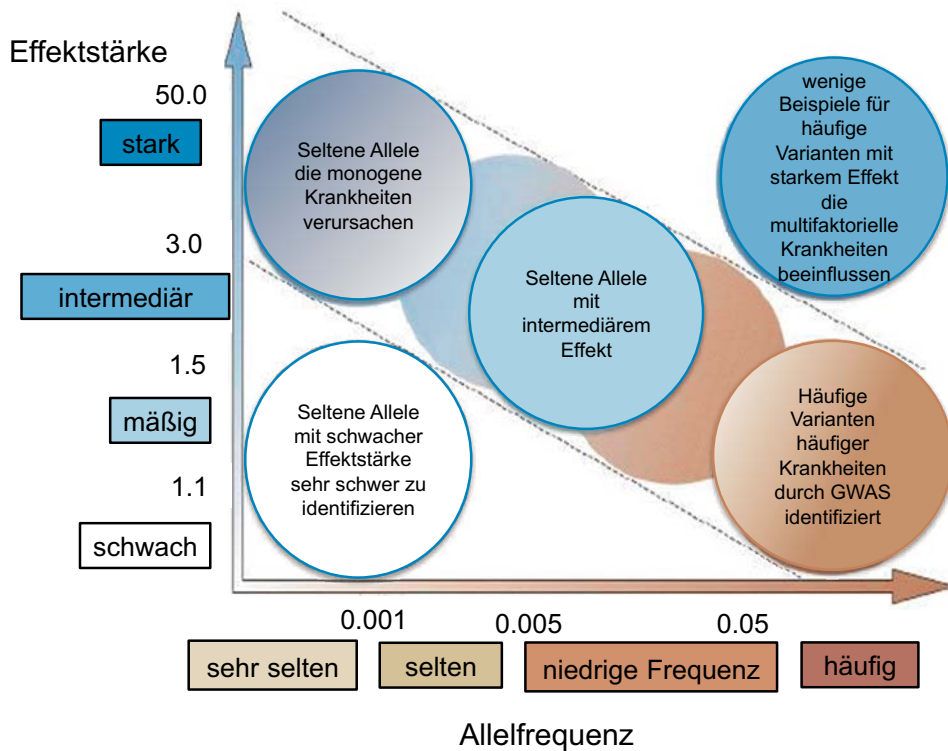
Die Aufschlüsselung der genetischen Architektur komplexer Erkrankungen und damit die Identifizierung und Charakterisierung von Risikovarianten ist von großer Bedeutung für die Prävention und Behandlung von Krankheiten. Darüber hinaus verbessert das Verständnis zugrundeliegender Vererbungsmuster und die Kenntnis von Suszeptibilitätsfaktoren auf Grund ihrer Effektstärke die molekulargenetische Diagnostik.



## **2 Einleitung**

### **2.1 Genetische Analyse komplexer oder multifaktorieller Krankheiten**

Im Focus der Genetischen Epidemiologie steht die Erforschung genetisch multifaktorieller oder komplexer Erkrankungen wie z.B. PD. Im Gegensatz zu klassischen „Mendelschen“ monogenen Krankheiten, die auf die Veränderung an einem Locus zurückzuführen sind, ist an der Entstehung komplexer Erkrankungen eine Vielzahl von Genen beteiligt. Diese Gene können zum einen Hochrisiko-Allele beherbergen, die den mendelschen Vererbungsmustern entsprechend dominant, rezessiv und/oder geschlechtschromosomal vererbt werden, und zum anderen sog. Suszeptibilitätsfaktoren, welche die Entstehung und Entwicklung einer Krankheit beeinflussen, jedoch nicht alleine zur Manifestation führen. Nach wie vor kennt man die genaue Zahl der beteiligten Faktoren sowie deren Wechselwirkungen untereinander und mit der Umwelt nicht. Man geht aber davon aus, dass alle zusammen, d.h. die sog. genetische Architektur (Anzahl, Art, Häufigkeit und Effektgröße von kausalen Varianten) für jede Person das individuelle Ausmaß seiner Empfindlichkeit einer Krankheit gegenüber bestimmt (T. A. Manolio et al., 2009). Abbildung 1 zeigt den Zusammenhang zwischen Allelfrequenz und Effektgröße von Varianten, die durch genetische Studien identifiziert werden können.



**Abbildung 1: Allelfrequenz und Effektstärke identifizierter Risikovarianten**

Seltene Allele mit starkem Effekt werden bei monogenen Krankheiten identifiziert, wohingegen häufige Allele mit schwachem oder mäßigem Effekt durch genomweite Assoziationsstudien (*genome-wide association study*, GWAS) identifiziert werden können. Abb. modifiziert nach (T. A. Manolio et al., 2009).

Methoden zur Analyse komplexer Erkrankungen sind Kopplungsanalysen und Assoziationsstudien. Beide beruhen auf dem Prinzip der gemeinsamen Vererbung benachbarter DNA Marker. Kopplungsanalysen konzentrieren sich dabei auf gemeinsam ererbte Segmente eines Chromosoms, welche bei Erkrankten innerhalb einer Familie gehäuft vorkommen, wohingegen Assoziationsstudien auf der Suche nach über Generationen gemeinsam ererbten Segmenten in der Bevölkerung basieren. In den letzten Jahren haben Kopplungsanalysen bei monogenen Krankheiten bereits große Erfolge erzielt; auch eine Vielzahl der Hochrisiko-Allele monogen vererbter Gene komplexer Krankheiten wurden auf diesem Weg aufgedeckt. Diese Hochrisiko-Allele sind jedoch selten und erklären nur einen kleinen Teil der Prävalenz der Krankheit in der Bevölkerung (Carlson, Eberle, Kruglyak, & Nickerson, 2004; Pritchard, 2001). Die wesentlich häufigeren Suszeptibilitätsfaktoren weisen jedoch keine familiäre Häu-



fung auf und ihr Erbgang entspricht nicht den klassischen „Mendelschen“ Regeln. Daher erfordert ihre Identifikation verfeinerte Analysemethoden wie z.B. Assoziationsstudien.

### **2.1.1 Assoziationsstudien**

Assoziationsstudien basieren auf dem Prinzip der Bestimmung von Häufigkeiten genetischer Marker in Patienten- und Kontrollstichproben aus ethnisch ähnlichen Populationen mit dem Ziel, Hinweise auf genetische Risikofaktoren zu erhalten. Man geht davon aus, dass gleiche Genotypen mit der Erkrankung in der gesamten Bevölkerung assoziiert sind. Diese Assoziation entsteht direkt durch die genetische Risikovariante oder durch Marker Polymorphismen (*tagSNPs*), die mit dem eigentlichen Krankheitsgen im Kopplungsungleichgewicht stehen und dieses somit näherungsweise abbilden (Ardlie, Kruglyak, & Seielstad, 2002; Zondervan & Cardon, 2004). Hat man eine geeignete Anzahl an Probanden zur Verfügung können mit Assoziationsstudien viel schwächere Anfälligkeitsfaktoren nachgewiesen werden als mit Kopplungsanalysen (Hirschhorn & Daly, 2005; Risch & Merikangas, 1996).

#### **2.1.1.1 Kandidatengen – Assoziationsstudie**

Bei einer Kandidatengen-Assoziationsstudie werden sogenannte Kandidatenregionen im Genom, die auf Grund von vorangegangenen Kopplungsanalysen einen Zusammenhang zwischen Region und Krankheit vermuten lassen, weiter eingegrenzt und Kandidatengene direkt auf die verantwortlichen Varianten untersucht. Ziel ist der Nachweis einer signifikanten Assoziation zwischen dem betreffenden Allel und der Erkrankung und dadurch der Nachweis einer Beteiligung des Kandidatengens an der Krankheit (Bieckeböller & Fische, 2007).

#### **2.1.1.2 Genomweite Assoziationsstudie**

Bedingt durch den Technologiesprung in der molekularen Genetik der 90er Jahre (Haines, 2013; S Lesage & Brice, 2012) können heute mit Hilfe von Hochdurchsatz-Genotypisierungs-Verfahren Millionen SNPs kostengünstig und mit geringer Fehlerrate parallel analysiert werden. Die Grundlage hierfür bilden sogenannter SNP-Chips bekannter Biochip-Hersteller wie Illumina und Affymetrix, die basierend auf den öffentlich zugänglichen Daten des SNP-Konsortium (Thorisson & Stein, 2003) und der HapMap (International HapMap, 2005; International & Consortium, 2003), ausgewählte *tagSNPs* über das Genom verteilt, beinhal-

ten. Hierdurch werden Assoziationsstudien auf genomweiter Ebene möglich. Für eine genomweite Analyse ist eine Anzahl von 600.000 bis 3.000.000 SNPs erforderlich, da die gemeinsam vererbten Segmente in einer Bevölkerung sehr klein sind und SNPs in Abständen von 5kb nötig sind, um alle möglichen Chromosomenabschnitte nachzuweisen.

Im Gegensatz zur Kandidatengen-Assoziationsstudie wird hier unabhängig von bekannten Genen (hypothesenfrei) die Lokalisation eines neuen disponierenden Markers im Genom angestrebt (Bieckeböller & Fische, 2007). Um signifikante Ergebnisse zu erhalten, werden allerdings sehr große Patienten- und Kontrollgruppen benötigt (Baker, 2008; Carlson et al., 2004).

## 2.2 Das Problem der „fehlenden Erbllichkeit“

Sowohl bei PD als auch anderen komplexen Erkrankungen erklärt die bislang aus molekular-genetischen Untersuchungen ermittelte genetische Variabilität zusammengenommen nur einen geringen Teil, in der Regel weniger als 10 %, der vorhergesagten Erbllichkeit (Keller et al., 2012). Aus Studien von Familien und eineiigen Zwillingen weiß man jedoch, dass bei den meisten Krankheiten etwa 50 % des Risikos genetischen Faktoren zuzurechnen ist (Collins, 2011). Ein Problem, das als „das Problem der fehlenden Erbllichkeit“ (*missing heritability*) bekannt wurde (T. A. Manolio et al., 2009). Nach heutiger Meinung gibt es verschiedene Erklärungen hierfür: fehlende Information über die Wechselwirkung zwischen Genvarianten, fehlende Information über die Wechselwirkung von Gen und Umwelt (Gibson, 2010; Song, Wang, & Slatkin, 2010) und fehlende Information über Varianten, die von den üblichen Arrays nicht abgedeckt werden, wie z.B. seltene Varianten mit mittlerer Effektstärke, Varianten im nicht kodierenden Bereich des Genoms, die eine Rolle bei der Regulierung der Genexpression spielen (z.B. sind Varianten im Insulin-Genpromotor mit einem erhöhten Risiko für Typ 1 Diabetes assoziiert) oder CNVs. Darüber hinaus werden Verzerrungen (*bias*) der Ergebnisse beim Messen der Eigenschaften als Problem angesehen, sowie ein möglicher Einfluss epigenetischer Faktoren diskutiert (Eichler et al., 2010; Gibson, 2010; T. A. Manolio et al., 2009; Marian, 2012). Die Hauptüberlegungen zur fehlenden Erbllichkeit werden in der „*Common Disease - Common Variant*“ (CDCV-) -Hypothese und der „*Common Disease - Multiple Rare Variant*“ (MRV-) -Hypothese zusammengefasst (Schork, Murray, Frazer, & Topol, 2009).

### 2.2.1 „Common Disease, Common Variant“ (CDCV) - Hypothese häufige

#### Erkrankungen, häufige Varianten

Die CDCV- Hypothese geht davon aus, dass das Zusammentreffen von einer geringen bis moderaten Anzahl von Genvarianten mehrerer Gene, die aber jeweils mit einer relativ hohen Frequenz in einer Population vorkommen (>5 %), unter bestimmten Umweltbedingungen zu häufigen, polygenen Erkrankungen, wie z.B. PD, führen (Bodmer & Bonilla, 2008). Das Erkrankungsrisiko wird durch jede einzelne Variante oder jeden einzelnen Genort nur geringfügig erhöht (durchschnittliche OR 1.36) (Bodmer & Bonilla, 2008). Sie gewinnen jedoch auf Populationsebene an Relevanz, da sie aufgrund ihres geringen Risikos nicht der natürlichen Selektion unterliegen, über viele Generationen weitergegeben werden, und in der Bevölkerung als genetische Varianten gehäuft auftauchen (Hindorff et al., 2009).

### **2.2.2 „Common Disease, Multiple Rare Variant“ (MRV) - Hypothese häufige Krankheiten, seltene Varianten**

Eine Alternativerklärung für fehlende Erblichkeit besagt, dass eine große Zahl von verschiedenen Genen mit jeweils sehr vielen verschiedenen, aber seltenen (privaten) Allelen ( $< 1\%$ ;  $> 5\%$ ), unter bestimmten Umweltbedingungen zu häufigen, polygenen Erkrankungen führen (Azzopardi et al., 2008; Walsh et al., 2008). Die durchschnittliche OR liegt bei 3,74 (Bodmer & Bonilla, 2008). Diese Varianten sind nicht unveränderlich in der Bevölkerung, wodurch der Selektionsdruck die Effektgröße nicht begrenzen kann (Singleton & Hardy, 2011).

Diese beiden Theorien konkurrieren miteinander. Inzwischen weiß man jedoch, dass sie sich nicht gegenseitig ausschließen. Man geht sogar davon aus, dass beide Theorien an einem Ort zusammen treffen können, den sog. pleomorphen Risikobereichen.

### **2.2.3 Pleomorphe Risikobereiche (pleomorphic risk loci - PRL)**

Die Hypothese der PRL besagt, dass bei komplexen Erkrankungen viele einzelne, häufig vorkommenden Risikovarianten und einige private, seltene Varianten am selben Locus koexistieren, miteinander agieren und zur Modifizierung des Krankheitsrisikos beitragen. Das individuelle Risiko eines Locus setzt sich also aus der Allelfrequenz und der Effektgröße dieser verschiedenen Varianten zusammen (Singleton & Hardy, 2011) (siehe Abbildung 1).

Wie bereits beschrieben, können die häufigeren Genvarianten gut über GWAS identifiziert werden. Die Identifizierung von seltenen Varianten in der Bevölkerung mit mittlerem Effekt ist auf diesem Weg jedoch schwierig (Pritchard, 2001). Meist lassen sich die kleinen Effekte erst im Rahmen von Metaanalysen eindeutig absichern. Alternativ müssen Tausende von Personen genotypisiert werden. Dabei ist es eine besondere Schwierigkeit, große phänotypisch einheitliche Kollektive zu generieren (Baker, 2008). Studien in geographisch isolierten Bevölkerungen können hier von Vorteil sein. Darüber hinaus sind neue Methoden wie z.B. *Phasing* und *Imputation* notwendig, um diese Voraussetzungen zu erfüllen.

### **2.3 Island: Forschung in einer Isolatpopulation**

Vor mehr als 1.100 Jahren besiedelten die Skandinavier und Kelten Island (Helgason et al., 2009; Smith, 1995). Seither gab es wenig Immigration. Aufgrund von Naturkatastrophen und Krankheiten wurde die Bevölkerung bis heute immer wieder dezimiert. Darüber hinaus hat der genetische *Drift*, die zufällige langsame Veränderung von Häufigkeiten bestimmter Varianten und dadurch die Elimination von Allelen, die genetische Vielfalt über viele Generationen im Vergleich zu Europa reduziert (Helgason, Nicholson, Stefánsson, & Donnelly, 2003). Die heute rund 320.000 Bewohner der Insel sind daher genetisch eine relativ homogene Bevölkerung, aufgrund dessen ihr Erbgut eine seltene und beeindruckende Quelle für Forscher darstellt. Genetische Unterschiede und Ursachen von Krankheiten können hier leichter identifiziert und erforscht werden. Auch die Tatsache, dass in Island seit 1915 ein staatliches Gesundheitssystem existiert, in welches seither jede Krankheitsgeschichte umfassend eingetragen wurde, kommt der Forschung zugute, genauso wie die Tatsache, dass Ahnenforschung ein Hobby fast jeden Isländers ist. Schon früh nach der Landnahme wurde sehr genau über Verwandtschaftsverhältnisse Buch geführt. So können heute ohne Probleme Familientafeln über viele Jahrhunderte zurückverfolgt werden (Peltonen, Palotie, & Lange, 2000).

#### **2.3.1 Rekombination in kleinen Populationen**

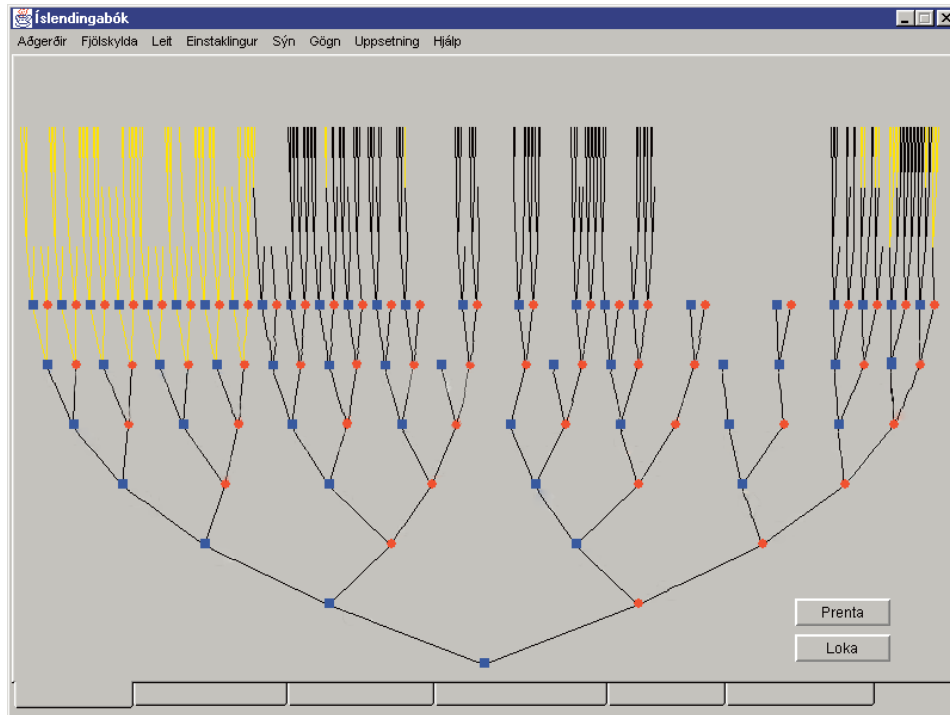
Gene und genetische Polymorphismen liegen linear angeordnet auf den Chromosomen. Daher werden eng zusammen liegende Loci meist auch zusammen von Generation zu Generation vererbt. Sie werden während der Meiose seltener durch ein Rekombinationsereignis getrennt, als durch Zufall in der Population zu erwarten wäre - zwischen ihnen besteht ein Kopplungsungleichgewicht. Auf Grund der Rekombinationshäufigkeit kann das Genom in sog. LD-Blöcke oder auch Haplotypen (HT) eingeteilt werden. Innerhalb eines Blocks findet eine niedrige Rekombinationshäufigkeit statt (*cold spots*), zwischen den Blöcken eine hohe (*hot spots*). Daher bezeichnet man Gene und Allele, die gemeinsam in einem Block liegen, als gekoppelt. Sie werden meist als Ganzes von Eltern auf Kinder vererbt (Strachan & P. Read, 2005). Die Größe und Verteilung von LD-Blöcken unterscheidet sich in verschiedenen Bevölkerungen. Sie sind Bestandteil vieler Studien (Gabriel et al., 2002), deren Ergebnisse im International HapMap Projekt zusammengeführt werden (<http://hapmap.ncbi.nlm.nih.gov/>).

Innerhalb eines Haplotyp-Blocks reicht also die Typisierung weniger informativer Marker um häufige Haplotypen zu identifizieren. Vor diesem Hintergrund werden die Proben für Assoziationsstudien mit SNPChips analysiert.

Kleine, isolierte „junge“ Populationen, die auf eine kleine Gruppe von Gründern zurückgehen (Gründereffekt), weisen in der Regel eine geringere Diversität der Haplotypen auf und stärker ausgeprägte Kopplungsungleichgewichte, was sie für Assoziationsstudien sehr geeignet macht (Kristiansson, Naukkarinen, & Peltonen, 2008). Darüber hinaus lebt der letzte gemeinsame Vorfahre wahrscheinlich in einer jüngeren Vergangenheit als in einer größeren Bevölkerung, und entsprechend größer sind dann auch die gemeinsamen Chromosomenabschnitte, die sich dann wiederum mit weniger Markern nachweisen lassen (Service et al., 2006). Man geht davon aus, dass kleine Völker wie Island bis zu 30 % weniger Marker benötigen, um das Genom abzudecken. Diese Kenntnisse bilden die Grundlage für deCODE Genetics's Gesamt-Genom-Sequenzierungs-Projekt.

### **2.3.2 DeCODE Genetics und das Gesamt-Genom-Sequenzierungs-Projekt**

Eine homogene Bevölkerung, gut dokumentierte Verwandtschaftsverhältnisse und Krankenakten sind eine seltene, hervorragende Voraussetzung, um den Zusammenhang zwischen Genen und Krankheiten aufzudecken. Mit dem Ziel, diese Ressourcen zu nutzen, um neue Hilfsmittel für die Diagnose, die Behandlung und Prävention von Krankheiten zu erstellen, gründete Kári Stefánsson 1996 das Pharmaunternehmen deCODE Genetics (<http://www.decode.com/>). Durch die Zusammenarbeit mit Friðrik Skúlason, einem Software-Entwickler, wurde eine riesige Datenbank ins Leben gerufen - *Íslendingabók*, das Buch der Isländer (<https://www.islendingabok.is/>). Hierbei handelt es sich um eine Stammbaum-Datenbank, bestehend aus Informationen aus Volkszählungen, Kirchenregistern, alte Schriften, Annalen, Nachrufen und Stammbüchern. Später beschloss das isländische Parlament die Erfassung und Speicherung sämtlicher Patientenakten durch deCODE Genetics (Gulcher & Stefánsson, 2000). Die geplante Krankendatenbank der ganzen Nation konnte nach massiven Protesten zwar nicht umgesetzt werden – allerdings stellten 140.000 Isländer, fast 50 % der Bevölkerung, ihre Blut- und Gewebeproben freiwillig zur Verfügung. Heute enthält diese genealogische Datenbank 819.410 registrierte Individuen. Das sind mehr als die Hälfte der Menschen, die je in Island gelebt haben, davon sind 410.551 um 1900 und danach geboren.



**Abbildung 2:** *Íslendingabók* - Das Busch der Isländer (<https://www.islendingabok.is/>).

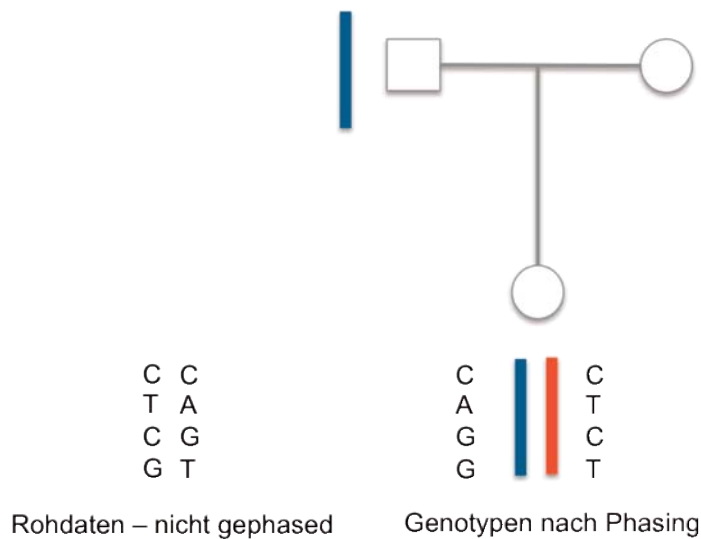
Um die Privatsphäre zu bewahren, werden Namen dreimal verschlüsselt, die DNA-Proben verlassen die Insel nicht, und die Zugangsrechte werden durch einen isländischen Ethik-Ausschuss überwacht.

Das Ziel von deCODE Genetics ist die Sequenzierung des gesamten Genoms der isländischen Bevölkerung. Da dieses Vorhaben sehr zeitaufwendig und teuer ist, wurden in den letzten Jahren verschiedene Verfahren entwickelt, mit deren Hilfe deCODE mit wenigen sequenzierten oder genotypisierten Proben großangelegte, statistisch signifikante Assoziationsstudien durchführen kann. Im Folgenden wird kurz auf diese Verfahren eingegangen.

### 2.3.2.1 Long-Range-Phasing

Entscheidend für die Bestimmung einer Rekombination ist die *Phase*, was heißt, welcher DNA-Strang mütterlicher bzw. väterlicherseits vererbt wurde. Bei einer Genotypisierung werden zwar die Genotypen bestimmt, nicht aber die beiden Haplotypen (Abbildung 3). Viele statistische Methoden nutzen stark korrelierende Allele in LD-Blocks, um ihre Proben zu *phasen* (Myers, Bottolo, Freeman, McVean, & Donnelly, 2005). Allerdings können hiermit

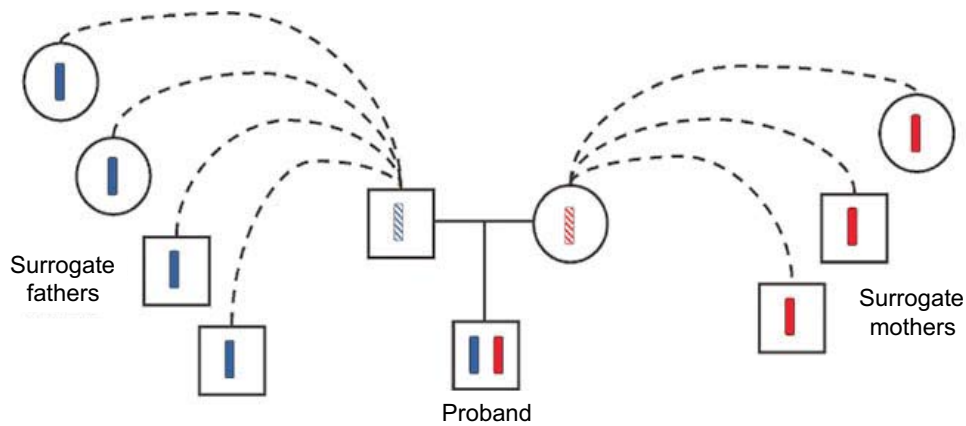
keine SNPs, die über mehrere LD-Blocks getrennt sind, sicher bestimmt werden. Andere verwenden SNP Daten von verwandten Individuen (Trios) (Coop, Wen, Ober, Pritchard, & Przeworski, 2008). In den meisten Fällen sind jedoch die benötigten Voraussetzungen hierfür nicht gegeben.



**Abbildung 3: Phasing**

DeCODE hat die Methode des *Long-range Phasing* (LRP) entwickelt (Kong et al., 2008). Diese ermöglicht die Zuordnung von Haplotypen, die abstammungsidentisch (der Herkunft nach gleich „*identical by descent*“ IBD) sind. IBD-Allele sind nachweisbar Kopien desselben ursprünglichen normalerweise elterlichen Alleles (Strachan & P. Read, 2005). Da Isländer selten blutsverwandte Partnerschaften eingehen, ist der Verwandtschaftskoeffizient, nämlich die Wahrscheinlichkeit, dass ein zufällig gewähltes Allel zwischen zwei Individuen IBD ist, nicht extrem hoch. Er ist aber hoch genug, dass eine beträchtliche Zahl der genotypisierten Individuen erwartungsgemäß abstammungsidentische Haplotypen gemeinsam haben. Für einen Probanden haben im Durchschnitt 17.6 (paternal) und 18.1 (maternal) andere Individuen den Locus IBD geerbt (Kong et al., 2008). Wenn Verwandte abstammungsidentische Haplotypen gemeinsam haben, können sie unabhängig vom Verwandtschaftsgrad zum *phasen* eines entsprechenden Probanden verwendet werden. Der Verwandte agiert als sogenannter „*surrogate father*“ oder „*surrogate mother*“ je nachdem ob er den paternalen oder den maternalen Haplotyp trägt.





**Abbildung 4: Ersatzelternschaft ("surrogate parenthood")**

Abb. modifiziert nach (Kong et al., 2008).

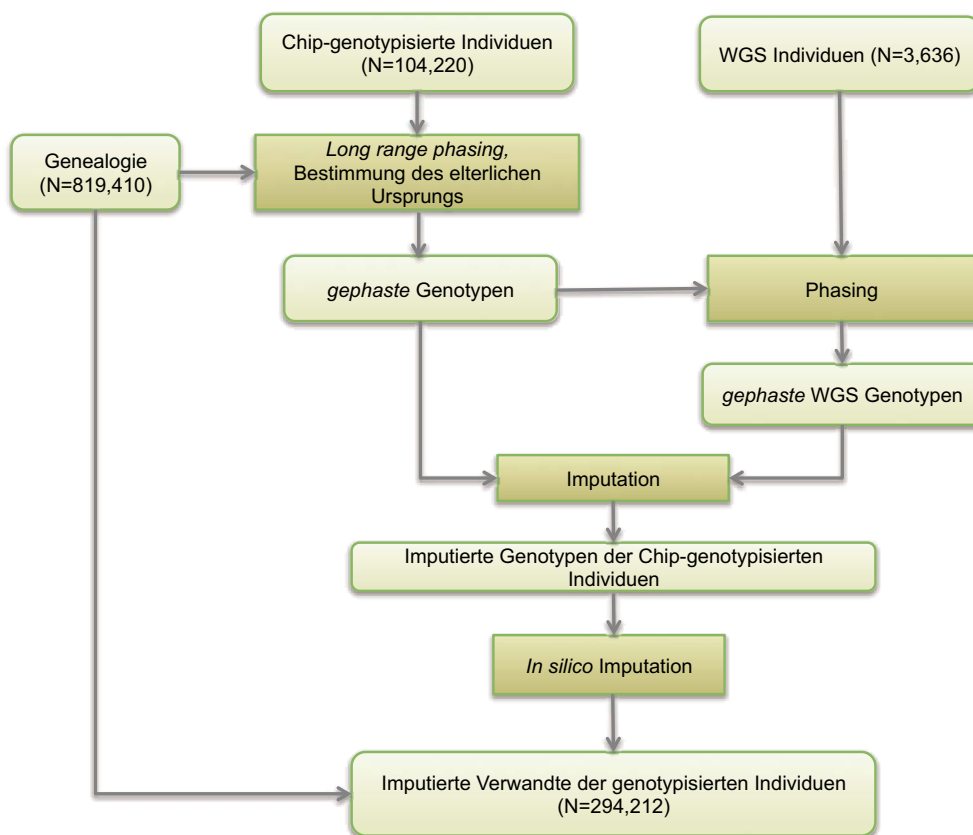
Die „Surrogat Elternschaft“ ist eine ungerichtete Beziehung. D.h. der Proband kann wiederum als „surrogate father“ oder „surrogate mother“ seines „surrogate father“ oder seiner „surrogate mother“ dienen. Durch diese Methode können auch heterozygote SNPs eines Probanden *gephased* werden, sobald ein „surrogate“ Verwandter homozygot ist (Kong et al., 2009).

### 2.3.2.2 Imputation von Genotypen und Haplotypen

Unter Imputation versteht man die Vervollständigung fehlender Daten durch statistische Erhebungen. Die Grundidee besteht darin, dass mit Hilfe von genotypisierten benachbarten Markern, die in hohem Kopplungsungleichgewicht zu den nicht genotypisierten Markern stehen, die Genotypen der nicht genotypisierten Marker geschätzt werden (Ziegler & Inke, 2010). DeCODE hat 2.636 Proben bis zu einer Tiefe von 10-30x mit Illumina Genom Analyser Iix und HiSeq 2000 Geräten komplett genomisch sequenziert. Dabei wurden 34,2 Millionen Sequenzvarianten (30,6 Millionen SNPs und 3,6 Millionen Indels) identifiziert. Diese Varianten werden anschließend in die 104.220 Individuen, die unter der Verwendung von

Illumina SNP Chips (HumanHap300; HumanHap CNV370; Hap 610, 1M oder Omni-1 Quad bead chips) genotypisiert wurden und deren *Phase* durch *long-range Phasing* bekannt ist, imputiert.

In einem zweiten Imputationsschritt werden die wahrscheinlichen Genotypen von 294.212 nahen Verwandten der 104.220 Chip-genotypisierten Individuen vorhergesagt, sogenannte *in silico* Genotypen.



**Abbildung 5: Übersicht des Imputationsvorgangs der Sequenzvarianten.**

Abb. modifiziert nach (Gudbjartsson et al., 2015)

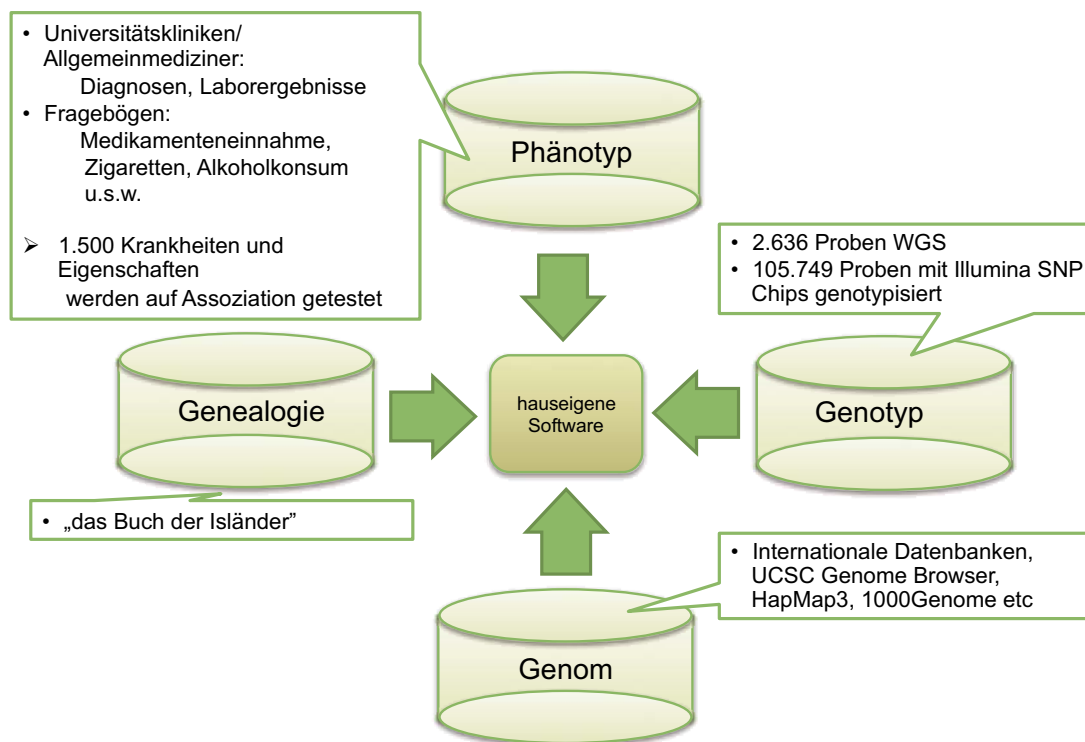
Auf gleichem Weg können auch Haplotypen in Personen mit unbekannter *Phase* imputiert werden wenn zwei genotypisierte Verwandte einen Haplotyp IBD teilen und die Genealogie daraufhin deutet, dass der Weg durch den Probanden führt.

Insgesamt gibt es ungefähr 78.000 Haplotypen, die zurechenbar sind. Diese Zahl bezieht sich auf das gesamte von SNPs gut bedeckte Genom. Welche Probanden und welche Haplotypen imputiert werden können, hängt allerdings von verschiedenen Faktoren ab und variiert von

Locus zu Locus. In ungefähr 10 % der Fälle kann ein Haplotyp trotz genotypisiertem Kind oder Enkel nicht zuverlässig imputiert werden (Kong et al., 2008).

### 2.3.2.3 Assoziationsstudien

Mit Hilfe dieser verschiedenen Ressourcen, die für Mitarbeiter über eine hauseigene Software bereitgestellt werden (Abbildung 6), hat deCODE Genetics bereits viele seltene und häufige genetische Risikovarianten für häufige Erkrankungen aufgedeckt.



**Abbildung 6: Zusammenfassende Darstellung der Ressourcen von deCODE Genetics**

Alle 2–3 Monate wird die Methode der logistischen Regression angewandt, um SNPs und mehr als 1.500 verschiedene Phänotyp-Listen auf Assoziation zu testen. Hierfür werden Genotypen und *in silico* Genotypen verwendet, basierend darauf, wie informativ die kalkulierten Genotypen sind.

In genetischen Fall/Kontroll-Studien können positive Testergebnisse entstehen, ohne dass ein biologischer Zusammenhang zwischen dem untersuchten Marker und der Krankheit existiert, geringe Effekte können verstärkt oder wahre Assoziationen maskiert werden. Wenn dieser

Effekt durch Subpopulationen oder Populationsmischung verursacht wird, spricht man von Populationsstratifikation oder genetischem Confounding (Bieckeböller & Fische, 2007). Zur Kontrolle der Verwandtschaft und Populationsstratifikation innerhalb der isländischen Fall- und Kontrollgruppen, wird daher die Methode der Genomischen Kontrolle angewendet (Devlin & Roeder, 1999).

## 2.4 Statistisches Testen

Da bei Assoziationsstudien immer eine Stichprobe (Teilmenge) analysiert wird und die Ergebnisse auf die Grundgesamtheit, die Bevölkerung, übertragen werden, müssen statistische Methoden angewendet werden. Zur Interpretation der Ergebnisse muss abgeschätzt werden in wie weit der Zufall einen Einfluss auf die Ergebnisse nimmt.

### 2.4.1 p-Wert

Um statistisch zu überprüfen, ob eine Variante mit einer Krankheit assoziiert ist, wird eine Nullhypothese aufgestellt, die besagt, dass die Variante nicht mit der Krankheit assoziiert ist. Der p-Wert, auch Signifikanzwert genannt, gibt die Wahrscheinlichkeit an, mit der das Testergebnis (oder ein noch extremeres Ergebnis) auch dann zu beobachten ist, wenn die Variante nicht mit der Krankheit assoziiert wäre, also die Nullhypothese zutrifft (Foulkes, 2009). Im Voraus wird das Signifikanzniveau festgelegt, d.h. die Grenze, ab der man ein Ergebnis als signifikant bezeichnet. Wenn p kleiner ist als das zuvor festgelegte Signifikanzniveau, wird dies als Beleg interpretiert, dass die Nullhypothese verworfen werden kann und die Alternativhypothese zutrifft. D.h. je kleiner der p-Wert desto mehr spricht das Ergebnis gegen die Nullhypothese. Der p-Wert enthält keine Information über die Größe des gefundenen Effekts, er besagt lediglich, ob ein signifikanter Unterschied existiert (Bieckeböller & Fische, 2007). Bei Biostatistikern hat sich ein Signifikanzniveau von 5 % etabliert. D.h. wird ein Test durchgeführt, beträgt die Wahrscheinlichkeit, dass keine Assoziation trotz hohem Wert vorliegt, 5 % oder 0,05 (Bieckeböller & Fische, 2007). Liegt das Ergebnis höher, ist es nicht signifikant. Die Nullhypothese ist deshalb nicht zwingend richtig, es wurde nur kein Unterschied nachgewiesen. Nach Lander und Kruglyak (1995) entspricht hoch signifikant einer Wahrscheinlichkeit von 0,001 und signifikant von 0,05 (E. Lander & Kruglyak, 1995).

### 2.4.2 Multiples Testen - Bonferroni-Korrektur

Wenn man jetzt allerdings mit mehreren Markern das gesamte Genom testet, ist die Wahrscheinlichkeit, dass man mit einem Marker zu einem falsch positiven Ergebnis gelangt, höher, als bei der Benutzung eines Markers. Mit der Bonferroni-Korrektur wird der p-Wert für das multiple Testen korrigiert. Hierbei wird für n Tests jeweils das Signifikanzniveau durch die Anzahl der Tests geteilt (Bieckeböller & Fische, 2007). Ein P-Wert unter  $5 \times 10^{-8}$  wird allge-

mein als genomweites Signifikanzniveau verwendet (Pe'er, Yelensky, Altshuler, & Daly, 2008).

### **2.4.3 Relatives Risiko und Odds Ratio**

Unter Relativem Risiko (RR) versteht man das Risiko (die Wahrscheinlichkeit), dass eine zufällig ausgewählte Person aus der Bevölkerung, wenn sie einem bestimmten Risikofaktor ausgesetzt ist, innerhalb einer zeitlich begrenzten Periode an einer Krankheit erkrankt, relativ zum Risiko ohne Exposition (d.h. einer nicht exponierten Gruppe) (Bieckeböller & Fische, 2007).

Bei einer Fall-Kontroll-Studie werden in einem Beobachtungszeitraum Erkrankte (Fälle) und Nichterkrankte (Kontrollen) rekrutiert. Hier kann kein RR geschätzt werden, da Fälle und Kontrollen nicht mehr die Anteile aus der Grundgesamtheit repräsentieren. Stattdessen wird die *Odds Ratio* (OR) betrachtet. Das OR vergleicht die Quote des Expositionsfaktors unter den Erkrankten mit der unter den nicht Erkrankten. Sie ist eine Schätzung des Relativen Risikos. Beide Werte beschreiben den Zusammenhang zwischen Genotyp und Phänotyp (Bieckeböller & Fische, 2007; Foulkes, 2009). Aufgrund dieser geschätzten Wirkungsstärke werden genetische Varianten in verschiedene Gruppen eingeteilt: Seltene pathogene Mutationen mit starker Wirkung ( $OR > 50$ ), seltene Risikovarianten mit intermediärer Wirkung ( $OR 1.5-50$ ) und häufig auftretende Risikovarianten mit einem niedrigen Risiko ( $OR < 1,5$ ) (T. A. Manolio et al., 2009).

### **2.4.4 Fisher-Test und logistische Regression**

Bei einem genetischen Marker mit zwei Allelen (z.B. SNP) kann man die OR über eine Mehrfeldertafel zusammenfassen (Foulkes, 2009), die Berechnung erfolgt mittels *Fisher's exact test*. Für die Beschreibung des Zusammenhangs zweier oder mehrerer Variablen dient das Regressionsmodell. Der Phänotyp ist die Zielvariable, alle weiteren genetischen und nicht genetischen Variablen bezeichnet man als Einflussvariablen (Kovariablen) (Bieckeböller & Fische, 2007).

### 2.4.5 Meta-Analysen

In einer Meta-Analyse werden die Ergebnisse mehrerer Studien zur gleichen Fragestellung zusammengefasst. Es handelt sich dabei um eine rein statistische Methode um eine höhere Teststärke der zu bestimmenden Effektgröße zu erreichen. Um einen möglichst aussagekräftigen Gesamteffekt darstellen zu können, ist eine gründliche und umfangreiche Recherche der einzelnen Studien notwendig sowie eine genaue Formulierung der Auswahlkriterien.

Die Ergebnisse, in unserem Fall die OR, der Einzelstudien werden entweder direkt aus den Publikationen entnommen oder aus den Angaben der Publikationen neu berechnet. Hierfür wird meist der *Fisher's exact test* oder ein logistisches oder lineares Regressionsmodell verwendet. Zur anschließenden Berechnung der gepoolten Effektstärke wird entweder ein Modell mit festen Effekten (*fixed effect*) oder ein Modell mit zufälligen Effekten (*random effect*) durchgeführt. Sie unterscheiden sich durch unterschiedliche Annahmen zur statistischen Heterogenität zwischen den Studien (Ziegler & Inke, 2010). Bei Modellen mit festen Effekten wird davon ausgegangen, dass jeder Studie der gleiche Effekt zugrunde liegt, d.h. dass alle Faktoren, die die Effektstärke beeinflussen, in allen Studienpopulationen identisch sind und Unterschiede nur durch zufällige Abweichungen zustande kommen. Die Schätzungen der Gewichte der einzelnen Studien sind umgekehrt proportional zu den einzelnen Studienvarianzen. Bei Modellen mit zufälligen Studieneffekten sind hingegen auch Variabilitäten zwischen den Studien erlaubt. Die Schätzungen der Gewichte der einzelnen Studien umfassen nicht nur die Studienvarianz, sondern auch die Heterogenität zwischen den Studien (Ziegler & Inke, 2010). Bei Meta-Analysen mit wenigen Studien, liefern diese Modelle allerdings keine angemessenen Schätzer (Effektstärken). In diesem Fall wird der Cochran-Mantel-Haenszel Test verwendet (Mantel & Haenszel, 1959). Dieser entspricht wiederum einem Modell mit festen Effekten.

## **2.5 Parkinson Erkrankung**

### **2.5.1 Definition und Prävalenz des Parkinson-Syndroms**

Die Parkinson Erkrankung ist eine langsam fortschreitende, neurodegenerative, Erkrankung des zentralen Nervensystems, welche die Kontrolle der Willkürbewegungen stört. Sie wurde 1817 erstmals durch den Britischen Psychiater James Parkinson in seiner Arbeit „An Essay on the Shaking Palsy“, „Schüttellähmung“, (J. Parkinson, 2002) als klinisch definiertes Krankheitsbild beschrieben. Die Parkinson Erkrankung gehört zu einer heterogenen Familie verschiedener Krankheiten, die überbegrifflich als Parkinson Syndrome oder Parkinsonismus bezeichnet werden. Die eigentliche Parkinsonkrankheit kennt man als idiopathisches Parkinson-Syndrom (IPS) oder primäre Parkinsonkrankheit. Im Gegensatz zu den sekundären Parkinsonkrankheiten oder Parkinson-Syndromen im Rahmen von anderen neurodegenerativen Erkrankungen, deren Ursache man gut kennt, ist diese beim IPS weiter unklar.

IPS kommt in allen ethnischen Gruppen nahezu ohne Geschlechtsunterschiede vor; die Prävalenz steigt nach dem 65. Lebensjahr stark an, so dass ungefähr 0,3 % der Gesamtbevölkerung an einem IPS leiden (3,4 % der über 75-jährigen; 3,6 % der über 85-jährigen (Alves, Forsaa, Pedersen, Dreetz Gjerstad, & Larsen, 2008; Bennett et al., 1996; Coppedè, 2012; de Rijk et al., 1997; A Elbaz et al., 1999)). Damit ist das IPS eine der häufigsten neurologischen Erkrankungen, insbesondere im höheren Alter. In Deutschland geht man von einer Gesamtzahl von 250.000 bis 400.000 Patienten aus.

### **2.5.2 Klinisches Bild des idiopathischen Parkinson-Syndroms**

Die klinischen Symptome des IPS beginnen in der Regel schleichend und unspezifisch wodurch sie oft zu einer jahrelangen Fehldiagnose führen. Nach den derzeit gebräuchlichen Kriterien der englischen „Hirnbank“ (United Kingdom Parkinson’s Disease Society Brain Bank) ist der Parkinsonismus jedoch durch die Verlangsamung der Initiation von Willkürbewegungen mit progressiver Abnahme von Geschwindigkeit und Amplitude bei repetitiven Aktionen (Bradykinese) oder der Reduktion von Bewegungsamplituden und Spontanbewegungen (Hypokinese) und mindestens einem der folgenden Symptome gekennzeichnet: Ruhetremor, Muskelrigor und/oder eine Störung der Halte- bzw. Stellreflexe. Der Ruhetremor, der meist einseitig beginnt, ist das markanteste dieser Kardinalsymptome und führte zum Begriff der „Schüttellähmung“. Im Verlauf der Erkrankung weisen die Patienten eine gebeugte Haltung auf, hinzu kommen das kleinschrittige Gangbild und eine erhöhte Fallneigung (Posturale

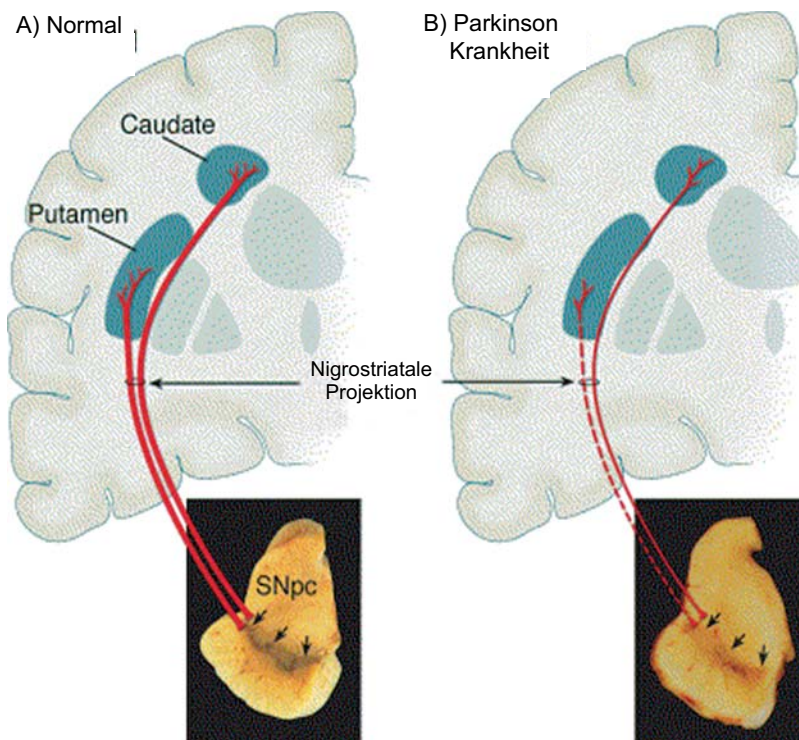


Instabilität). Darüber hinaus spielen auch nicht-motorische Symptome wie z.B. Schluck-, Riech-, Seh- und Schlafstörungen eine Rolle. Viele Patienten klagen im Laufe der Erkrankung über Konzentrations- und Merkschwäche. 40% weisen eine Depression auf (Dooneief et al., 1992) und mehr als 20% eine Demenz (Aarsland, Andersen, Larsen, Lolk, & Kragh-Sørensen, 2003; Biggins et al., 1992). Durch die Therapie bedingt, kommt es zu Halluzinationen meist visueller Natur (Fénelon, Mahieux, Huon, & Ziegler, 2000). Allerdings kann auch die Grunderkrankung zu Halluzinationen führen.

### **2.5.3 Pathologische Veränderungen beim Parkinson-Syndrom**

Das Parkinson-Syndrom (PS) ist neuropathologisch durch den progredienten Verlust der nigrostriatalen dopaminergen Neurone der Substantia nigra pars compacta (SNpc) und dem daraus resultierenden Dopaminmangel im Striatum gekennzeichnet, weiterhin durch die Präsenz von eosinophilen Einschlusskörperchen, genannt Lewy Körperchen, im Zytoplasma der verbleibenden Neuronen (Baba et al., 1998; Riess, Kuhn, & Kruger, 2000; Spillantini et al., 1997; Spillantini, Crowther, Jakes, Hasegawa, & Goedert, 1998). Die Zusammensetzung dieser Einschlusskörperchen und ihre Rolle bei der Pathogenese der Parkinson Erkrankung ist bis heute noch nicht vollständig aufgeklärt. Sie sind Proteinaggregate, die aus zahlreichen Proteinen bestehen.  $\alpha$ -Synuclein stellt die Hauptkomponente der Lewy-Körperchen dar, darüber hinaus wurden Neurofilamente, Ubiquitin, UCH-L1,  $\alpha$ 2-Makroglobulin,  $\alpha$ B-Kristallin,  $\beta$ -Amyloid-Precursor-Protein, Mikrotubuli-assoziiertes Protein Tau, Tyrosin-Hydroxylase, Glutathion-S-Transferase, Topomyosin, Tubulin, Synphilin 1 sowie das Protein 14-3-3 als Bestandteile gefunden (Spillantini et al., 1997; Ubl et al., 2002; Wakabayashi et al., 2000).

Der nigrostriatale Signalweg ist einer der wichtigsten Signalwege im Gehirn. Er besteht aus dopaminergen Neuronen, deren Zellkörper in der SNpc lokalisiert sind. Diese Neuronen senden Projektionen zu den Basalganglien und Synapsen in das Striatum, welches aus Putamen und caudate nucleus bestehen (Abbildung 7). Sie vermitteln Signale, die auf die Kontrolle und Modulation von Willkürbewegungen Einfluss nehmen. Bei Auftritt der ersten Symptome sind bereits etwa 60 % der dopaminergen Neuronen der SNpc degeneriert sowie nahezu 80 % des striatalen Dopamingehalts verloren (Hornykiewicz, 1982).



**Abbildung 7: Schematische Darstellung der nigrostriatalen Projektion**

A) Kontrollhirn mit normaler nigrostriataler Projektion. B) Gehirn eines Parkinson-Patienten mit degenerierter nigrostriataler Projektion. Abb. modifiziert nach (Dauer & Przedborski, 2003).

#### 2.5.4 Theorien der Pathogenese

Wie bereits besprochen, ist die pathologische Veränderung von PD durch den Untergang der dopaminergen Neuronen in der Substantia nigra und die daraus resultierende gestörte Projektion dieser Zellen auf das Striatum bekannt. Die Ursache für diesen Zelluntergang ist allerdings bis heute unklar. Basierend auf der Erforschung von Wirkungsmechanismen dopaminerg Neurotoxine (MPTP) und molekularer Untersuchungen der bereits mit PD assoziierten, monogen vererbten Gene, existieren mehrere Hypothesen bezüglich molekularer Prozesse, die zum Nervenzelluntergang führen. Zu ihnen gehören Alterungsprozesse, oxidativer Stress, mitochondriale Dysfunktion, Apoptose, Protein-Aggregation, Umwelteinflüsse und exogene und endogene Toxine. Die zugrundeliegenden Mechanismen stehen vermutlich miteinander im Zusammenhang. Darüber hinaus haben epidemiologische Untersuchungen verschiedene Risikofaktoren, wie z.B. das Alter, Verhalten (Rauchen, Ernährung, körperliche Aktivität), genetische Prädisposition, oder männliches Geschlecht in Zusammenhang mit der Erkrankung gebracht.

## **2.6 Genetik der Parkinson Erkrankung**

Lange Zeit wurde PD als eine sporadische Krankheit angesehen, die durch verschiedene Umweltfaktoren ausgelöst wird. Die Forschung der letzten Jahre hat jedoch gezeigt, dass nach dem Altern, erbliche Faktoren den stärksten Risikofaktor darstellen (Sellbach, Boyle, Silburn, & Mellick, 2006). Familienangehörige von Parkinson-Patienten haben ein 3-fach erhöhtes Risiko, an PD zu erkranken als die Normalbevölkerung der über 65-jährigen.

### **2.6.1 Monogene Formen der Parkinson Erkrankung**

Seit 1980 haben genetische Kopplungsanalysen in Familien mit erblichem PD zur Identifikation von Krankheitsgenen, und damit entscheidend zur Aufklärung der Pathogenese beigetragen (T Gasser, Hardy, & Mizuno, 2011). Bis heute wurden 15 Genorte als PARK Loci bestimmt (OMIM #168600). 11 dieser Gene können autosomal-dominanten und autosomal-rezessiven Erbgängen zugeordnet werden. Ein Zusammenhang der übrigen Loci ist noch unsicher.

Eine monogenetische oder Mendelsche Vererbung der Parkinson Erkrankung ist relativ selten und variiert zwischen verschiedenen Bevölkerungsgruppen. Zusammengefasst geht man heute davon aus, dass ca. fünf bis max. zehn Prozent der Parkinson-Patienten eine genetisch monogen determinierte Form des IPS haben.

Im Folgenden werden nur die in dieser Arbeit relevanten Gene näher beschrieben. Weitere bekannte PD-Loci sind in Tabelle 1 - 5 im Anhang zusammengefasst.

#### **2.6.1.1 Autosomal dominante Formen von PD**

Bei einer dominanten Vererbung ist eine heterozygote Mutation ausreichend, um die Krankheit auszulösen. In der Regel sind mehrere Familienmitglieder über einige Generationen an PD erkrankt. Ist ein Elternteil heterozygoter Mutationsträger, wird das Krankheitsallel mit 50 % Wahrscheinlichkeit an ein Kind weiter gegeben.

##### **2.6.1.1.1 *PARK 8 Dardarin (LRRK2)***

Mutationen im *Leucin-rich Repeat Kinase 2 (LRRK2)* Gen wurden erstmals bei einer japanischen Familie beschrieben (Zimprich et al., 2004). Das krankheitsverursachende Gen liegt auf

Chromosom 12 (12q12). Es kodiert für das große, 2.527 Aminosäuren umfassende Protein *Dardarin*, welches aus mehreren funktionellen Domänen aufgebaut ist (N-terminale Ankyrin-Domäne (*Ankyrin repeats*), Leucin-reiche Wiederholung von Aminosäuren (*leucin rich repeat*), Roc-Domäne (*Ras of complex proteins*), COR-Domäne (*C-terminal of Roc*), MAPKKK-Domäne (*mitogen-activated protein kinase kinase kinase*) und WD40-Domäne (WD40 repeats) (Taylor, Mata, & Farrer, 2006)). *LRRK2*-Mutationen sind die häufigste Ursache für die dominante familiäre Form der Parkinson Erkrankung mit später Manifestation (5-15 %) (Berg et al., 2005), darüber hinaus sind sie für 1-3 % der sporadischen Fälle verantwortlich (Gilks et al., 2005). Fünf der 19 identifizierten *Missense*-Mutationen sind nachweislich pathogen (p.G2019S, p.R1441C, p.R1441G, p.R1441H, p.Y1699C, p.I2020T), darunter die häufigste Mutation Gly2019Ser (Healy et al., 2008). Ihre Frequenz variiert stark je nach untersuchter Population (Ashkenazi Jewish 30 %, Nord Afrika Berber 40 %, Portugal und Spanien 10 %). In Asien und Europa (ausgenommen Spanien und Portugal) weisen sie eine niedrige Frequenz auf (1-2 %) (Papapetropoulos et al., 2008). Diese deutlichen Häufigkeitsunterschiede sind auf Gründereffekte zurückzuführen (Suzanne Lesage et al., 2005; Ozelius et al., 2006). Zusammengenommen kommt die G2019S-Mutation bei 5-7 % aller Patienten mit autosomal-dominantem Parkinson-Syndrom vor (Di Fonzo et al., 2005) und bei 2-7 % aller sporadischen kaukasischen IPS-Patienten (Thomas Gasser, 2015; Gilks et al., 2005; Mata et al., 2005), was die Bedeutung dieses Proteins verdeutlicht und auch eine diagnostische genetische Untersuchung bei PD-Patienten nahelegt. Die Funktion des Proteins ist unbekannt. Man vermutet, dass es für die Phosphorylierung von  $\alpha$ -Synuclein und/oder Tau verantwortlich ist. Darüber hinaus kann die Kinase-Aktivität eine entscheidende Rolle in der Akkumulation und Aggregation von fehlgefalteten Proteinen in den degenerierenden Neuronen spielen. Die häufige G2019S Mutation befindet sich in der Kinasedomäne und erhöht vermutlich die Kinaseaktivität (Gloeckner et al., 2006; A. B. West et al., 2005).

*LRRK2*-Patienten erkranken typischerweise um das 50. Lebensjahr. Sie sind durch ein großes Spektrum an unterschiedlich ausgeprägten, pathologischen Merkmalen gekennzeichnet und kaum vom Phänotyp des IPS zu unterscheiden (Ishihara et al., 2006). Sie reagieren positiv auf  $L$ -DOPA, und weisen Lewy-Körperchen auf.

### 2.6.1.2 Autosomal rezessive Formen von PD

Bei einer autosomal rezessiven Vererbung verursachen zwei mutierte Allele in einem Gen (entweder homozygot oder zusammengesetzt heterozygot) die Krankheit. Das Auftreten einer

dieser Mutationen hat keine krankmachende Wirkung. Man geht davon aus, dass beide Eltern eines betroffenen Kindes heterozygote Anlageträger am Krankheitsort sind. Im Gegensatz zu den dominanten PD Genen, in denen immer nur ungefähr 1-7 Mutationen bekannt sind, gibt es bei rezessiven Genen viele bekannte Mutationen. Allerdings sind einige von ihnen nicht sicher als pathogen nachgewiesen.

#### 2.6.1.2.1 *PARK 2 Parkin*

Mutationen im *PARK2* Gen wurde erstmals 1998 in einer japanischen Familie beschrieben (Kitada et al., 1998) Das krankheitsverursachende Gen liegt auf dem langen Arm des Chromosoms 6 (6q25) (Kitada et al., 1998). Es ist ein 1,5-Megabasen großes Gen mit 12 Exons (eines der größten Gene des humanen Genoms) und kodiert für ein Protein von 465 Aminosäuren (Kitada et al., 1998), bestehend aus verschiedenen Domänen (N-terminale *Ubiquitin-like* (UBL)-Domäne, C-terminale *Really Interesting New Gene* (RING)-Box, bestehend aus zwei RING-Motiven und der dazwischen liegenden *In-between* RING (IBR)-Domäne (Pfeiffer, Wszolek, & Ebadi, 2012)). *PARK2*-Mutationen sind die häufigste Ursache für die rezessive familiäre Form der Parkinson Erkrankung mit früher Manifestation (50 %) (Abbas et al., 1999). In sporadischen Fällen hängt die Wahrscheinlichkeit, *PARK2*-Mutationen zu finden, stark vom Erkrankungsalter ab (Lucking et al., 2000). Man geht davon aus, dass die meisten sehr jungen Patienten mit IPS einen *PARK2*-Gendefekt aufweisen. Bis heute sind mehr als 100 Mutationen als Ursache für das autosomal-rezessive juvenile Parkinson-Syndrom beschrieben, darunter Deletionen, Duplikationen und Punktmutationen (Abbas et al., 1999). Das Protein, eine E3-Ubiquitin-Ligase, dient dazu, Proteine, die im Proteasom abgebaut werden sollen, mit Ubiquitin zu markieren (Shimura et al., 2000). Eine Störung dieser Funktion durch die Reduktion der Enzymaktivität führt vermutlich zu einer Akkumulation der Proteine, die nicht markiert werden wie z.B.  $\alpha$ -Synuclein oder Synphilin-1, wodurch es zum Zelltod kommt. Der exakte Mechanismus des Parkin assoziierten Zelluntergangs bleibt aber weiter unklar.

*PARK2*-Patienten erkranken schon ab dem 20. Lebensjahr; es handelt sich dabei um eine Form des juvenilen Parkinsonismus. Das klinische Erscheinungsbild ist in der Regel nicht von der sporadischen Parkinson Erkrankung zu unterscheiden und durch ein gutes Ansprechen auf L-DOPA gekennzeichnet. Die Progression der Beschwerden ist sehr langsam. Kognitive Einschränkungen werden nicht beobachtet. Dafür entwickeln die Patienten relativ schnell motorische Fluktuationen. Auffallend ist, dass diese Patienten keine Lewy-Körperchen in der SN

aufweisen, aber doch eine selektive Degeneration in der SN vorhanden ist. Inzwischen wurden aber auch, abhängig von der Position der Mutation im Gen (Lim et al., 2005), Patienten mit Lewy-Körperchen beschrieben (Farrer et al., 2001).

### 2.6.1.3 Loci mit unsicherer PD-Assoziation

#### 2.6.1.3.1 *PARK 18 EIF4G1*

Mutationen im eukaryotischen Initiationsfaktoren 4- $\gamma$  1 (*EIF4G1*)-Gen wurden erstmals bei einer französischen Familie mit autosomal-dominantem PD beschrieben (Chartier-Harlin et al., 2011). Das krankheitsverursachende Gen liegt auf Chromosom 3 (3q27.1) und kodiert für den eukaryotischen Initiationsfaktor 4- $\gamma$  1, ein 1.606 Aminosäuren umfassendes Protein, welches aus mehreren funktionellen Domänen aufgebaut ist (Bindungsstellen für das Poly(A)-binding Protein (PABP), EIF4E,- EIF4A und EIF3 und einer MNK1 (*mitogen-activated protein kinase signal integrating kinase*)-Domäne) (Prévôt, Darlix, & Ohlmann, 2003; Prévôt, Décimo, et al., 2003). Aufgrund der geringen Häufigkeit von Mutationen in diesem Gen, ist eine klare Assoziation mit PD nicht möglich. Auch die beschriebenen Mutationen (p.Ala502Val und p.Arg1205His) wurden inzwischen in Kontrollen gefunden (S Lesage et al., 2012; Nuytemans et al., 2013; Schulte et al., 2012; Tucci et al., 2012).

*EIF4G1* kodiert für einen Transkriptionsfaktor und ist Bestandteil eines als eIF4F bezeichneten Proteinkomplex. Dieser ist für die Translation von wachstumsregulierenden mRNAs verantwortlich. Veränderungen im Gen könnten daher die Formierung dieses Komplexes stören und damit zu einer Translationsrepression führen, welche zu einer verschlechterten Stressantwort der Zelle führt, darüber hinaus zur Herunterregulierung der mitochondrialen Biogenese und einer gesteigerten Autophagie (Cardenas, Cutler, Lorenz, Di Como, & Heitman, 1999; Keiper, Gan, & Rhoads, 1999; Prévôt, Décimo, et al., 2003; Ramírez-Valle, Braunstein, Zavadil, Formenti, & Schneider, 2008).

Die betroffenen Familienmitglieder zeigen die klassische klinische Symptomatik des PD. Das beschriebene durchschnittliche Erkrankungsalter beträgt ca. 64 Jahre. Die Erkrankung weist einen langsamen und milden Verlauf auf und ein gutes Ansprechen auf L-DOPA.

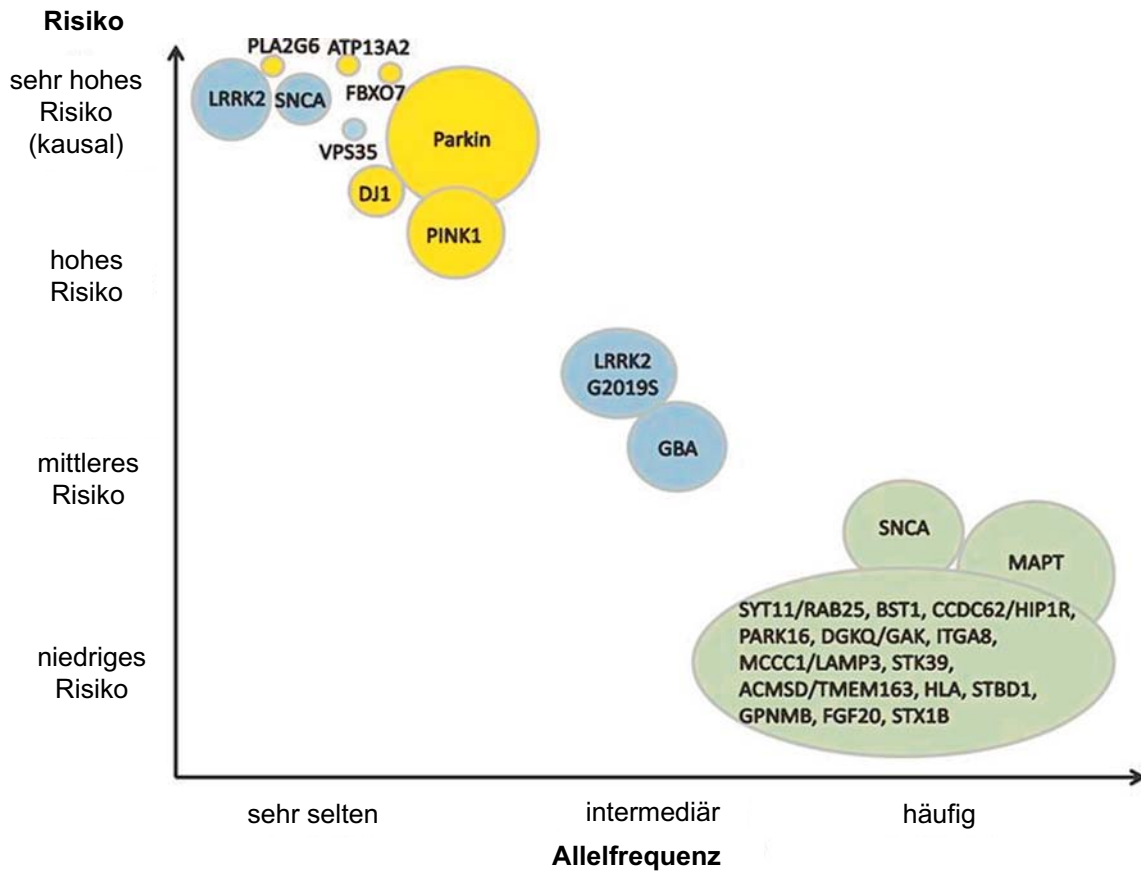
## 2.6.2 Genetische Risikofaktoren der sporadische Parkinson Erkrankung

Genetische Risikofaktoren sind Varianten, die das Risiko einer Erkrankung erhöhen, deren Effektstärke jedoch nicht ausreicht, sie alleine auszulösen. Man vermutet, dass sie für den erblichen Teil der nicht-familiären (sporadischen) PD verantwortlich sind und darüber hinaus die Vielfalt bezüglich des Krankheitsbeginns, Schweregrads und der klinischen Subtypen bei Trägern der gleichen Mutation bewirken. Auch bei atypischen Parkinson-Syndromen wie der Multisystematrophie (MSA) (Scholz et al., 2009) oder der progressiven supranukleären Blickparese (PSP) spielen sie eine Rolle.

Bis heute ließen sich 28 genetische Risiko-Loci für sporadisches PD identifizieren (Keller et al., 2012; M. A. Nalls et al., 2011; M. a Nalls et al., 2014). Interessanterweise haben nur Gene, die bereits mit monogenen Formen von PD in Zusammenhang gebracht werden, SNCA, LRRK2, und MAPT, dauerhaft Assoziation mit PD gezeigt, ein klarer Hinweis auf einen geteilten pathologischen Signalweg zwischen familiärer und sporadischer PD Form (Thomas Gasser, 2015). Die Wahrscheinlichkeit, PD zu entwickeln, wird durch diese Varianten nur um einen Faktor 1.2-2.0 erhöht (Thomas Gasser, 2015). Da sie in der Bevölkerung jedoch häufig vorkommen, können sie auf Populationsebene etwa 10 % des Erkrankungsrisikos erklären.

Zusätzlich zu diesen häufigen Risikofaktoren mit schwacher Effektstärke wurden Risikofaktoren mit intermediärer Effektstärke gefunden (Faktor 5). Im Besonderen heterozygote Mutationen im *GBA*-Gen, das für die Gaucherkrankheit verantwortlich ist, wurden als Suszeptibilitätsfaktoren mit intermediärer Effektstärke validiert (Sidransky et al., 2009). Auch die G2019S Variante kann zu dieser Kategorie gezählt werden.

Die genetische Architektur der Parkinson Erkrankung ist in Abbildung 8 zusammengefasst.



**Abbildung 8: Genetische Architektur der Parkinson Erkrankung**

Zusammenhang der Varianten mit unterschiedlicher Effektstärke und Allelfrequenz. Die Größe der Blasen entspricht ungefähr der Häufigkeit eines Allels in der Bevölkerung. Die Farbe steht für den Vererbungsmodus: dominant (blau), rezessive (gelb), Risikofaktoren (grün). Abb. modifiziert nach (Thomas Gasser, 2015).



## 2.7 Ziel der Arbeit

Durch die Einführung neuer Sequenzierungstechnologien (NGS) haben sich die Möglichkeiten im Bereich der molekulargenetischen Forschung sehr verbessert. Mit der zunehmenden Identifikation von prädisponierenden Genvarianten, die in der Bevölkerung häufig sind, ist auch das Interesse an einer molekularen Diagnostik gestiegen. Daher ist es heute von großer Wichtigkeit, Penetranz und Vererbungsmodus der verschiedenen krankheitsassoziierten Varianten zu kennen, um ein stabiles Fundament für eine gute Beratung zu ermöglichen.

Die vorliegende Arbeit baut auf das Gesamt-Genom-Sequenzierungs-Projekt des isländischen Pharmaunternehmens deCODE Genetics auf. Das Ziel ist die Identifikation und Charakterisierung genetischer Risikofaktoren der Parkinson Erkrankung in der isländischen Bevölkerung. Aufgrund der geografischen und kulturellen Eigenständigkeit dieser Population nimmt man eine geringere genetische Heterogenität an. Folglich eignet sich diese Inselbevölkerung sehr gut zur Erforschung komplexer Merkmale. Drei bekannte PD-assoziierte Gene werden hierbei näher beschrieben.

Im ersten Teil der Arbeit wurde mit Hilfe von konventionellen Genotypisierungsverfahren das *EIF4G1*-Gen in einem großen Europäischen Patientenkollektiv untersucht mit dem Ziel, die Häufigkeit und Relevanz bekannter und neu identifizierter Varianten zu bestimmen. Zum anderen wurde die p.Arg1205His-Mutation im Rahmen einer großen Fall/Kontrollstudie verifiziert und auf Assoziation getestet.

Im zweiten Teil der Arbeit soll die Rolle heterozygoter CNVs im rezessiven *PARK2*-Gen anhand einer großen Fall/Kontrollstudie geklärt werden. Um falsch-positive Mutationsträger auf Grund von *Compound*-Heterozygotie auszuschließen, wurde das isländische Kollektiv zusätzlich auf *Missense*-Mutationen untersucht und alle Patienten mit heterozygoten CNVs auf Mutationen in *trans* überprüft. Anschließend wurde eine Meta-Analyse zur Sicherung unserer Daten durchgeführt.

Zudem wurde in einer dritten Studie der Einfluss häufiger Risikovarianten auf den Phänotyp von Gly2019Ser -Mutationsträgern in einem kumulativen Risikomodell untersucht.



### 3 Ergebnisse

#### 3.1 Publikation 1: „*EIF4G1* is neither a strong nor a common risk factor for Parkinson's disease: evidence from large European cohorts”

**Huttenlocher, J.**, Krüger, R., Capetian, P., Lohmann, K., Brockmann, K., Csoti, I., Klein, C., Berg, D., Gasser, T., Bonin, M., Riess, O., Bauer, P. (2015). *EIF4G1* is neither a strong nor a common risk factor for Parkinson's disease: evidence from large European cohorts. *Journal of Medical Genetics*, 52(1), 37–41. doi:10.1136/jmedgenet-2014-102570

Seit der Erstbeschreibung des PARK18-Locus auf Chromosom 3q26-q28 in einer französischen Familie mit autosomal-dominantem Erbgang und der Identifikation des eukaryotischen Initiationsfaktors 4-γ 1 (*EIF4G1*)-Gens als PD-assoziiertes Gen (Chartier-Harlin et al., 2011), wird dessen Rolle bei der Entstehung der Erkrankung kontrovers diskutiert (S Lesage et al., 2012; Schulte et al., 2012; Tucci et al., 2012). Die anfänglich als schädigend für die Funktion des Proteins beschriebenen *Missense*-Mutationen, eine Arginin zu Histidin Substitution an Position 1205 des Polypeptids (p.Arg1205His) und eine Alanin zu Valin Substitution an Position 502 des Polypeptids (p.Ala502Val) (Chartier-Harlin et al., 2011), wurden in der Zwischenzeit wiederholt auch in gesunden Kontrollen beschrieben (S Lesage et al., 2012; Schulte et al., 2012; Tucci et al., 2012).

Das Ziel unserer Studie war, die Relevanz dieser zwei *EIF4G1*-Mutationen in einer großen europäischen Kohorte sporadischer Patienten zu untersuchen sowie die p.Arg1205His-Variante in der isländischen Bevölkerung auf Assoziation mit PD zu testen. Darüber hinaus wurde die Prävalenz anderer *EIF4G1*-Mutationen in Patienten mit familiengebundener Erkrankung untersucht.

#### 3.1.1 Ergebnisse

##### 3.1.1.1 Genetische Analysen von *EIF4G1*

Im ersten Teil unserer Studie sollte sowohl die Frequenz der vermeintlich pathogenen Mutationen (p.Ala502Val und p.Arg1205His) in Exon 12 und 24 bestimmt, als auch neue kausale

Mutationen in diesen Bereichen detektiert werden. Durch die Untersuchung mittels hochauflösender Schmelzkurvenanalyse und direkter Sanger-Sequenzierung von 2.051 nicht verwandten, sporadischen Patienten europäischen Ursprungs (Lübeck (n = 672, Erkrankungsalter 51,3 ±14,4, 41 % weiblich, davon 69 % mit familiärer PD, von ihnen 23 % mit positiver PD-Diagnose), Tübingen (n = 1.379, Erkrankungsalter 56,1 ±12,0, 41 % weiblich, davon 22,3 % mit familiärer PD)) konnte die p.Arg1205His Variante in einem Patienten heterozygot nachgewiesen werden, wohingegen die p.Ala502Val Variante nicht gefunden wurde. Darüber hinaus konnten wir sechs seltene kodierende SNPs (c.1294A>G; p.Met432Val, c.1456C>T; p.Pro486Ser, c.1384\_1392delGAAGAA; p.Gly465\_Arg467del, c.967C>T; p.Pro323Ser, c.1300C>T; p.Pro434Ser und c.3497G>A; p.Arg1166Gln) und vier bekannte Polymorphismen identifizieren. Bei p.Met432Val (Chartier-Harlin et al., 2011; S Lesage et al., 2012; K. Li et al., 2013; Nuytemans et al., 2013; Tucci et al., 2012), p.Pro486Ser (Nuytemans et al., 2013; Schulte et al., 2012; Tucci et al., 2012) und p.Gly465\_Arg467del (Chartier-Harlin et al., 2011; S Lesage et al., 2012; Nuytemans et al., 2013) handelt es sich um bekannte benigne Varianten, wohingegen p.Pro323Ser, p.Pro434Ser und p.Arg1166Gln noch nicht beschrieben sind. Ob die letzteren die Proteinfunktion beeinträchtigen, muss über funktionelle Analysen erst noch ermittelt werden.

Im zweiten Teil unserer Studie wurde der gesamte kodierende Bereich von 95 Patienten mit familiengebundenem PD aus Deutschland (n = 80) und Island (n = 15) mittels direkter Sanger-Sequenzierung analysiert. In dieser Kohorte wurden zusätzlich drei Varianten und vier bekannte Polymorphismen detektiert. Alle Varianten (c.780C>T; p.Ser260Leu, c.1696C>T; p.Arg566Cys, c.2416A>G; p.Ile807Val) waren bereits beschrieben.

### 3.1.1.2 Assoziation der p.Arg1205His Variante in Island

Während wir die p.Arg1205His Variante in unserer Kohorte nur einmal nachweisen konnten, wurde diese im Zuge des Gesamt-Genom-Sequenzierungs-Projektes bei deCODE Genetics in vier Patienten detektiert. Durch Imputation der Varianten in die gephasen Haplotypen der 93.698 genotypisierten Proben, erhöhte sich die Zahl der Träger auf insgesamt 76, von denen lediglich fünf an PD erkrankt waren. Nach Verifizierung von 200 Proben mit imputierter rs 112176450 Variante mittels direkter Genotypisierung, wurden Patienten und Kontrollen (≤ 65) auf Assoziation mit der p.Arg1205His Variante getestet. Es konnte kein signifikanter Unterschied (OR 1.3, p = 0.67) nachgewiesen werden.

### 3.1.2 Diskussion

Die Identifizierung neuer PD-assoziiierter Gene ist von großer Bedeutung. Sie ermöglicht neue Einblicke in die molekulargenetischen Ursachen und die Entstehung der Erkrankung. Darüber hinaus können Rückschlüsse auf den Phänotyp gezogen werden und infolge dessen neue Angriffspunkte für Medikamente entdeckt werden.

Das *EIF4G1*-Gen wurde 2011 mit einem autosomal-dominanten, spät manifestierenden PD in Zusammenhang gebracht. Besonders die p.Ala502Val und die p.Arg1205His Substitutionen standen im Focus des Interesses. Funktionelle Untersuchungen deuteten darauf hin, dass diese Sequenzveränderungen die Formierung des Translations-Initiations-Komplexes beeinflussen, wodurch die Zelle daran gehindert wird, schnell und effektiv auf Stress zu reagieren. Darüber hinaus wurde gezeigt, dass die mitochondriale Biogenese gehemmt wird und die Autophagie zunimmt (Chartier-Harlin et al., 2011; LeFebvre et al., 2006; Ramirez-Valle, Braunstein, Zavadil, Formenti, & Schneider, 2008; Zhou, Huang, & Przedborski, 2008). Aufgrund dieser Erkenntnisse haben wir uns anfänglich auf die zwei *Hotspot* Exons konzentriert, in denen diese Mutationen lokalisiert sind. Wie bereits in früheren Studien beschrieben, sind beide Mutationen sehr selten (Y. Chen et al., 2013; S Lesage et al., 2012; Nuytemans et al., 2013; Schulte et al., 2012; Siitonen et al., 2013; Sudhama et al., 2013; Tucci et al., 2012; Zhao et al., 2013), übereinstimmend damit konnten wir die p.Ala502Val Variante in unseren Proben (familiär und sporadisch) nicht detektieren, und die p.Arg1205His Variante wurde nur in einem Patienten mit sporadischem PD gefunden. Um die Relevanz dieser seltenen Mutation weiter zu klären, war also ein großer Datensatz erforderlich. Wir führten daher eine Assoziationsstudie mit 2.730 PD-Patienten und 39.046 Kontrollen ( $\leq 65$ ) isländischer Herkunft durch. Unsere Ergebnisse konnten keinen signifikanten Unterschied der beiden Gruppen nachweisen (OR = 1.3,  $P = 0.50$ ). Schwache, frühe PD-Symptome können jedoch in der Kontrollgruppe übersehen worden sein. Trotzdem deuten unsere Ergebnisse nicht auf eine pathogene Wirkung der p.Arg1205His Variante hin. Zumindest scheint sie eine reduziertere Penetranz zu haben als bis jetzt vermutet. Diese Beobachtung wird von den Ergebnissen früherer Studien, die diese Variante auch in Kontrollen gefunden haben, unterstützt (S Lesage et al., 2012; Schulte et al., 2012; Tucci et al., 2012).

Um die genetisch Variabilität des gesamten *EIF4G1*-Gens beurteilen zu können, haben wir zusätzlich nach anderen Mutationen im Gen gesucht. Unter den detektierten Varianten haben wir bei einem Indexpatienten eine heterozygote Arginin zu Cystein Substitution an Position 566 des Polypeptids (p.Arg566Cys) gefunden. Diese Variante wurde mit der Software Poly-

Phen-2 als schädlich für das Protein eingestuft. PolyPhen-2 ist eine bioinformatische Plattform, um den Einfluss eines Aminosäureaustausches auf die Funktion und Struktur eines entstehenden Proteins vorherzusagen. Zur Vorhersage werden die Informationen von Proteinstrukturen sowie Proteinsequenz verwendet. Auch die Position der p.Arg566Cys Variante im Gen nahe der *EIF4E* Bindungs-Stelle ließ uns einen funktionellen Einfluss vermuten. Zwei betroffene Geschwister des Mutationsträgers wurden anschließend sequenziert. Die Variante konnte jedoch nur in einem der beiden Brüder nachgewiesen werden. Um ein mögliches Rekombinationsereignis, welches die Kopplung zu einer anderen relevanten Variante unterbrechen könnte, in der *EIF4G1*-Region ausschließen zu können, wurde zusätzlich eine Mikrosatelliten Haplotypanalyse durchgeführt. Ein Rekombinationsereignis konnte dabei nicht bestätigt werden. Eine Erklärung für dieses Ergebnis könnte eine hohe Toleranz für genetische Variabilität in dieser Region sein, was bereits von Tucci et al nach der Analyse des EVS (<http://evs.gs.washington.edu/EVS/>) Datensatz vermutet wurde (Tucci et al., 2012). Darüber hinaus wurde diese Variante bereits in einer europäisch-amerikanischen und in einer afrikanisch-amerikanischen Kontroll-Probe des EVS (<http://evs.gs.washington.edu/EVS/>) gefunden. Eine eher unwahrscheinliche, aber mögliche pathogene Wirkung muss daher erst noch durch funktionelle Analysen oder Studien in großen Kohorten nachgewiesen werden.

Bei unserer genetischen Analyse des *EIF4G1*-Gens haben wir einige kodierende Varianten gefunden, doch auch diese sprechen nicht dafür, dass Varianten in diesem Gen einen hohen Risikofaktor darstellen. Zu diesem Ergebnis sind auch Nuytemans und seine Kollegen (Nuytemans et al., 2013) gekommen. Sie konnten keine Assoziation des *EIF4G1* Gens mit PD nachweisen.

Unsere Ergebnisse zeigen, dass Mutationen im *EIF4G1*-Gen sowohl in der deutschen als auch in der isländischen Bevölkerung selten sind, was nicht ausschließt, dass das Gen in der Pathogenese von PD eine Rolle spielt. Allerdings handelt es sich vermutlich um einen Risikofaktor mit reduzierter Penetranz.

### 3.2 Publikation 2: „Heterozygote carriers for CNVs in PARK2 are at increased risk of Parkinson’s disease”

**Huttenlocher, J.**, Stefansson, H., Steinberg, S., Helgadóttir, H. T., Sveinbjörnsdóttir, S., Riess, O., & Bauer, P. (2015). Heterozygote carriers for CNVs in PARK2 are at increased risk of Parkinson’s disease. *Human Molecular Genetics*, ddv277–. doi:10.1093/hmg/ddv277

Mutationen im *PARK2*-Gen sind für eine autosomal-rezessiv vererbte Form der Parkinson Erkrankung mit juvenilem Krankheitsbeginn verantwortlich (Kitada et al., 1998). Entsprechend dem klassischen Verständnis eines rezessiven Erbganges ist bekannt, dass homozygote Punktmutationen und Exon-Deletion bzw. Duplikationen in diesem Gen ein Risiko für PD darstellen. Darüber hinaus wird diskutiert, ob auch heterozygote Mutationsträger ein erhöhtes Erkrankungsrisiko haben (Klein, Lohmann-Hedrich, Rogaeva, Schlossmacher, & Lang, 2007). Zur weiteren Klärung dieser Fragestellung haben wir CNVs im *PARK2*-Gen in einem Datensatz von 105.749 genotypisierten Isländern identifiziert und in einer Stichprobe von 1.415 Parkinson-Patienten und 40.474 Kontrollen ( $\leq 65$ ) auf Assoziation mit PD getestet.

#### 3.2.1 Ergebnisse

##### 3.2.1.1 Genetische Analyse von PARK2

###### 3.2.1.1.1 CNVs

CNVs wurden mittels des Computerprogramms *PennCNV-Caller* identifiziert. Dieser Algorithmus nutzt die Signalintensität verschiedener SNPs, um die Kopienzahl einer genomischen Region zu schätzen (K. Wang et al., 2007). Jeder CNV wurde anschließend zusätzlich einer visuellen Kontrolle unterzogen. Unsere Analyse resultierte in 989 CNVs, darunter größere Exon Rearrangements und mehrere kleine Deletionen und Duplikationen, vor allem im Bereich um Exon 2. Bei 24 der CNV-Trägern (1,69 %) handelte es sich um PD-Patienten. Die Häufigkeit heterozygoter CNVs in Kontrollen älter als 65 Jahre (0,91 %) deckte sich mit der Frequenz früherer Studien (D M Kay et al., 2010). Zusätzlich zu CNVs im heterozygoten Mutationsstatus fanden wir in unserem Kollektiv zwei homozygote CNV-Träger (Exon2Del / große Del) in der Patienten-Gruppe, sowie ein Geschwisterpaar älter als 65 Jahre mit einer

homozygoten Exon 2 Duplikation. Sie zeigten zum Zeitpunkt der Untersuchung keine PD-Symptome.

Alle CNV-Träger wurden in einem Kollektiv von 1.415 Patienten und 40.474 Kontrollen älter als 65 Jahre auf Assoziation getestet. Wie erwartet waren homozygote CNVs mit PD assoziiert ( $P = 0.001$ ). Interessanterweise konnten wir aber auch für heterozygote CNV-Träger eine signifikante Risikoerhöhung nachweisen ( $OR = 1.69$ ,  $P = 0.03$ ).

#### 3.2.1.1.2 Ausschluss von Missense-Mutationen bei CNV-Trägern

Um zu klären, ob mögliche *Compound*-heterozygote Mutationsträger diese Assoziation erklären könnten, analysierten wir zusätzlich Daten von 2.261 Proben, die im Zuge des Gesamt-Genom Sequenzierungs-Projekts sequenziert wurden und deren Genotypen in die mittels LRP gephasen Haplotypen imputiert worden waren. Insgesamt konnten wir auf Daten von 1.387 PD-Patienten und 93.698 Kontrollen zurückgreifen. In Übereinstimmung mit früherer Literatur (Bruggemann et al., 2009; D M Kay et al., 2007; Lincoln et al., 2003) fanden wir SNPs mit einer Frequenz von 2,45 % in der isländischen Bevölkerung. Insgesamt wurden acht *Missense*-Mutationen identifiziert. Die p.Met327Ile Variante war noch nicht in der Literatur beschrieben und wurde in Island nur in Kontrollen gefunden. Die anderen sieben, bekannten Varianten wurden sowohl in Patienten als auch in Kontrollen detektiert. Die Varianten p.Ser167Asn, p.Val380Leu und p.Asp394Asn tauchen in der Literatur als häufige, benigne Polymorphismen auf (Lucking et al., 2003), wohingegen p.Ala82Glu, p.Arg402Cys und p.Pro409Leu möglicherweise ein erhöhtes Risiko für PD darstellen (Oliveira et al., 2003). Die Arg275Trp Variante war die einzige, die von der Software PolyPhen-2 als schädigend für das Protein eingestuft wurde. Darüber hinaus wurde sie bereits vielfach in der Literatur beschrieben und dabei entweder als kausale Variante oder Suszeptibilitätsfaktor eingestuft (Abbas et al., 1999; Khan et al., 2002; Oliveira et al., 2003; Ruffmann et al., 2012). In Übereinstimmung damit konnten wir eine Assoziation dieser Variante im homozygoten Status nachweisen ( $OR = 23.1$ ,  $P = 0.079$ ). Eine Assoziation aller *Missense*-Mutationen mit PD konnte jedoch, wie bereits in früheren Studien vermutet (D M Kay et al., 2007), nicht gezeigt werden.

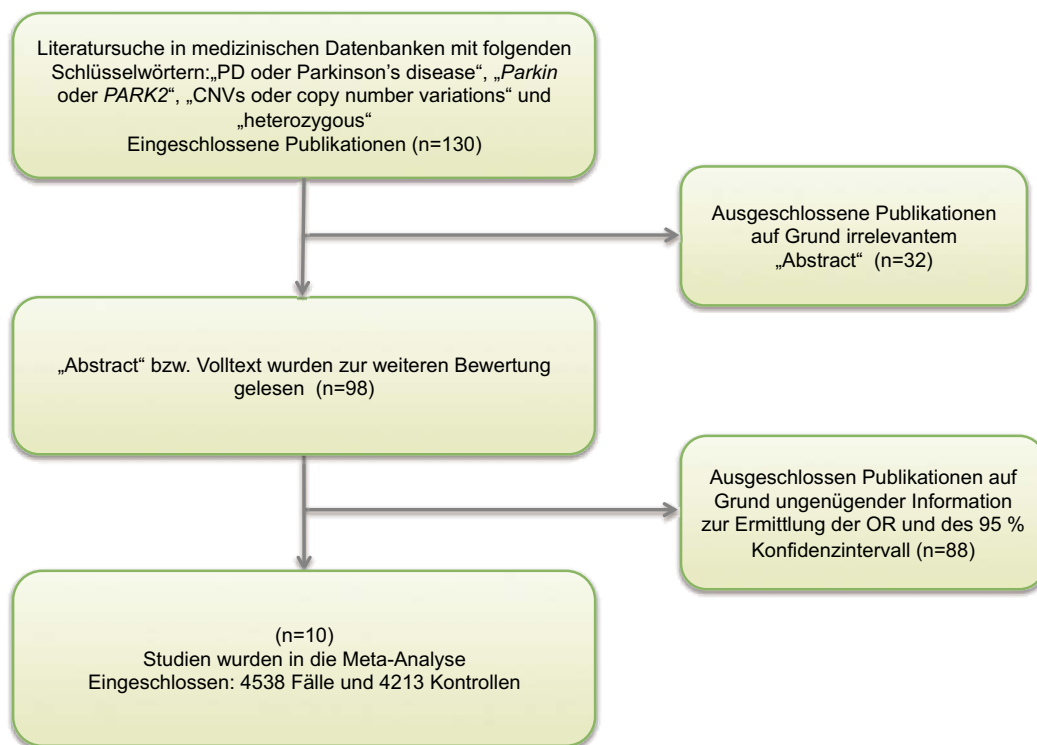
Da durch Imputation nur Mutationen mit einer Frequenz  $> 0,01$  % aufgedeckt werden können, wurden mittels Sanger Sequenzierung mögliche sehr seltene, pathogene Mutationen bei allen Patienten mit CNV oder Arg275Trp Mutation, ausgeschlossen.



Mit Hilfe der gephassten Haplotypen konnte anschließend bestimmt werden, ob sich die veränderten Allele von Trägern zweier Mutationen auf dem gleichen oder verschiedenen Chromosomen befanden. Kein Patient zeigte eine *Compound*-Heterozygotie. Allerdings waren zwei Kontrollen älter als 65 Jahre *compound*-heterozygot für eine Exon2 Duplikation und eine Arg275Trp Variante.

### 3.2.1.2 Meta-Analyse

Für eine Metaanalyse haben wir alle Zeitschriftenartikel der Medline Datenbank (<http://www.ncbi.nlm.nih.gov/pubmed/>) und der PD Mutation Datenbank ([www.molgen.ua.ac.be/PDmutDB](http://www.molgen.ua.ac.be/PDmutDB)) nach Artikeln durchsucht. Die Durchführung und die Ergebnisse der Literatursuche sind in Abbildung 9 dargestellt.



**Abbildung 9: Flussdiagramm zur Literatursuche und der Extraktion relevanter Studienergebnisse**

Die in die Meta-Analyse eingeschlossenen Publikationen beinhalteten ausreichend Information zur Ermittlung der studienspezifischen OR und des 95 % Konfidenzintervall (*confidence*

*interval*, CL) mittels *Fisher's exact test*. Insgesamt wurden Daten von 8.751 Individuen (4.538 Patienten und 4.213 Kontrollen) einbezogen. Für die Durchführung der gepoolten Analyse wurde der Cochran-Mantel-Haenszel Test gewählt. Diese Berechnungen erfolgten mit dem Programm R (<https://www.r-project.org/>). Die isländischen Daten, welche verwandte Individuen beinhalteten, wurden mit einer hauseigenen Software analysiert. Anschließend wurden die publizierten und die isländischen Ergebnisse mittels eines Modells mit festen Effekten und der Inverse-Varianz-Methode zusammengefasst. In der gepoolten Analyse konnten wir eine signifikante Assoziation zwischen heterozygoten CNV Trägern und PD (OR = 2.53,  $P = 1.67 \times 10^{-5}$ ) beobachten, welche in der Kombination mit den Ergebnissen aus Island noch an Signifikanz gewann (OR = 2.11,  $P = 2.54 \times 10^{-6}$ ).

### 3.2.2 Diskussion

Ein rezessiver Erbgang impliziert, dass Mutationen auf beiden Allelen entweder die gleichen Mutationen (homozygot) oder unterschiedliche (*compound*-heterozygot), eine Manifestation der Erkrankung bewirken; wohingegen Mutationen auf nur einer Genkopie keine Auswirkung haben. Da komplexe Krankheiten jedoch fast nie dem klassischen Modell der Mendelschen Segregation folgen und immer wieder heterozygote Mutationsträger mit voll ausgeprägtem Krankheitsspektrum beschrieben (Klein et al., 2007) werden, ist diese Sichtweise Inhalt vieler Diskussionen.

Wir haben uns in unserer Studie auf die Identifikation heterozygoter CNVs im *PARK2*-Gen konzentriert mit dem Ziel, ihre Beteiligung an der Entstehung der Parkinson Erkrankung weiter zu klären.

Die Identifikation von CNVs ist zeitaufwendiger und teurer. Darüber hinaus sind CNVs aufgrund ihrer niedrigen Frequenz in der Bevölkerung (1 %) (D M Kay et al., 2010), nur selten Bestandteil von Assoziationsstudien. Um einen signifikanten Effekt nachzuweisen, werden zur Identifikation sehr große Datensätze benötigt. Auf Grund dieser erschwerten Umstände haben sich bis jetzt die meisten Studien mit einem möglichen Risiko heterozygoter Punktmutationsträger beschäftigt (Clark et al., 2006; D M Kay et al., 2007; S Lesage et al., 2008; Lincoln et al., 2003; Nuytemans et al., 2009), und weniger mit dem von heterozygoten CNV-Trägern, obgleich hinter diesen Sequenzveränderungen eine größere Pathogenität vermutet wird (Pankratz et al., 2009).

In Übereinstimmung mit unseren Ergebnissen haben zwei Studien eine Assoziation heterozygoter CNV-Träger mit PD beschrieben (Pankratz et al., 2009, 2011). Andere Gruppen kamen

jedoch zu gegenteiligen Ergebnissen (D M Kay et al., 2010; Kim et al., 2012; L. Wang et al., 2013). Ursache hierfür könnten die relativ kleinen Kollektive dieser Studien sein, die es erschweren, zuverlässige Vorhersagen des Krankheitsrisikos zu machen.

Allerdings gelangten auch Segregationsanalysen großer Familien zu unterschiedlichen Erkenntnissen. Heterozygote Familienmitglieder von homozygoten Mutationsträgern zeigten leichte PD-Symptome (Abbas et al., 1999; Farrer et al., 2001; Klein et al., 2000; Pramstaller et al., 2002), wohingegen eine andere Studie in heterozygoten Mutationsträgern keine klinischen Symptome feststellen konnte (L. Wang et al., 2013). Auch die Tatsache, dass typische *PARK2*-assoziierte klinische Symptome, wie z.B. Dystonie und Dyskinesie, mit der Zahl der mutierten Allele korreliert (Grünewald, Kasten, Ziegler, & Klein, 2013), spricht für eine Rolle der Mutationen im heterozygoten Zustand. Darüber hinaus wurde ein späteres Erkrankungsalter heterozygoter Mutationsträger im Vergleich zu homozygoten Trägern, die typischerweise vor dem 40. Lebensjahr erkranken, beschrieben (Foroud et al., 2003; Hedrich et al., 2002; Nuytemans et al., 2009; Sun et al., 2006; M. Wang et al., 1999). Jedoch liegt dieses immer noch früher als bei PD-Patienten ohne *PARK2*-Mutationen (Marder et al., 2010; Oliveira et al., 2003; Sun et al., 2006).

Die Vermutung, dass heterozygote Mutationsträger ein erhöhtes Risiko tragen, wird auch durch die Neuro-Bildgebung und elektrophysiologische Befunde gestützt. Vorklinische Veränderungen, wie z.B. eine reduzierte Fluorodopa-Aufnahme in das Striatum, wurde bei symptomlosen heterozygoten Mutationsträgern beschrieben, die zu einem späteren Zeitpunkt an PD erkrankten (Hilker et al., 2001; Hiller et al., 2007; Khan et al., 2002, 2005).

Wie in der Literatur bereits gezeigt wurde, sind *compound*-heterozygote Mutationen im *PARK2*-Gen für einen Teil der an PD erkrankten Mutationsträger verantwortlich (Kim et al., 2012). Es ist sehr schwierig heraus zu finden, ob beide Genkopien mutiert sind, oder die Mutationen vom selben Elternteil vererbt wurden. Hierzu sind Informationen über die Phase der Genotypen erforderlich. In vielen Studien stehen jedoch keine genotypisierten Eltern zur Verfügung. Darüber hinaus handelt es sich in den meisten Fällen um sporadisch Erkrankte. Wir konnten mit Hilfe von gephasen Haplotypen die Struktur der CNVs bestimmen und feststellen, ob die identifizierte *Missense*-Mutation p.Arg275Trp (Klein et al., 2005; Periquet et al., 2001; Ruffmann et al., 2012), möglicherweise in *trans* lokalisiert ist. Da wir in unserem Datensatz keine *compound*-heterozygoten Mutationsträger ausfindig machen konnten, gehen wir davon aus, dass unsere OR von 1.69 allein in heterozygoten Mutationen begründet ist.

Die Durchführung einer Meta-Analyse lieferte zusätzliche Hinweise auf eine Beteiligung heterozygoter *PARK2*-Mutationen an der Manifestation der Krankheit (Abbildung 10). Die

meisten Studien, die sich mit heterozygoten CNV-Trägern im *PARK2*-Gen beschäftigen, sind durch eine geringe Anzahl an Individuen in einer signifikanten Datenerhebung eingeschränkt. Trotzdem resultierte unsere Literatursuche zusammengenommen in 4.538 Patienten und 4.213 Kontrollen. Durch diese Meta-Analyse konnten wir eine Assoziation mit PD bestätigen. Allerdings muss beachtet werden, dass zwischen verschiedenen genetisch-epidemiologischen Studien ethnische Unterschiede oder leichte Abweichungen im Studiendesign, z.B. bzgl. ihrer Ein- und Ausschlusskriterien, existieren. Darüber hinaus haben sich die Analyse-Methoden für die Detektion von CNVs in den letzten Jahren zunehmend verbessert. Einige Wiederholungsstudien sprechen dafür, dass in früheren Studien falsch-positive Proben mit eingeschlossen wurden. Auch Publikationsbias, worunter man eine statistisch verzerrte Darstellung der Datenlage versteht, spielt eine Rolle (Colhoun, McKeigue, & Davey Smith, 2003). Hierbei geht es um das selektive Publizieren bzw. Nichtpublizieren, zugunsten „positiver“ sowie signifikanter Ergebnisse, von ganzen Untersuchungen. Demnach ist davon auszugehen, dass sich weitaus mehr Wissenschaftler mit *PARK2* befasst haben, aufgrund nicht signifikanter Ergebnisse jedoch nicht darüber berichtet wurde. Diese Verzerrungen würden die Ergebnisse der Meta-Analyse stark in Richtung eines größeren Effekts verfälschen. Die hier jedoch relevanteste Einschränkung ist die Tatsache, dass die meisten Studien ihre Probanden nicht nach *Mis-sense*-Mutationen untersucht haben und daher *compound*-heterozygote Mutationsträger nicht ausgeschlossen werden können.

Im Allgemeinen sollten Studienergebnisse die auf einer größeren Anzahl an Versuchspersonen basieren, stärker in dem Gesamtergebnis einer Meta-Analyse berücksichtigt werden. Um diese Präzision der einzelnen Effekte mit einzubeziehen, erhalten die Einzelstudien ein unterschiedliches Gewicht. Der Gesamteffekt unserer Meta-Analyse wird demnach größtenteils durch unsere Studie beeinflusst, was die Bedeutung und Signifikanz unserer Ergebnisse unterstreicht. Zur graphischen Darstellung der gewichteten Einzelergebnisse aller eingeschlossenen Studien mit den Ergebnissen der kombinierten Analysen wurde ein Forest-Plot erstellt (Abbildung 10).

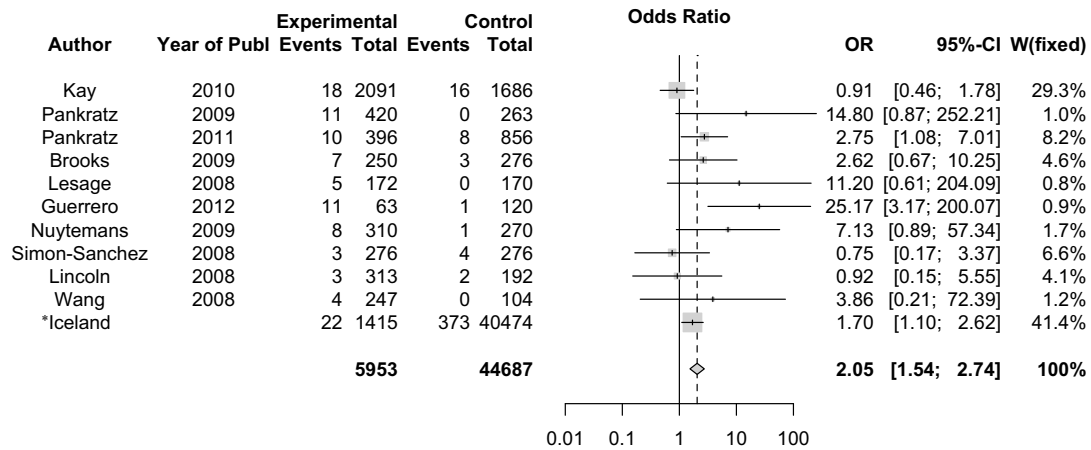


Abbildung 10: Forest-Plot zur Darstellung der Meta-Analyse

\*Iceland: das CI der isländischen Studie wird auf Grund der Nichtberücksichtigung der Verwandtschaftsverhältnisse beim Forest-Plot unterschätzt (*Overdispersion*-Problem).

Zusammenfassend haben wir weitere Hinweise dafür geliefert, dass heterozygote *PARK2* CNV-Träger ein erhöhtes Risiko tragen, an PD zu erkranken. Diese Sequenzveränderungen können als Suszeptibilitätsfaktoren zum Krankheitsbild beitragen und erklären einen weiteren Teil der Erblichkeit sowohl der familiären als auch der sporadische Parkinson Erkrankung. Der Einfluss heterozygoter CNVs sollte nicht unterschätzt werden und weiter im Fokus der Forschung stehen.

### 3.3 Studie 3: “Common genetic PD-risk variants do not affect the penetrance of the *LRRK2*-gene G2019S mutation”

Mutationen im *Leucin-rich Repeat Kinase 2 (LRRK2)*- Gen sind die häufigste, genetische Ursache des sporadischen und des familiären PD. Die *missense*-Veränderung Gly2019Ser (G2019S) ist dabei die häufigste Mutation der dominant erblichen PD. Trotzdem gibt es viele heterozygote Mutationsträger, die keine Krankheitssymptome zeigen, was auf eine reduzierte Penetranz von *LRRK2*-Mutationen hinweist (Healy et al., 2008). Das Phänomen der reduzierten Penetranz wird auf ein Zusammenspiel zusätzlicher genetischer und umweltbedingter Risikofaktoren zurückgeführt (Alexis Elbaz, 2008). Suszeptibilitätsfaktoren, die in der Bevölkerung häufig sind, könnten einen Einfluss auf die pathogene Wirkung der G2019S-Mutation haben. Wir haben die Frequenz der G2019S-Mutation in der isländischen Bevölkerung bestimmt, auf Assoziation zur PD hin getestet und das kumulative Erkrankungsrisiko der Mutationsträger geschätzt, um eine mögliche modifizierende Wirkung von Suszeptibilitätsfaktoren auf die Penetranz der Mutation zu zeigen.

#### 3.3.1 Ergebnisse

##### 3.3.1.1 Genetische Analyse des *LRRK2*

Für unsere Analyse der G2019S-Mutation in der isländischen Bevölkerung konnten wir auf Daten von 1.387 PD-Patienten und 93.698 Kontrollen zurückgreifen. Mittels Gesamt-Genom-Sequenzierung, Haplotyp-Analysen und anschließender Imputation wurden 130 heterozygote G2019S-Mutationsträger identifiziert. Ungefähr 1/3 der über 70-jährigen Mutationsträger waren PD-Patienten. Alle Mutationsträger wurden in einem Kollektiv von 1.415 PD-Patienten und 40.474 Kontrollen  $\geq 65$  Jahre auf Assoziation getestet. Wie erwartet, war die G2019S-Variante signifikant mit PD assoziiert (OR = 2.075,  $P = 0,0062$ ).

### 3.3.1.2 Kumulatives Risikomodell

Im Rahmen eines kumulativen Risikomodells wird angenommen, dass nicht ein isolierter Risikofaktor zu einem negativen Entwicklungsverlauf einer Krankheit beiträgt, sondern besonders die negative Wechselwirkung von mehreren Risiken gleichzeitig (Atzaba-Poria, Pike, & Deater-Deckard, 2004). Um herauszufinden, ob andere Suszeptibilitätsfaktoren, die in der Bevölkerung häufig sind, eine größere pathogene Wirkung der G2019S-Mutation verursachen können, haben wir das Studiendesign eines kumulativen Risikomodells durchgeführt. Als unabhängige Variablen wurden alle kürzlich publizierten, mit PD assoziierten DNA-Marker verwendet (Do et al., 2011; Lill et al., 2012; M. A. Nalls et al., 2011; I. Parkinson, Consortium, Trust, & Control, 2011). In die Analyse wurden nur voneinander unabhängige SNPs ( $r\text{-squared} < 0.2$ ) (<http://www.broadinstitute.org/mpg/snap/ldsearch.php>) einbezogen, und wo möglich wurde die OR einer Replikationsstudie verwendet. Nach Ausschluss der Varianten, die wir in der isländischen Bevölkerung nicht identifizieren konnten, ergab dies eine Liste von 23 Markern (Tabelle 6). Das kumulative Risiko wurde anschließend mit Hilfe eines logistischen Regressionsmodells für folgende Gruppen vergleichend bestimmt: G2019S-Mutationsträger mit PD, G2019S-Mutationsträger ohne PD älter als 70 Jahre, PD-Patienten, familiäre PD-Patienten und sporadische PD-Patienten. Ein signifikanter Unterschied im Vergleich der Gruppen konnte nicht festgestellt werden. Die Ergebnisse sind in Tabelle 7 zusammen gefasst.

### 3.3.2 Diskussion

Ein dominanter Erbgang impliziert, dass Mutationen auf nur einem Allel ausreichen, um zur Manifestation einer Erkrankung zu führen. Heute weiß man jedoch, dass der dominante Erbgang durch verschiedene Einflüsse protektiver genetischer und nicht genetischer Faktoren modifiziert werden kann, was zur unvollständigen Penetranz einer Mutation führt. Eine besondere Form reduzierter Penetranz liegt vor, wenn die Krankheitswahrscheinlichkeit bei vorliegendem Risikotyp vom Alter abhängt. G2019S-Mutationen zeigen ein breites Spektrum des Alters beim Krankheitsbeginn und eine ausgeprägte unvollständige Penetranz (Goldwurm et al., 2007). Obwohl die pathologische Wirkung der Mutation nachvollziehbar ist und die meisten Fälle einer klar dominanten Vererbung zugeordnet werden können (Ishihara et al., 2006; Denise M. Kay et al., 2006; Nichols et al., 2005; Zabetian et al., 2005), gibt es viele Träger, auch im homozygoten Status, die keine klinischen Symptome zeigen (Ishihara et al., 2006).

Darüber hinaus weisen homozygote Mutationsträger im Vergleich mit heterozygoten keine frühere Krankheitsmanifestation oder einen stärker ausgeprägten Phänotyp auf (Klein et al., 2007). Daher vermutet man zusätzliche genetische Faktoren oder Umwelteinflüsse, welche bei der Expression des Phänotyps der G2019S-Träger eine Rolle spielen (Nichols et al., 2005). Auch die Tatsache, dass die G2019S-Mutation mit vielen atypischen PD-Symptomen in Verbindung gebracht wird (Dächsel et al., 2007; Ross et al., 2006; Zhao et al., 2009) spricht für die Beteiligung weiterer Faktoren.

Wir konnten keinen signifikanten Einfluss häufiger Risikovarianten auf den Phänotyp der G2019S-Mutationsträger nachweisen. Allerdings wurde nur etwa die Hälfte der für unser Risikomodell verwendeten SNPs, N370S/i4000416 im *GBA*-Gen (Han et al., 2015), rs11248060 im *DGKQ*-Gen (Yong Ping Chen et al., 2013), rs356219 im *SNCA*-Gen (N.-N. Li, Mao, et al., 2013; Pan et al., 2012), chr8:89442157 im *MMP16*-Gen (Hernandez et al., 2012), rs1491942 im *LRRK2*-Gen (Hernandez et al., 2012; Lill et al., 2012), rs4130047 im *RIT2/SYT4*-Gen, rs2390669 im *STK39*-Gen, chr4:911311/rs6599389 im *GAK*-Gen (M. a Nalls et al., 2014), rs708723 im *RAB7LI*-Gen (Hernandez et al., 2012; Liu et al., 2013), rs156429 im *GPNMB*-Gen (M. a Nalls et al., 2014), rs7077361 im *ITGA8*-Gen und rs10513789 im *MCCCI*-Gen (M. a Nalls et al., 2014), auch in weiteren Studien signifikant repliziert.

Eine Assoziation der Variante rs6710823 im *ACMSD*-Gen mit PD hat sich in Folgestudien nicht bestätigen lassen. Darüberhinaus wurden folgende Varianten nur in manchen Studien signifikant repliziert: rs34016896 im *NMD3*-Gen (Liu et al., 2013), rs11724635 im *BST1*-Gen (K.-H. Chang et al., 2015; Miyake et al., 2012; Sharma et al., 2012; S. Wang et al., 2015), rs823156 im *SLC41A1*-Gen (X. L. Chang et al., 2011; H. Li, Teo, & Tan, 2015), rs12817488 im *CCDC62/HIP1R*-Gen (N.-N. Li, Tan, et al., 2013; Shulman et al., 2014; Yu et al., 2015) und rs4889603 im *STX1B*-Gen (YongPing P. Chen et al., 2015; Hernandez et al., 2012). Einige Polymorphismen wurden sowohl als protektiv als auch als Suszeptibilitätsfaktoren beschrieben: rs591323 im *FGF20*-Gen (Hernandez et al., 2012; Jing et al., 2015; M. a Nalls et al., 2014), rs11868035 im *SREBF1/RAII*-Gen (M. a Nalls et al., 2014; Shulman et al., 2014), rs2942168 im *MAPT*-Gen (Sharma et al., 2012; Shulman et al., 2014), rs6812193 im *SCARB2/FAM47E/STBD1*-Gen (S. Chen et al., 2012; YongPing P. Chen et al., 2015; Gan-Or et al., 2015; Hernandez et al., 2012; Kalinderi, Bostantjopoulou, Katsarou, & Fidani, 2013; Liu et al., 2013; M. a Nalls et al., 2014) und rs2823357 im *MMP16*-Gen (Gan-Or et al., 2015). Diese unterschiedlichen Ergebnisse lassen sich zum einen auf verschiedene Studiendesigns zurückführen und zum anderen auf Heterogenität zwischen Bevölkerungsgruppen. Sowohl die



Effektstärke (Sharma et al., 2012), als auch die Frequenz der Varianten, 30 % in Europäern vs 10 % in Asiaten (Huo, Li, Zhao, & Wang, 2015), kann hierbei stark variieren. Wie für die Region PARK16 (rs823156) nachgewiesen wurde (H. Li et al., 2015), kann diese große Variabilität auf unterschiedliche LD-Muster zurückgeführt werden (Ardlie et al., 2002). Schließlich kann sich auch der Einfluss von Umweltfaktoren im Zusammenspiel mit bestimmten Varianten verstärken (Wirdefeldt, Adami, Cole, Trichopoulos, & Mandel, 2011). Für die Variante rs11724635 im *BST1*-Gen wurde in einer Studie eine Assoziation mit PD nur in Kombination mit dem Trinken von Brunnenwasser nachgewiesen (M. L. Chen, Lin, Lee, & Wu, 2014). Eine Einschränkung unserer Studie besteht darin, dass einige Meta-Analysen die für die SNP-Liste verwendet wurden, Daten der isländischen Bevölkerung beinhalten, was zu einer Verfälschung der Ergebnisse führen kann. Daher können wir auf Grund unserer Ergebnisse eine modifizierende Wirkung von häufigen Risikovarianten auf den Phänotyp der G2019S-Mutationsträger nicht ausschließen. Viele der beschriebenen Varianten stehen vielleicht nur im LD mit der eigentlich krankheitsverursachenden Variante. Es gilt nun diese zu identifizieren und die Interaktion der verschiedenen Varianten miteinander und der Umwelt zu beschreiben



## 4 Diskussion

Mit der vollständigen Entschlüsselung des humanen Genoms (E. S. Lander et al., 2001; Venter et al., 2001) wurden Genetikern neue Möglichkeiten gegeben. Die Entschlüsselung der genetischen Architektur (Anzahl, Art, Häufigkeit und Effektgröße von kausalen Varianten) komplexer, multigen bedingter Krankheiten ist eine davon (T. A. Manolio et al., 2009).

Durch die Verfügbarkeit von Millionen charakterisierter Polymorphismen über existierende Datenbanken wie z.B. 1000 Genomes (<http://www.1000genomes.org/>) oder HapMap (<http://www.hapmap.org/>) wurden die SNP-Chip Hochdurchsatztechnologien ausgereift, und Forscher in die Lage versetzt, aussagekräftige, genomweite Assoziationsstudien durchzuführen. Diese besitzen das Potential, genetische Einflussfaktoren zu identifizieren, die nicht über herkömmliche Kopplungsanalysen aufzufinden sind und können somit sowohl in Bezug auf die fehlende Erblichkeit als auch die molekulargenetische Diagnostik zu hilfreichen Erkenntnissen führen.

In den letzten Jahren wurde viel über Vor- und Nachteile, Möglichkeiten und Probleme von GWAS diskutiert (Hardy & Singleton, 2009; T. a Manolio & Collins, 2009; McCarthy et al., 2008; Ripke et al., 2014; Sham & Purcell, 2014; Spencer, Su, Donnelly, & Marchini, 2009; Visscher, Brown, McCarthy, & Yang, 2012). Im Vergleich zu Kopplungsanalysen ist die Rekrutierung von Nichtverwandten aus der allgemeinen Bevölkerung ein klarer Vorteil. Hierdurch können große Kollektive untersucht werden, die eine größere statistische Aussagekraft ermöglichen und die Identifikation von Varianten mit geringer Effektstärke ermöglichen (Hirschhorn & Daly, 2005). Kritisiert wird, dass man in der Regel nicht funktionell relevante krankheitsfördernde oder krankheitsauslösende Varianten detektiert, sondern nur einen benachbarten tag SNP. Auf der anderen Seite sind Regionen, die zwischen Nichtverwandten der Abstammung nach gleich sind, kleiner als IBD-Regionen bei Verwandten, wodurch die anschließende Identifikation der kausalen Variante im Vergleich zu familienbasierten Untersuchungen vereinfacht wird (Carlson et al., 2004).

Einer der größten limitierenden Faktoren von GWAS ist sicherlich die Notwendigkeit der Untersuchung eines sehr großen Kollektivs und der damit verbundene logistische und finanzielle Aufwand. Da bei der Auswertung Assoziationstests für alle Marker einzeln eingesetzt werden, steht zudem das Problem des multiplen Testens im Vordergrund, was wiederum die Anforderungen an die Struktur und Größe der Testgruppe verschärft. Eine kürzlich veröffentlichte Studie des Schizophrenie-Konsortiums zeigt jedoch auf, welchen Beitrag Assoziations-

studien für den wissenschaftlichen Erkenntnisprozess leisten können, solange sie genügend Individuen beinhalten und somit eine hohe statistische Power erreichen (Psychiatric, Consortium, & Committee, 2009; Ripke et al., 2014). Ripke et al. untersuchte insgesamt 38.131 Patienten und 114.674 Kontrollen aus 48 Untersuchungszentren, die alle bereits Teil vorangegangener Studien waren. Durch die gemeinsame Auswertung der Proben erlangten 108 Loci genomweite Signifikanz, 83 dieser Bereiche waren bisher noch nicht beschrieben. Weiter konnte er mit seiner global angelegten genomweiten Assoziationsstudie zeigen, dass sich die mit einer Krankheit assoziierten Gene nicht zufällig über das Genom verteilen, sondern alle in spezifischen Geweben exprimiert werden. Darüber hinaus waren die meisten auffälligen Genloci bereits früher mit Schizophrenie in Verbindung gebracht worden, woraus sich das mögliche Potential von neu-assoziierten Loci für das Verständnis der Signalwege und die Entwicklung von Medikamenten ableiten lässt. Die Auswertung von großen Datensätzen kann also die Komplexität der zu Grunde liegenden Erkrankung widerspiegeln und so zu einem besseren Verständnis der Ätiologie und der Pathophysiologie dieser Erkrankung beitragen.

Trotz erfolgreicher Identifikation vieler Suszeptibilitätsfaktoren in den letzten Jahren (T. A. Manolio, Brooks, & Collins, 2008), haben Forscher die erhofften Ziele noch nicht erreicht: Alle über GWAS identifizierten Varianten erklären nur einen sehr kleinen Teil der Erkrankungen in der Bevölkerung (Schork et al., 2009) und damit nicht wie angenommen einen Großteil der fehlenden Erblichkeit.

In Bezug auf die Parkinson Erkrankung konnten über Assoziationsstudien 28 Suszeptibilitätsfaktoren identifiziert werden (M. A. Nalls et al., 2011). Die meisten dieser Varianten erhöhen die Wahrscheinlichkeit PD zu entwickeln um einen Faktor 1.2-2.0 (Thomas Gasser, 2015). Zusammengenommen sind sie nur mit einer OR von 3 bis 4 assoziiert (M. a Nalls et al., 2014) und erklären somit auf Populationsebene nur etwa 10 % des Erkrankungsrisikos (Gibson, 2012; Keller et al., 2012; M. A. Nalls et al., 2011). In Kombination mit den Hochrisikoallelen der monogen vererbten Gene können demnach bis heute nur ca. 30 % der Erblichkeit von PD erklärt werden (T Gasser et al., 2011).

Wie bereits besprochen, gibt es verschiedene Hypothesen zur fehlenden Erblichkeit. Unter anderem wird sie in seltenen Varianten mit mittlerer Effektstärke oder CNVs gesucht, da diese Sequenzveränderungen von den üblichen GWAS-Designs nicht ausreichend abgedeckt werden. DeCODEs Gesamt-Genom-Sequenzierungs-Projekt bietet Möglichkeiten, Erkenntnisse über diese schwer zu identifizierenden, möglichen Risikofaktoren zu erlangen. Der Erfolg einer GWAS hängt von der Größe der untersuchten Gruppen ab (Wei et al., 2013;

Willems, Mihaescu, Sijbrands, Van Duijn, & Janssens, 2011) und davon, wie gut der verwendete Markersatz die Variation im Genom erfasst, d.h. wie groß das LD zwischen den disponierenden Varianten und den typisierten SNPs ist. Folglich bieten Populationen wie die Isländer große Vorteile durch größere genetische Übereinstimmung mit weniger Haplotypvarianten, höherem Kopplungsungleichgewicht und der Möglichkeit, große Kollektive zu generieren (Kristiansson et al., 2008; Service et al., 2006).

In der vorliegenden Arbeit wurden diese Vorteile genutzt, um drei mit PD assoziierte Bereiche des Genoms der isländischen Bevölkerung zu untersuchen.

#### **4.1 Fehlende Erbllichkeit - *EIF4G1* und *PARK2***

Die erste Publikation befasst sich mit der Rolle des 2011 beschriebenen PD-assoziierten *EIF4G1*-Gens (Chartier-Harlin et al., 2011), zu dem in den letzten Jahren widersprüchliche Ergebnisse publiziert wurden (S Lesage et al., 2012; Nuytemans et al., 2013; Schulte et al., 2012; Siitonen et al., 2013; Tucci et al., 2012). Auch unsere Ergebnisse erlauben keine klare pathogene Zuordnung des Gens und zeigen, wie komplex die genetische Architektur der Parkinson Erkrankung ist.

Wir konnten die ursprünglich beschriebene Variante p.Arg1205His in unserer Kohorte mit sehr niedriger Frequenz nachweisen, allerdings keine Assoziation zu PD in der isländischen Bevölkerung bestätigen. Auch die zusätzlich identifizierten Varianten sprechen nicht für eine entscheidende Rolle des Gens in der Ätiologie der Parkinson Erkrankung.

Nach heutigen Erkenntnissen ist die Interpretation eines solchen Gens sehr komplex und schwierig. Mutationen im *EIF4G1*-Gen scheinen sehr selten zu sein. Gerade in den letzten Jahren ist das Interesse an diesen seltenen Genvarianten, die in weniger als 5 % der Bevölkerung vorkommen und einen mittleren Effekt auf das Krankheitsrisiko haben, gestiegen (MRCV-Hypothese). Diese Varianten scheinen schon lange zu existieren, ihre Identifikation gestaltet sich jedoch schwierig, da sie nur in Kombination mit anderen Risikofaktoren zur Manifestation der Erkrankung führen. Darüber hinaus sind sie zu selten für GWAS, tauchen aber auch nicht so gehäuft in Familien auf, dass sie über Kopplung gefunden werden können. Auch die Tatsache, dass bei den meisten Assoziationsstudien zuerst den signifikantesten, häufigsten Varianten an einem Locus Beachtung geschenkt wird, trägt dazu bei, dass andere Varianten, die in diesem Genabschnitt auf demselben oder aber auf einem anderen Haplotyp-Block liegen, regelmäßig übersehen werden (Hardy & Singleton, 2009; Singleton, Myers, & Hardy, 2004). Möglicherweise sind viele dieser seltenen Sequenzveränderungen immer noch

nicht entdeckt. Der Fall der beiden Gene *LRRK2* und *GBA* zeigt jedoch, dass diese Allele als genetische Risikofaktoren der PD existieren (Goker-Alpan et al., 2004; Tan, 2006). Die gemeinsame Analyse mehrerer genomweiter Assoziationsstudien – der Arbeit von Ripke et al. für die Schizophrenie folgend - könnte helfen, sie zu finden.

Nach heutigen Erkenntnissen gibt es nicht einen bestimmten Haplotyp für das Krankheitsrisiko, eher eine Zahl verschiedener Haplotypen, die sich am selben Locus überschneiden (M. A. Nalls et al., 2011). Diese verschiedenen Kombinationen führen zu unterschiedlichen Phänotypen, welche es wiederum erschweren, die kausalen Varianten ausfindig zu machen. Würde man das *EIF4G1*-Gen als sog. pleomorphen Risikobereich einordnen, könnten Varianten mit medianem Effekt also auch in Kombination mit häufigen, durch GWAS identifizierten Risikovarianten im selben Bereich für das individuelle Risiko verantwortlich sein. Die Interaktion dieser Varianten miteinander ist jedoch sehr komplex und schwer zu erfassen. Eine Mutation könnte z.B. nur eine Wirkung zeigen, sofern ein physiologischer Prozess bereits durch eine andere Mutation geschwächt wurde. D.h. die Kombination von nur zwei dieser Mutationen reicht möglicherweise zur Manifestation der Erkrankung aus, womit die reduzierte Penetranz der p.Arg1205His Variante erklärbar wäre. Auch Ripke et al. konnte eine Übereinstimmung von Genen mit seltenen Hochrisikoallelen und Genen mit erhöhter Suszeptibilität gegenüber PD nachweisen (Ripke et al., 2014).

In der zweiten Publikation der vorliegenden Arbeit haben wir uns mit der Rolle heterozygoter CNVs im *PARK2*-Gen beschäftigt.

Trotz verbesserter Techniken im molekulargenetischen Bereich ist es nach wie vor sehr schwierig, neue Erkenntnisse über CNVs zu erlangen. Mit einer geringen Frequenz in der Allgemeinbevölkerung können sie nur in Studien mit einer entsprechend hohen Kollektivgröße detektiert werden, und ihre Identifikation ist zeitaufwendig und teuer. Auf Grund dieser Schwierigkeiten wird ihre Frequenz und ihr Anteil an der genetischen Variabilität wahrscheinlich deutlich unterschätzt.

Die Ergebnisse unserer Studie deuten jedoch darauf hin, dass sie einen Teil dieser fehlenden Erbllichkeit erklären können. Wir konnten in einer genomweiten Assoziationsstudie zeigen, dass heterozygote CNVs in einem rezessiv vererbten Gen ein erhöhtes Risiko darstellen, an PD zu erkranken. Das klinische Bild dieser heterozygoten Patienten entsteht nach heutigen Erkenntnissen aus dem Zusammenspiel der Genprodukte der beiden unterschiedlichen Allele. Hier sind verschiedene Wirkungsweisen beschrieben: *Haploinsuffizienz*, *Loss-of-function*,

*Gain-of-function* oder ein *Dominant negativ Effekt* (Sidransky, 2006; von Coelln, Dawson, & Dawson, 2004; A. West et al., 2002).

Durch Ausschluss von *Compound*-Heterozygotie konnten wir zeigen, dass das erhöhte Risiko der heterozygoten Mutationsträger auf einem Dosiseffekt beruht. Auch eine vor kurzem veröffentlichte Studie hat gezeigt, dass typische mit PD assoziierte Symptome mit der Anzahl der mutierten Gene korrelieren (Grünewald et al., 2013). Darüber hinaus weiß man, dass Triplikationen einen größeren Effekt haben als Duplikationen (Fuchs et al., 2007; Singleton et al., 2004).

Auf der anderen Seite können homozygote *LRRK2* G2019S-Mutationsträger, wie auch in unserer Studie, teilweise nicht von Heterozygoten unterschieden werden (Ishihara et al., 2006). Hier führt eine zweite Mutation in einem dominanten Gen nicht zu einer ernsteren Erkrankung oder einem früheren Erkrankungsalter. Und viele heterozygote Mutationsträger zeigen bis ins hohe Alter keinerlei Symptome. Die Interaktion zwischen verschiedenen pathogenen Mutationen ist demnach nicht vorher zu sagen. Der Grund hierfür ist bisher nicht bekannt und ein breites Feld für weitere Forschung. Es ist jedoch offensichtlich, dass es andere bis jetzt nicht identifizierte genetische oder umweltbedingte *Modifier* geben muss, die beispielsweise im Fall der *LRRK2* G2019S-Mutationsträger den Phänotyp beeinflussen. Bei multifaktoriellen Erkrankungen ist durchaus zu erwarten, dass die Wahrscheinlichkeit für das Entstehen der Erkrankung oder ihre Ausprägung von vielen Gen-Gen und Gen-Umwelt-Wechselwirkungen beeinflusst wird (Bieckeböller & Fische, 2007). Doch genomweite Interaktionsanalysen stellen eine große Herausforderung dar (Cordell, 2009).

#### **4.2 Diagnostik - *EIF4G1* und *PARK2***

Auf Grund des Fortschritts durch „Panel Sequenzierung“, d.h. die gleichzeitige Bestimmung der genetischen Variabilität in einer großen Zahl verschiedener Gene, die mit der Krankheit mehr oder weniger stark assoziiert sind, oder auch durch Exom-Sequenzierung, ist in Bezug auf das Parkinson-Syndrom das Interesse an genetischer Beratung und Risikovorhersage gestiegen (Falcone, Wood, Xie, Siderowf, & Van Deerlin, 2011). Ein geringes Risiko, ein Risiko in Verbindung mit Umweltfaktoren oder ein hohes Risiko auch ohne zusätzliche Faktoren: in welchem Bereich dieser möglichen Kombinationen das Risiko einer individuellen Person anzusiedeln ist, ist jedoch schwer vorherzusagen. Eine direkte Testung zur Ermittlung des Risikos gibt es nur in seltenen Fällen mit monogenetischer PD. Demzufolge ist heute die genetische Testung in PD auf einige wenige Fälle reduziert, in denen der familiäre Hintergrund

und das Erkrankungsalter einen Hinweis auf eine der monogen vererbten Formen gibt (Alcalay et al., 2010). Auch wenn im Moment noch keine präventive Therapie zur Verfügung steht, kann eine genetische Testung und Beratung Unsicherheiten beseitigen und, falls nötig, eine frühe pharmakologische Intervention bei Mutationsträgern ermöglichen. Gerade in Familien mit gehäuftem Auftreten der Erkrankung hilft eine sichere Diagnose, die Symptome besser zu verstehen. Patienten mit einer frühmanifestierenden Parkinson Erkrankung zweifeln oft an der Diagnose einer in der Regel spätmanifestierenden Erkrankung. Sie können durch eine Testung Klarheit erlangen.

Aufgrund der Komplexität der Krankheit ist die Interpretation der Testergebnisse jedoch eine große Herausforderung für genetische Berater. Die Forschung der letzten Jahre hat gezeigt, dass die genetische und phänotypische Variabilität um vieles größer ist als man dachte. Darüber hinaus wird man immer mehr Patienten mit seltenen Mutationen finden, deren Signifikanz man zunächst noch nicht kennt. Damit diese große Zahl an genetischen Varianten in der Diagnostik und der Medizin ihren Nutzen finden kann, muss man die pathogenen Mutationen von den seltenen Varianten unterscheiden lernen, die möglicherweise einen funktionellen Effekt haben, aber nicht pathogen sind. Neben den festgelegten Kriterien zur Beurteilung der Pathogenität einer Mutation (MacArthur et al., 2014), ist es von großer Wichtigkeit, so viel wie möglich über die Penetranz der Varianten und den jeweiligen Vererbungsmodus zu erfahren (Berardelli et al., 2013). Dieses Wissen bildet das Fundament, auf dem eine solide medizinische Interpretation nach genetischer Testung steht.

Es gibt eine Vielzahl von Genen, die wie das *EIF4G1*-Gen bis jetzt noch nicht eindeutig in ihrer pathogenen Wirkung bestätigt wurden (Tabelle 4). Alle Ergebnisse bezüglich dieser Gene bei einer genetischen Testung sollten mit großer Vorsicht interpretiert werden. Eine falsche Zuordnung von Varianten kann rasch negative Folgen für den Patienten und seine Familie haben. Auch wenn die anfängliche Einordnung eines Gens als pathogen meist stimmt, wird oft die Kausalität der einzelnen Varianten falsch interpretiert (MacArthur et al., 2014). Wie bereits besprochen, werden bei Assoziationsstudien in der Regel nur die stärksten Signale näher beschrieben. Krankheitsrelevante Loci mit plausibler Effektgröße sind jedoch möglicherweise auch in großen Studien nicht unter den besten Kandidaten (Craddock, O'Donovan, & Owen, 2008). Auch unsere Analyse des *EIF4G1*-Gens konnte keine klare Assoziation mit PD bestätigen, vielmehr fanden wir keine Hinweise auf eine pathogene Wirkung. Da es sich aber offensichtlich um einen Risikofaktor mit reduzierter Penetranz handelt, sollte der Berater die Ergebnisse mit Vorsicht interpretieren.



In Hinblick auf *PARK2* wird im Moment die Testung bei Patienten mit typischem PD, die entweder einen familiären Hintergrund haben, der mit einem rezessiven Erbgang übereinstimmt oder die im Alter von  $\leq 40$  erkrankten, angeboten (Kalinderi, Bostantjopoulou, & Fidani, 2016). Auch unsere Ergebnisse befürworten kein Routinescreening auf heterozygote Mutationsträger, da das Erkrankungsrisiko zu gering ist. Für eine Person, die einen nahen Verwandten mit bestätigter *PARK2*-Mutation hat, kann sie jedoch durchaus von Bedeutung sein. Ungeachtet dessen wird eine genetische Testung immer häufiger sowohl für familiäre als auch sporadische PD angeboten. Daher ist es von großer Wichtigkeit, das Vererbungsmuster und die Penetranz besser zu verstehen. Die Interpretation der Testergebnisse kann hierdurch verbessert werden.

Zusammenfassend lässt sich sagen, dass in den meisten Fällen eine Kombination aus Hochrisiko- und Niedrigrisikofaktoren für die Parkinson Erkrankung verantwortlich ist, worüber keine begründete Voraussage über ein individuelles Krankheitsrisiko möglich ist. Dies stellt den Nutzen einer genetischen Testung bei Parkinson außerhalb der familiären Fälle in Frage. Im Blick auf die Zukunft kann eine Testung jedoch von großem Nutzen sein. Die Forschung der letzten Jahre hat gezeigt, dass die Genotyp-Phänotyp Korrelation um Vieles variabler ist als man dachte, und der Gebrauch von Exom-Sequenzierung wird das Wissen um diese Varianz noch stark ausweiten. Die Parkinson Erkrankung ist klinisch eine sehr heterogene Krankheit und viele der Symptome sind auch in ähnlichen neurologischen Erkrankungen zu beobachten. Klarheit über diese verschiedenen Subtypen kann teilweise durch mehr molekulare Informationen erreicht werden.

Auch in Bezug auf mögliche Therapien und klinische Studien kann die genetische Untersuchung aus wissenschaftlicher Sicht ausgesprochen wichtig sein. Nicht nur Patienten mit familiengebundenem PD, sondern auch sporadische Patienten können auf Grund ihrer individuellen Kombination genetischer Risikofaktoren gruppiert werden (genetische Stratifizierung), wodurch neue Erkenntnisse hinsichtlich des zugrunde liegenden krankheitsverursachenden Signalwegs und der klinischen Merkmale bestimmter Gruppen gewonnen werden können. Beispielsweise wurde gezeigt, dass nicht-motorische Merkmale, wie z.B. kognitive Fehlfunktionen in Patienten mit *GAB*-Mutationen häufiger sind (K. Brockmann et al., 2011), wohingegen Varianten im 3'-untranslatierten Bereich des *SNCA*-Gens mit einem früheren Erkrankungsalter und einem schnelleren Fortschreiten der Krankheit assoziiert sind (Kathrin Brockmann et al., 2013). Beim Vorhandensein des ApoE4 Allels ist die Demenz-Entwicklung

in PD Patienten erhöht (Mata et al., 2014). Je mehr man also über diese Genotyp-Phänotyp-Korrelationen weiß, desto besser kann man die Ursache einer Erkrankung eingrenzen, den spezifischen Krankheitsverlauf begleiten und womöglich mit einer frühen Intervention beginnen.

Eine personalisierte Behandlung ist bei heterogenen Erkrankungen sehr wichtig und diese individuelle Behandlung ist nur auf der Grundlage einer bekannten Genetik und Biochemie möglich (Wurtman, 2013).

## 5 Ausblick

Auch wenn genomweite Assoziationsstudien in der Kritik stehen, nicht genug zum Verständnis und der Vorhersagbarkeit von komplexen Krankheiten beigetragen zu haben, konnten wir in dieser Arbeit zeigen, dass GWAS – insbesondere in hoch-informativen Kohorten – ein wichtiges Werkzeug zur Identifikation und Charakterisierung von Risikovarianten sind. Da die Menge an Daten eine große Herausforderung für Forschung und Diagnostik darstellt, kommt es in Zukunft darauf an, Wege zu finden, mit dieser enormen Datenflut umzugehen, die wesentlichen genetischen Varianten herauszufiltern und ihre funktionelle und pathophysiologische Bedeutung aufzuklären. Auch wenn es im Moment noch sehr schwierig scheint, diese neuen Erkenntnisse in biologisches Wissen zu verwandeln, wird auf lange Sicht das Verständnis der Pathogenese erweitert werden und GWAS wird wesentlich dazu beitragen, die Ätiologie der untersuchten Erkrankung in ihrer Gesamtheit besser zu verstehen. Genau dieses tiefere Verständnis der genetischen und neuronalen Grundlagen (International, Sclerosis, & Consortium, 2011) verspricht große Fortschritte bei der Entwicklung neuer Medikamente und Therapien. Außerdem bahnen die Verknüpfung von Krankheit und Genom den Weg für personalisierte Behandlungen, denen besonders bei biologisch heterogenen Erkrankungen wie der Parkinson Erkrankung große Bedeutung zukommen wird.

Da man heute weiß, dass bei der PD jeder Risikofaktor für sich alleine genommen wenig Gewicht hat, und erst das Zusammenspiel verschiedener genetischer Veränderungen zur Manifestation der Erkrankung führt, wird es in Zukunft von großer Wichtigkeit sein, Gen-Gen und Gen-Umwelt Interaktionen besser zu erfassen und zu verstehen. Darüber hinaus sollte den protektiven Varianten mehr Aufmerksamkeit geschenkt werden (Thomas Gasser, 2015).



## 6 Zusammenfassung

Die Identifizierung genetischer Varianten mit unterschiedlichster Effektstärke spielt bei der Erforschung komplexer und multifaktorieller Erkrankungen, wie der Parkinson Erkrankung, eine entscheidende Rolle. Seit dem letzten Jahrzehnt stehen hierfür die sogenannten genomweiten Assoziationsstudien zur Verfügung. Sie haben bereits einen erheblichen Beitrag zum Verständnis der Ätiologie dieser Erkrankungen beigetragen. Dennoch ist die genetische Variabilität in ihrer Gesamtheit bei weitem nicht erklärt und man ist weiter auf der Suche nach kausativen Varianten, die die „fehlende Erbllichkeit“ erklären sollen. Stark vorangetrieben wird diese Suche durch das Projekt von deCODE Genetics. Das isländische Pharmaunternehmen nutzt die Ähnlichkeit der genetischen Struktur des Inselvolkes und deren Offenheit gegenüber neuen Technologien und Vorgehensweisen, zur Identifikation von Krankheitsgenen. Mittlerweile wurde das Genom von 2.600 Isländern vollständig sequenziert. Mit Hilfe dieser Informationen und dem Wissen zu genealogischen Daten war es möglich, die Genomvarianten fast aller Isländer in einer proprietären Datenbank abzubilden.

In der vorliegenden Arbeit, wurden diese Ressourcen genutzt um drei bereits kartierte Genorte und ihre Beteiligung bei der Pathogenese der Parkinson Erkrankung weiter zu klären.

Das erste Teilprojekt befasste sich mit der Charakterisierung des *EIF4G1*-Gens. Durch die Untersuchung einer großen europäischen Kohorte sporadischer Patienten konnten wir die in der Literatur beschriebene niedrige Frequenz der pathogenen Mutationen (p.Ala502Val und p.Arg1205His) in Exon 12 und 24 des Gens, bestätigen. Zusätzlich wurde eine Testung der p.Arg1205His-Variante auf Assoziation mit PD in der isländischen Bevölkerung durchgeführt. Hierbei konnte keine signifikante Assoziation nachgewiesen werden. Die Prävalenz anderer *EIF4G1*-Mutationen wurde durch Exom-Sequenzierung in Patienten mit familiengebundener Erkrankung untersucht. Auch diese Ergebnisse deuten nicht darauf hin, dass Varianten in diesem Gen einen hohen Risikofaktor darstellen. Vielmehr handelt es sich um einen Risikofaktor mit reduzierter Penetranz. Bei einer genetischen Beratung sollten Varianten im *EIF4G1*-Gen daher mit Vorsicht interpretieren werden.

Im zweiten Teil der Arbeit wurden CNVs im *PARK2*-Gen in einem Datensatz von 105.749 genotypisierten Isländern identifiziert und in einer Stichprobe von 1.415 Parkinson-Patienten und 40.474 Kontrollen auf Assoziation mit PD getestet. Wir konnten für heterozygote CNV-Träger eine signifikante Risikoerhöhung bezüglich des Auftretens von PD nachweisen (OR =

1.69,  $P = 0.03$ ). *Compound-heterozygote* Mutationsträger wurden durch die zusätzliche Suche nach *Missense*-Mutationen im Gen ausgeschlossen. Die Annahme einer Beteiligung heterozygoter *PARK2*-CNVs an der Manifestation der Erkrankung konnten wir durch die Durchführung einer Meta-Analyse zusätzlich stärken. Unsere Ergebnisse verdeutlichen, dass bei einem autosomal rezessiven Gen die Rolle heterozygoter Mutationen nicht unterschätzt werden sollte. Diese Sequenzveränderungen können als Suszeptibilitätsfaktoren zum Krankheitsbild beitragen und erklären einen weiteren Teil der Erblichkeit.

Im dritten Teil der Arbeit konnten wir eine Assoziation der bekannten und häufigen *missense*-Veränderung p.Gly2019Ser im *LRRK2* - Gen zur PD nachweisen. Trotz autosomal-dominantem Erbgang gibt es viele heterozygote Mutationsträger, die keine Krankheitssymptome zeigen, was auf zusätzliche genetische und umweltbedingte Faktoren zurückzuführen ist. Eine mögliche modifizierende Wirkung von bekannten Suszeptibilitätsfaktoren der PD auf die Penetranz dieser Mutation konnten wir im Rahmen eines kumulativen Risikomodells ausschließen.

## 7 Summary

Identification of genetic risk factors can shed new light on important aspects of pathogenesis of complex diseases such as PD. Over the last years an increasing number of data on common risk alleles have become available, made possible by genome wide association studies. However, the genetic diversity of PD has only been partly explained and researchers remain focused on discovering new risk alleles to explain the “missing heritability” of the disorder.

The Icelandic pharmaceutical company deCODE Genetics uses the similarity of genetic architecture of the insular people and their openness to new technologies, for identification of new human disease genes. At the time of writing, the genome of 2.600 Icelanders has been whole-genome sequenced. This information combined with carefully kept genealogical records, which go back to the time when the island was settled and genealogical tables, allow deCODE to “impute” the genomes of nearly everyone born of Icelandic parents.

In the present thesis, these resources have been used to more closely characterize three known candidate genes of PD.

The first study focused on the characterization of the *EIF4G1* gene. Therefore we screened a large German sample of sporadic PD patients for p.Ala502Val and p.Arg1205His mutations, the most promising variants, by means of high resolution melting curve analysis. In addition we tested the p.Arg1205His variant in the Icelandic population for association with PD. Subsequently we assessed the prevalence of other *EIF4G1* mutations by sequencing the complete coding region of this gene in PD patients with a dominant family history. In summary, we provide a large number of data that do not easily support *EIF4G1* as high-risk PD locus. Rather, a risk increase for the seminal p.Arg1205His mutation is not associated with PD in the Icelandic population. Although our observations do not rule out *EIF4G1* as an important genetic modifier of PD, caution should be taken when interpreting genotyping results in isolated patients and PD families.

In the second part of the thesis our objective was to investigate the role of heterozygous and compound heterozygous CNVs in the *PARK2* gene as a risk factor for PD in the Icelandic population. Focusing on the samples tested for CNVs we had access to 1415 PD cases and 104.320 controls. Our association study results support the hypothesis that heterozygous exon rearrangements in *PARK2* contribute to disease susceptibility (OR=1,7; p=0,03). To exclude that compound heterozygosity explains the risk, we identified all missense mutations in the

gene through whole genome sequencing and analyzed the mutational phase of all samples that had been phased using long-rang-phasing information. No PD patients were found compound heterozygous for missense mutations and CNVs. To further delineate our signal, we conducted a meta analysis of published studies on heterozygous CNV carriers, which showed the same results. Our work provides further evidence that heterozygous *PARK2* CNV carriers are at increased risk of developing PD. Thus, the impact of heterozygous mutations in recessive genes should not be underestimated and should be the focus of future investigation. The mutations may act as susceptibility factors or disease modifiers and might further contribute to the heritability of familial and sporadic PD.

Inheritance in *LRRK2*-associated PD is thought to be autosomal dominant. Nevertheless several heterozygous carriers are unaffected. This is due to reduced penetrance of *LRRK2* mutations. Reasons that can explain this phenomenon, are additional genetic and environmental risk factors, or a combination of both. In the third study a possible modifying effect of common markers on penetrance of the common missense mutation p.Gly2019Ser in the *LRRK2* gene could be excluded by using a cumulative risk model based on all published PD risk variants.



## 8 Literatur

- Aarsland, D., Andersen, K., Larsen, J. P., Lolk, A., & Kragh-Sørensen, P. (2003). Prevalence and characteristics of dementia in Parkinson disease: an 8-year prospective study. *Archives of Neurology*, *60*, 387–392. doi:10.1001/archneur.60.3.387
- Abbas, N., Lucking, C. B., Ricard, S., Durr, A., Bonifati, V., De Michele, G., ... Brice, A. (1999). A wide variety of mutations in the parkin gene are responsible for autosomal recessive parkinsonism in Europe. French Parkinson's Disease Genetics Study Group and the European Consortium on Genetic Susceptibility in Parkinson's Disease. *Hum Mol Genet*, *8*(4), 567–574.
- Alcalay, R. N., Caccappolo, E., Mejia-Santana, H., Tang, M. X., Rosado, L., Ross, B. M., ... Clark, L. N. (2010). Frequency of known mutations in early-onset Parkinson disease: implication for genetic counseling: the consortium on risk for early onset Parkinson disease study. *Archives of Neurology*, *67*(9), 1116–1122. doi:10.1001/archneurol.2010.194
- Alves, G., Forsaa, E. B., Pedersen, K. F., Dreetz Gjerstad, M., & Larsen, J. P. (2008). Epidemiology of Parkinson's disease. In *Journal of Neurology* (Vol. 255, pp. 18–32). doi:10.1007/s00415-008-5004-3
- Ardlie, K. G., Kruglyak, L., & Seielstad, M. (2002). Patterns of linkage disequilibrium in the human genome. *Nature Reviews. Genetics*, *3*(4), 299–309. doi:10.1038/nrg777
- Atzaba-Poria, N., Pike, A., & Deater-Deckard, K. (2004). Do risk factors for problem behaviour act in a cumulative manner? An examination of ethnic minority and majority children through an ecological perspective. *Journal of Child Psychology and Psychiatry*, *45*(4), 707–718.
- Azzopardi, D., Dallosso, A. R., Eliason, K., Hendrickson, B. C., Jones, N., Rawstorne, E., ... Cheadle, J. P. (2008). Multiple rare nonsynonymous variants in the Adenomatous polyposis coli gene predispose to colorectal adenomas. *Cancer Research*, *68*, 358–363. doi:10.1158/0008-5472.CAN-07-5733
- Baba, M., Nakajo, S., Tu, P. H., Tomita, T., Nakaya, K., Lee, V. M., ... Iwatsubo, T. (1998). Aggregation of alpha-synuclein in Lewy bodies of sporadic Parkinson's disease and dementia with Lewy bodies. *The American Journal of Pathology*, *152*(4), 879–84.

- Retrieved from  
<http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=1858234&tool=pmcentrez&rendertype=abstract>
- Baker, M. (2008). Genome studies: genetics by numbers. *Nature*. doi:10.1038/451516a
- Bennett, D. A., Beckett, L. A., Murray, A. M., Shannon, K. M., Goetz, C. G., Pilgrim, D. M., & Evans, D. A. (1996). Prevalence of parkinsonian signs and associated mortality in a community population of older people. *The New England Journal of Medicine*, 334, 71–76. doi:10.1056/NEJM199601113340202
- Berardelli, a, Wenning, G. K., Antonini, a, Berg, D., Bloem, B. R., Bonifati, V., ... Vidailhet, M. (2013). EFNS/MDS-ES/ENS [corrected] recommendations for the diagnosis of Parkinson's disease. *European Journal of Neurology : The Official Journal of the European Federation of Neurological Societies*, 20(1), 16–34. doi:10.1111/ene.12022
- Berg, D., Schweitzer, K. J., Leitner, P., Zimprich, A., Lichtner, P., Belcredi, P., ... Gasser, T. (2005). Type and frequency of mutations in the LRRK2 gene in familial and sporadic Parkinson's disease\*. *Brain : A Journal of Neurology*, 128(Pt 12), 3000–3011. doi:10.1093/brain/awh666
- Bieckeböller, H., & Fische, C. (2007). *Einführung in die Genetische Epidemiologie*. Springer.
- Biggins, C. A., Boyd, J. L., Harrop, F. M., Madeley, P., Mindham, R. H., Randall, J. I., & Spokes, E. G. (1992). A controlled, longitudinal study of dementia in Parkinson's disease. *Journal of Neurology, Neurosurgery, and Psychiatry*, 55, 566–571. doi:10.1136/jnnp.55.7.566
- Bodmer, W., & Bonilla, C. (2008). Common and rare variants in multifactorial susceptibility to common diseases. *Nature Genetics*, 40, 695–701. doi:10.1038/ng.f.136
- Brockmann, K., Schulte, C., Hauser, A. K., Lichtner, P., Huber, H., Maetzler, W., ... Gasser, T. (2013). SNCA: Major genetic modifier of age at onset of Parkinson's disease. *Movement Disorders*, 28(9), 1217–1221. doi:10.1002/mds.25469
- Brockmann, K., Srulijes, K., Hauser, A. K., Schulte, C., Csoti, I., Gasser, T., & Berg, D. (2011). GBA-associated PD presents with nonmotor characteristics. *Neurology*, 77(3), 276–280. doi:10.1212/WNL.0b013e318225ab77

- Bruggemann, N., Mitterer, M., Lanthaler, A. J., Djarmati, A., Hagenah, J., Wiegers, K., ... Lohmann, K. (2009). Frequency of heterozygous Parkin mutations in healthy subjects: need for careful prospective follow-up examination of mutation carriers. *Parkinsonism Relat Disord*, *15*(6), 425–429. doi:10.1016/j.parkreldis.2008.11.014
- Cardenas, M. E., Cutler, N. S., Lorenz, M. C., Di Como, C. J., & Heitman, J. (1999). The TOR signaling cascade regulates gene expression in response to nutrients. *Genes & Development*, *13*, 3271–3279. doi:10.1101/gad.13.24.3271
- Carlson, C. S., Eberle, M. A., Kruglyak, L., & Nickerson, D. A. (2004). Mapping complex disease loci in whole-genome association studies. *Nature*, *429*, 446–452. doi:10.1038/nature02623
- Chang, K.-H., Wu, Y.-R., Chen, Y.-C., Fung, H.-C., Lee-Chen, G.-J., & Chen, C.-M. (2015). STK39, But Not BST1, HLA-DQB1, and SPPL2B Polymorphism, Is Associated With Han-Chinese Parkinson's Disease in Taiwan. *Medicine (Baltimore)*, *94*(41), 1690.
- Chang, X. L., Mao, X. Y., Li, H. H., Zhang, J. H., Li, N. N., Burgunder, J. M., ... Tan, E. K. (2011). Association of GWAS loci with PD in China. *American Journal of Medical Genetics, Part B: Neuropsychiatric Genetics*, *156*(3), 334–339. doi:10.1002/ajmg.b.31167
- Chartier-Harlin, M. C., Dachsel, J. C., Vilarino-Guell, C., Lincoln, S. J., LePrete, F., Hulihan, M. M., ... Farrer, M. J. (2011). Translation initiator EIF4G1 mutations in familial Parkinson disease. *Am J Hum Genet*, *89*(3), 398–406. doi:10.1016/j.ajhg.2011.08.009
- Chen, M. L., Lin, C. H., Lee, M. J., & Wu, R. M. (2014). BST1 rs11724635 interacts with environmental factors to increase the risk of Parkinson's disease in a Taiwanese population. *Parkinsonism and Related Disorders*, *20*(3), 280–283. doi:10.1016/j.parkreldis.2013.11.009
- Chen, S., Zhang, Y., Chen, W., Wang, Y., Liu, J., Rong, T. Y., ... Chen, S. Di. (2012). Association study of SCARB2 rs6812193 polymorphism with Parkinson's disease in Han Chinese. *Neuroscience Letters*, *516*(1), 21–23. doi:10.1016/j.neulet.2012.03.035
- Chen, Y., Chen, K., Song, W., Chen, X., Cao, B., Huang, R., ... Shang, H. F. (2013). VPS35 Asp620Asn and EIF4G1 Arg1205His mutations are rare in Parkinson disease from southwest China. *Neurobiol Aging*, *34*(6), 1709.e7–8. doi:10.1016/j.neurobiolaging.2012.11.003

- Chen, Y. P., Song, W., Huang, R., Chen, K., Zhao, B., Li, J., ... Shang, H.-F. (2013). GAK rs1564282 and DGKQ rs11248060 increase the risk for Parkinson's disease in a Chinese population. *Journal of Clinical Neuroscience*, 20(6), 880–883.
- Chen, Y. P., Yuan, X. Q., Cao, B., Wei, Q. Q., Ou, R. W., Yang, J., ... Shang, H. F. (2015). No association of FAM47E rs6812193, SCARB2 rs6825004 and STX1B rs4889603 polymorphisms with Parkinson's disease in a Chinese Han population. *Journal of Neural Transmission*, 122(11), 1547–1552. doi:10.1007/s00702-015-1430-4
- Clark, L. N., Afridi, S., Karlins, E., Wang, Y., Mejia-Santana, H., Harris, J., ... Marder, K. (2006). Case-control study of the parkin gene in early-onset Parkinson disease. *Arch Neurol*, 63(4), 548–552. doi:10.1001/archneur.63.4.548
- Colhoun, H. M., McKeigue, P. M., & Davey Smith, G. (2003). Problems of reporting genetic associations with complex outcomes. *Lancet*, 361(9360), 865–872.
- Collins, F. (2011). *Meine Gene - mein Leben: Auf dem Weg zur personalisierten Medizin*. Spektrum Akademischer Verlag.
- Coop, G., Wen, X., Ober, C., Pritchard, J. K., & Przeworski, M. (2008). High-resolution mapping of crossovers reveals extensive variation in fine-scale recombination patterns among humans. *Science (New York, N.Y.)*, 319, 1395–1398. doi:10.1126/science.1151851
- Coppedè, F. (2012). Genetics and Epigenetics of Parkinson's Disease. *The Scientific World Journal*. doi:10.1100/2012/489830
- Cordell, H. J. (2009). Detecting gene-gene interactions that underlie human diseases. *Nature Reviews. Genetics*, 10(6), 392–404. doi:10.1038/nrg2579
- Craddock, N., O'Donovan, M. C., & Owen, M. J. (2008). Genome-wide association studies in psychiatry: lessons from early studies of non-psychiatric and psychiatric phenotypes. *Molecular Psychiatry*. doi:10.1038/mp.2008.45
- Dächsel, J. C., Ross, O. A., Mata, I. F., Kachergus, J., Toft, M., Cannon, A., ... Farrer, M. J. (2007). Lrrk2 G2019S substitution in frontotemporal lobar degeneration with ubiquitin-immunoreactive neuronal inclusions. *Acta Neuropathologica*, 113(5), 601–606. doi:10.1007/s00401-006-0178-1
- Dauer, W., & Przedborski, S. (2003). Parkinson's disease: mechanisms and models. *Neuron*,

- 39(6), 889–909. doi:S0896627303005683 [pii]
- de Rijk, M. C., Tzourio, C., Breteler, M. M., Dartigues, J. F., Amaducci, L., Lopez-Pousa, S., ... Rocca, W. A. (1997). Prevalence of parkinsonism and Parkinson's disease in Europe: the EUROPARKINSON Collaborative Study. European Community Concerted Action on the Epidemiology of Parkinson's disease. *Journal of Neurology, Neurosurgery, and Psychiatry*, *62*(1), 10–15. doi:10.1136/jnnp.62.1.10
- Den Dunnen, J. T., & Antonarakis, S. E. (2000). Mutation nomenclature extensions and suggestions to describe complex mutations: A discussion. *Human Mutation*, *15*(1), 7–12. doi:10.1002/(SICI)1098-1004(200001)15:1<7::AID-HUMU4>3.0.CO;2-N
- Devlin, B., & Roeder, K. (1999). Genomic control for association studies. *Biometrics*, *55*, 997–1004. doi:10.1111/j.0006-341X.1999.00997.x
- Di Fonzo, A., Rohé, C. F., Ferreira, J., Chien, H. F., Vacca, L., Stocchi, F., ... Bonifati, V. (2005). A frequent LRRK2 gene mutation associated with autosomal dominant Parkinson's disease. *Lancet*, *365*, 412–415. doi:10.1016/S0140-6736(05)17829-5
- Do, C. B., Tung, J. Y., Dorfman, E., Kiefer, A. K., Drabant, E. M., Francke, U., ... Eriksson, N. (2011). Web-based genome-wide association study identifies two novel loci and a substantial genetic component for Parkinson's disease. *PLoS Genetics*, *7*(6), e1002141. doi:10.1371/journal.pgen.1002141
- Dooneief, G., Mirabello, E., Bell, K., Marder, K., Stern, Y., & Mayeux, R. (1992). An estimate of the incidence of depression in idiopathic Parkinson's disease. *Archives of Neurology*, *49*, 305–307. doi:10.1001/archneur.1992.00530270125028
- Eichler, E. E., Flint, J., Gibson, G., Kong, A., Leal, S. M., Moore, J. H., & Nadeau, J. H. (2010). Missing heritability and strategies for finding the underlying causes of complex disease. *Nature Reviews. Genetics*, *11*, 446–450. doi:10.1038/nrg2809
- Elbaz, A. (2008). LRRK2: bridging the gap between sporadic and hereditary Parkinson's disease. *The Lancet Neurology*, *7*(7), 562–564.
- Elbaz, A., Grigoletto, F., Baldereschi, M., Breteler, M. M., Manubens-Bertran, J. M., Lopez-Pousa, S., ... Rocca, W. A. (1999). Familial aggregation of Parkinson's disease: a population-based case-control study in Europe. EUROPARKINSON Study Group. *Neurology*, *52*(9), 1876–1882.

- Falcone, D. C., Wood, E. M., Xie, S. X., Siderowf, A., & Van Deerlin, V. M. (2011). Genetic testing and Parkinson disease: Assessment of patient knowledge, attitudes, and interest. *Journal of Genetic Counseling*, 20(4), 384–395. doi:10.1007/s10897-011-9362-0
- Farrer, M., Chan, P., Chen, R., Tan, L., Lincoln, S., Hernandez, D., ... Langston, J. W. (2001). Lewy bodies and parkinsonism in families with parkin mutations. *Ann Neurol*, 50(3), 293–300.
- Fénelon, G., Mahieux, F., Huon, R., & Ziegler, M. Hallucinations in Parkinson's disease: prevalence, phenomenology and risk factors., 123 ( Pt 4 Brain : a journal of neurology 733–745 (2000). doi:10.1093/brain/123.4.733
- Foroud, T., Uniacke, S. K., Liu, L., Pankratz, N., Rudolph, A., Halter, C., ... Nichols, W. C. (2003). Heterozygosity for a mutation in the parkin gene leads to later onset Parkinson disease. *Neurology*, 60(5), 796–801.
- Foulkes, A. (2009). *Applied Statistical Genetics with R*. Springer.
- Fuchs, J., Nilsson, C., Kachergus, J., Munz, M., Larsson, E. M., Schüle, B., ... Farrer, M. J. (2007). Phenotypic variation in a large Swedish pedigree due to SNCA duplication and triplication. *Neurology*, 68(12), 916–922. doi:10.1212/01.wnl.0000254458.17630.c5
- Gabriel, S. B., Schaffner, S. F., Nguyen, H., Moore, J. M., Roy, J., Blumenstiel, B., ... Altshuler, D. (2002). The structure of haplotype blocks in the human genome. *Science (New York, N.Y.)*, 296, 2225–2229. doi:10.1126/science.1069424
- Gan-Or, Z., Girard, S. L., Noreau, A., Leblond, C. S., Gagnon, J. F., Arnulf, I., ... Rouleau, G. A. (2015). Parkinson's Disease Genetic Loci in Rapid Eye Movement Sleep Behavior Disorder. *Journal of Molecular Neuroscience : MN*, 56(3), 617–22. doi:10.1007/s12031-015-0569-7
- Gasser, T. (2015). Usefulness of Genetic Testing in PD and PD Trials: A Balanced Review. *Journal of Parkinson's Disease*, 1–7. doi:10.3233/JPD-140507
- Gasser, T., Hardy, J., & Mizuno, Y. (2011). Milestones in PD genetics. *Mov Disord*, 26(6), 1042–1048. doi:10.1002/mds.23637
- Gibson, G. (2010). Hints of hidden heritability in GWAS. *Nature Genetics*. doi:10.1038/ng0710-558
- Gibson, G. (2012). Rare and common variants: twenty arguments. *Nature Reviews Genetics*.

doi:10.1038/nrg3118

Gilks, W. P., Abou-Sleiman, P. M., Gandhi, S., Jain, S., Singleton, A., Lees, A. J., ... Wood, N. W. (2005). A common LRRK2 mutation in idiopathic Parkinson's disease. *Lancet*, 365, 415–416. doi:10.1016/S0140-6736(05)17830-1

Gloeckner, C. J., Kinkl, N., Schumacher, A., Braun, R. J., O'Neill, E., Meitinger, T., ... Ueffing, M. (2006). The Parkinson disease causing LRRK2 mutation I2020T is associated with increased kinase activity. *Human Molecular Genetics*, 15(2), 223–232. doi:10.1093/hmg/ddi439

Goker-Alpan, O., Schiffmann, R., LaMarca, M. E., Nussbaum, R. L., McInerney-Leo, A., & Sidransky, E. (2004). Parkinsonism among Gaucher disease carriers. *Journal of Medical Genetics*, 41(12), 937–40. doi:10.1136/jmg.2004.024455

Goldwurm, S., Zini, M., Mariani, L., Tesei, S., Miceli, R., Sironi, F., ... Pezzoli, G. (2007). Evaluation of LRRK2 G2019S penetrance: Relevance for genetic counseling in Parkinson disease. *Neurology*, 68(14), 1141–1143. doi:10.1212/01.wnl.0000254483.19854.ef

Grünewald, A., Kasten, M., Ziegler, A., & Klein, C. (2013). Next-generation phenotyping using the parkin example: time to catch up with genetics. *JAMA Neurology*, 70(9), 1186–91. doi:10.1001/jamaneurol.2013.488

Gudbjartsson, D. F., Helgason, H., Gudjonsson, S. a, Zink, F., Oddson, A., Gylfason, A., ... Stefansson, K. (2015). Large-scale whole-genome sequencing of the Icelandic population. *Nat Genet*, 47(5), 435–444. doi:10.1038/ng.3247

Gulcher, J., & Stefansson, K. (2000). The Icelandic Healthcare Database and informed consent. *N Engl J Med.*, 342(24), 1827–1830.

Haines, J. L. (2013). The importance of rare DNA variation in neurologic disease: cautionary tale. *Neurology*, 80(11), 974–975. doi:10.1212/WNL.0b013e318287289c

Han, F., Grimes, D. A., Li, F., Wang, T., Yu, Z., Song, N., ... Bulman, D. E. (2015). Mutations in the glucocerebrosidase gene are common in patients with Parkinson's disease from Eastern Canada. *International Journal of Neuroscience*, 126(5), 415–421.

Hardy, J., & Singleton, A. (2009). Genomewide association studies and human disease. *The New England Journal of Medicine*, 360(17), 1759–1768. doi:10.1056/NEJMra0808700

- Healy, D. G., Falchi, M., O'Sullivan, S. S., Bonifati, V., Durr, A., Bressman, S., ... Wood, N. W. (2008). Phenotype, genotype, and worldwide genetic penetrance of LRRK2-associated Parkinson's disease: a case-control study. *The Lancet Neurology*, *7*(7), 583–590. doi:10.1016/S1474-4422(08)70117-0
- Hedrich, K., Marder, K., Harris, J., Kann, M., Lynch, T., Meija-Santana, H., ... Klein, C. (2002). Evaluation of 50 probands with early-onset Parkinson's disease for Parkin mutations. *Neurology*, *58*(8), 1239–1246.
- Helgason, A., Lalueza-Fox, C., Ghosh, S., Sigurdardóttir, S., Sampietro, M. L., Gigli, E., ... Stefánsson, K. (2009). Sequences from first settlers reveal rapid evolution in Icelandic mtDNA pool. *PLoS Genetics*, *5*. doi:10.1371/journal.pgen.1000343
- Helgason, A., Nicholson, G., Stefánsson, K., & Donnelly, P. (2003). A reassessment of genetic diversity in Icelanders: Strong evidence from multiple loci for relative homogeneity caused by genetic drift. *Annals of Human Genetics*, *67*, 281–297. doi:10.1046/j.1469-1809.2003.00046.x
- Hernandez, D. G., Nalls, M. a., Ylikotila, P., Keller, M., Hardy, J. a., Majamaa, K., & Singleton, A. B. (2012). Genome wide assessment of young onset Parkinson's disease from Finland. *PLoS ONE*, *7*(7), 5–9. doi:10.1371/journal.pone.0041859
- Hilker, R., Klein, C., Ghaemi, M., Kis, B., Strotmann, T., Ozelius, L. J., ... Pramstaller, P. P. (2001). Positron emission tomographic analysis of the nigrostriatal dopaminergic system in familial parkinsonism associated with mutations in the parkin gene. *Ann Neurol*, *49*(3), 367–376.
- Hiller, A., Hagenah, J. M., Djarmati, A., Hedrich, K., Reetz, K., Schneider-Gold, C., ... Klein, C. (2007). Phenotypic spectrum of PINK1-associated parkinsonism in 15 mutation carriers from 1 family. *Mov Disord*, *22*(1), 145–147. doi:10.1002/mds.21059
- Hindorff, L. A., Sethupathy, P., Junkins, H. A., Ramos, E. M., Mehta, J. P., Collins, F. S., & Manolio, T. A. (2009). Potential etiologic and functional implications of genome-wide association loci for human diseases and traits. *Proceedings of the National Academy of Sciences of the United States of America*, *106*, 9362–9367. doi:10.1073/pnas.0903103106
- Hirschhorn, J. N., & Daly, M. J. (2005). Genome-wide association studies for common diseases and complex traits. *Nature Reviews. Genetics*, *6*, 95–108. doi:10.1038/nrg1521



- Hornykiewicz, O. (1982). Imbalance of brain monoamines and clinical disorders. *Progress in Brain Research*, 55, 419–429.
- Huo, Q., Li, T., Zhao, P., & Wang, L. (2015). Association between rs6812193 polymorphism and sporadic Parkinson's disease susceptibility. *Neurological Sciences*, 36(8), 1479–1481.
- International HapMap, C. (2005). A haplotype map of the human genome. *Nature*, 437(7063), 1299–1320. doi:10.1038/nature04226
- International, T., & Consortium, H. (2003). The International HapMap Project. *Nature*, 426(6968), 789–796. doi:10.1038/nature02168
- International, T., Sclerosis, M., & Consortium, G. (2011). Genetic risk and a primary role for cell-mediated immune mechanisms in multiple sclerosis. *Nature*, 476, 214–219. doi:10.1038/nature10251
- Ishihara, L., Warren, L., Gibson, R., Amouri, R., Lesage, S., Dürr, A., ... Hentati, F. (2006). Clinical features of Parkinson disease patients with homozygous leucine-rich repeat kinase 2 G2019S mutations. *Archives of Neurology*, 63(9), 1250–1254. doi:10.1001/archneur.63.9.1250
- Jing, C. C., Luo, X. G., Cui, H. G., Li, E. R., Li, P., Jiang, E. Z., ... Pang, H. (2015). Screening of polymorphisms located in the FGF20 and TMEM175 genes in North Chinese Parkinson's disease patients. *Genet. Mol. Res.*, 14(4), 13679 – 13687.
- Kalinderi, K., Bostantjopoulou, S., & Fidani, L. (2016). The genetic background of Parkinson's disease: current progress and future prospects. *Acta Neurologica Scandinavica*.
- Kalinderi, K., Bostantjopoulou, S., Katsarou, Z., & Fidani, L. (2013). Association study of rs6812193 polymorphism with Parkinson's disease in a Greek population. *Neuroscience Letters*, 541, 190–2. doi:10.1016/j.neulet.2013.02.048
- Kay, D. M., Moran, D., Moses, L., Poorkaj, P., Zabetian, C. P., Nutt, J., ... Payami, H. (2007). Heterozygous parkin point mutations are as common in control subjects as in Parkinson's patients. *Ann Neurol*, 61(1), 47–54. doi:10.1002/ana.21039
- Kay, D. M., Stevens, C. F., Hamza, T. H., Montimurro, J. S., Zabetian, C. P., Factor, S. A., ... Payami, H. (2010). A comprehensive analysis of deletions, multiplications, and copy

- number variations in PARK2. *Neurology*, 75(13), 1189–1194.  
doi:10.1212/WNL.0b013e3181f4d832
- Kay, D. M., Zabetian, C. P., Factor, S. A., Nut, J. G., Samii, A., Griffith, A., ... Payami, H. (2006). Parkinson's disease and LRRK2: Frequency of a common mutation in U.S. movement disorder clinics. *Movement Disorders*, 21(4), 519–523.  
doi:10.1002/mds.20751
- Keiper, B. D., Gan, W., & Rhoads, R. E. (1999). Protein synthesis initiation factor 4G. *International Journal of Biochemistry and Cell Biology*, 31, 37–41. doi:10.1016/S1357-2725(98)00130-7
- Keller, M. F., Saad, M., Bras, J., Bettella, F., Nicolaou, N., Simón-Sánchez, J., ... McCarthy, M. I. (2012). Using genome-wide complex trait analysis to quantify “missing heritability” in Parkinson's disease. *Human Molecular Genetics*, 21, 4996–5009.  
doi:10.1093/hmg/dd335
- Khan, N. L., Brooks, D. J., Pavese, N., Sweeney, M. G., Wood, N. W., Lees, A. J., & Piccini, P. (2002). Progression of nigrostriatal dysfunction in a parkin kindred: an [18F]dopa PET and clinical study. *Brain*, 125(Pt 10), 2248–2256.
- Khan, N. L., Scherfler, C., Graham, E., Bhatia, K. P., Quinn, N., Lees, A. J., ... Piccini, P. (2005). Dopaminergic dysfunction in unrelated, asymptomatic carriers of a single parkin mutation. *Neurology*, 64(1), 134–136. doi:10.1212/01.wnl.0000148725.48740.6d
- Kim, S. Y., Seong, M. W., Jeon, B. S., Ko, H. S., Kim, J. Y., & Park, S. S. (2012). Phase analysis identifies compound heterozygous deletions of the PARK2 gene in patients with early-onset Parkinson disease. *Clin Genet*, 82(1), 77–82. doi:10.1111/j.1399-0004.2011.01693.x
- Kitada, T., Asakawa, S., Hattori, N., Matsumine, H., Yamamura, Y., Minoshima, S., ... Shimizu, N. Mutations in the parkin gene cause autosomal recessive juvenile parkinsonism., 392 Nature 605–608 (1998). doi:10.1038/33416
- Klein, C., Djarmati, A., Hedrich, K., Schafer, N., Scaglione, C., Marchese, R., ... Pramstaller, P. P. (2005). PINK1, Parkin, and DJ-1 mutations in Italian patients with early-onset parkinsonism. *Eur J Hum Genet*, 13(9), 1086–1093. doi:10.1038/sj.ejhg.5201455
- Klein, C., Lohmann-Hedrich, K., Rogaeva, E., Schlossmacher, M. G., & Lang, A. E. (2007). Deciphering the role of heterozygous mutations in genes associated with parkinsonism.

- Lancet Neurol*, 6(7), 652–662. doi:10.1016/s1474-4422(07)70174-6
- Klein, C., Pramstaller, P. P., Kis, B., Page, C. C., Kann, M., Leung, J., ... Ozelius, L. J. (2000). Parkin deletions in a family with adult-onset, tremor-dominant parkinsonism: expanding the phenotype. *Ann Neurol*, 48(1), 65–71.
- Kong, A., Masson, G., Frigge, M. L., Gylfason, A., Zusmanovich, P., Thorleifsson, G., ... Stefansson, K. (2008). Detection of sharing by descent, long-range phasing and haplotype imputation. *Nat Genet*, 40(9), 1068–1075. doi:10.1038/ng.216
- Kong, A., Steinthorsdottir, V., Masson, G., Thorleifsson, G., Sulem, P., Besenbacher, S., ... Stefansson, K. (2009). Parental origin of sequence variants associated with complex diseases. *Nature*, 462(7275), 868–874. doi:10.1038/nature08625
- Kristiansson, K., Naukkarinen, J., & Peltonen, L. (2008). Isolated populations and complex disease gene identification. *Genome Biology*, 9(8), 109. doi:10.1186/gb-2008-9-8-109
- Lander, E., & Kruglyak, L. (1995). Genetic dissection of complex traits: guidelines for interpreting and reporting linkage results. *Nature Genetics*, 11(3), 241–247. doi:10.1038/ng1195-241
- Lander, E. S., Linton, L. M., Birren, B., Nusbaum, C., Zody, M. C., Baldwin, J., ... Chen, Y. J. (2001). Initial sequencing and analysis of the human genome. *Nature*, 409(6822), 860–921. doi:10.1038/35057062
- LeFebvre, A. K., Korneeva, N. L., Trutschl, M., Cvek, U., Duzan, R. D., Bradley, C. A., ... Rhoads, R. E. (2006). Translation initiation factor eIF4G-1 binds to eIF3 through the eIF3e subunit. *J Biol Chem*, 281(32), 22917–22932. doi:10.1074/jbc.M605418200
- Lesage, S., & Brice, A. (2012). Role of mendelian genes in “sporadic” Parkinson’s disease. *Parkinsonism Relat Disord*, 18 Suppl 1, S66–70. doi:10.1016/s1353-8020(11)70022-0
- Lesage, S., Condroyer, C., Klebe, S., Lohmann, E., Durif, F., Damier, P., ... Brice, A. (2012). EIF4G1 in familial Parkinson’s disease: pathogenic mutations or rare benign variants? *Neurobiol Aging*, 33(9), 2233.e1–2233.e5. doi:10.1016/j.neurobiolaging.2012.05.006
- Lesage, S., Ibanez, P., Lohmann, E., Pollak, P., Tison, F., Tazir, M., ... Brice, A. (2005). G2019S LRRK2 mutation in French and North African families with Parkinson’s disease. *Annals of Neurology*, 58(5), 784–787. doi:10.1002/ana.20636
- Lesage, S., Lohmann, E., Tison, F., Durif, F., Durr, A., & Brice, A. (2008). Rare

- heterozygous parkin variants in French early-onset Parkinson disease patients and controls. *J Med Genet*, 45(1), 43–46. doi:10.1136/jmg.2007.051854
- Li, H., Teo, Y., & Tan, E. (2015). Patterns of linkage disequilibrium at PARK16 may explain variances in genetic association studies. *Movement Disorders*, 30(10), 1335–1342(8).
- Li, K., Tang, B. S., Guo, J. F., Lou, M. X., Lv, Z. Y., Liu, Z. H., ... Yan, X. X. (2013). Analysis of EIF4G1 in ethnic Chinese. *BMC Neurol*, 13, 38. doi:10.1186/1471-2377-13-38
- Li, N.-N., Mao, X.-Y., Chang, X.-L., Zhao, D.-M., Zhang, J.-H., Liao, Q., ... Peng, R. (2013). SNCA rs356219 variant increases risk of sporadic Parkinson's disease in ethnic Chinese. *American Journal of Medical Genetics. Part B, Neuropsychiatric Genetics : The Official Publication of the International Society of Psychiatric Genetics*, 162B(5), 452–6. doi:10.1002/ajmg.b.32143
- Li, N.-N., Tan, E.-K., Chang, X.-L., Mao, X.-Y., Zhang, J.-H., Zhao, D.-M., ... Peng, R. (2013). Genetic association study between STK39 and CCDC62/HIP1R and Parkinson's disease. *PLoS One*, 8(11), e79211.
- Lill, C. M., Roehr, J. T., McQueen, M. B., Kavvoura, F. K., Bagade, S., Schjeide, B. M. M., ... Bertram, L. (2012). Comprehensive research synopsis and systematic meta-analyses in Parkinson's disease genetics: The PDgene database. *PLoS Genetics*, 8(3). doi:10.1371/journal.pgen.1002548
- Lim, K. L., Chew, K. C. M., Tan, J. M. M., Wang, C., Chung, K. K. K., Zhang, Y., ... Dawson, T. M. (2005). Parkin mediates nonclassical, proteasomal-independent ubiquitination of synphilin-1: implications for Lewy body formation. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience*, 25, 2002–2009. doi:10.1523/JNEUROSCI.4474-04.2005
- Lincoln, S. J., Maraganore, D. M., Lesnick, T. G., Bounds, R., de Andrade, M., Bower, J. H., ... Farrer, M. J. (2003). Parkin variants in North American Parkinson's disease: cases and controls. *Mov Disord*, 18(11), 1306–1311. doi:10.1002/mds.10601
- Liu, J., Xiao, Q., Wang, Y., Xu, Z. M., Wang, Y., Yang, Q., ... Chen, S. Di. (2013). Analysis of genome-wide association study-linked loci in parkinson's disease of mainland china. *Movement Disorders*, 28(13), 1892–1895. doi:10.1002/mds.25599
- Lucking, C. B., Chesneau, V., Lohmann, E., Verpillat, P., Dulac, C., Bonnet, A. M., ... Brice,

- A. (2003). Coding polymorphisms in the parkin gene and susceptibility to Parkinson disease. *Arch Neurol*, *60*(9), 1253–1256. doi:10.1001/archneur.60.9.1253
- Lucking, C. B., Durr, A., Bonifati, V., Vaughan, J., De Michele, G., Gasser, T., ... Brice, A. (2000). Association between early-onset Parkinson's disease and mutations in the parkin gene. *N Engl J Med*, *342*(21), 1560–1567. doi:10.1056/nejm200005253422103
- MacArthur, D. G., Manolio, T. A., Dimmock, D. P., Rehm, H. L., Shendure, J., Abecasis, G. R., ... Gunter, C. (2014). Guidelines for investigating causality of sequence variants in human disease. *Nature*, *508*(7497), 469–76. doi:10.1038/nature13127
- Manolio, T. a, & Collins, F. S. (2009). The HapMap and genome-wide association studies in diagnosis and therapy. *Annual Review of Medicine*, *60*, 443–456. doi:10.1146/annurev.med.60.061907.093117
- Manolio, T. A., Brooks, L. D., & Collins, F. S. (2008). A HapMap harvest of insights into the genetics of common disease. *Journal of Clinical Investigation*. doi:10.1172/JCI34772
- Manolio, T. A., Collins, F. S., Cox, N. J., Goldstein, D. B., Hindorff, L. A., Hunter, D. J., ... Visscher, P. M. (2009). Finding the missing heritability of complex diseases. *Nature*, *461*, 747–753. doi:10.1038/nature08494
- Mantel, N., & Haenszel, W. (1959). Statistical aspects of the analysis of data from retrospective studies of disease. *J Natl Cancer Inst*, *22*((4)), 719–48.
- Marder, K. S., Tang, M. X., Mejia-Santana, H., Rosado, L., Louis, E. D., Comella, C. L., ... Clark, L. N. (2010). Predictors of parkin mutations in early-onset Parkinson disease: the consortium on risk for early-onset Parkinson disease study. *Arch Neurol*, *67*(6), 731–738. doi:10.1001/archneurol.2010.95
- Marian, A. J. (2012). Elements of “missing heritability.” *Current Opinion in Cardiology*. doi:10.1097/HCO.0b013e328352707d
- Mata, I. F., Leverenz, J. B., Weintraub, D., Trojanowski, J. Q., Hurtig, H. I., Van Deerlin, V. M., ... Zabetian, C. P. (2014). APOE, MAPT, and SNCA genes and cognitive performance in Parkinson disease. *JAMA Neurology*, *71*(11), 1405–12. doi:10.1001/jamaneurol.2014.1455
- Mata, I. F., Taylor, J. P., Kachergus, J., Hulihan, M., Huerta, C., Lahoz, C., ... Alvarez, V. (2005). LRRK2 R1441G in Spanish patients with Parkinson's disease. *Neuroscience*

*Letters*, 382, 309–311. doi:10.1016/j.neulet.2005.03.033

- McCarthy, M. I., Abecasis, G. R., Cardon, L. R., Goldstein, D. B., Little, J., Ioannidis, J. P. A., & Hirschhorn, J. N. (2008). Genome-wide association studies for complex traits: consensus, uncertainty and challenges. *Nature Reviews. Genetics*, 9(5), 356–369. doi:10.1038/nrg2344
- Miyake, Y., Tanaka, K., Fukushima, W., Kiyohara, C., Sasaki, S., Tsuboi, Y., ... Nagai, M. (2012). Lack of association between BST1 polymorphisms and sporadic Parkinson's disease in a Japanese population. *Journal of the Neurological Sciences*, 323(1-2), 162–166. doi:10.1016/j.jns.2012.09.008
- Myers, S., Bottolo, L., Freeman, C., McVean, G., & Donnelly, P. (2005). A fine-scale map of recombination rates and hotspots across the human genome. *Science (New York, N.Y.)*, 310, 321–324. doi:10.1126/science.1117196
- Nalls, M. a, Pankratz, N., Lill, C. M., Do, C. B., Hernandez, D. G., Saad, M., ... Singleton, A. B. (2014). Large-scale meta-analysis of genome-wide association data identifies six new risk loci for Parkinson's disease. *Nature Genetics*, 056, 1–7. doi:10.1038/ng.3043
- Nalls, M. A., Plagnol, V., Hernandez, D. G., Sharma, M., Sheerin, U.-M., Saad, M., ... Wood, N. W. (2011). Imputation of sequence variants for identification of genetic risks for Parkinson's disease: a meta-analysis of genome-wide association studies. *Lancet*, 377(9766), 641–9. doi:10.1016/S0140-6736(10)62345-8
- Nichols, W. C., Pankratz, N., Hernandez, D., Paisán-Ruíz, C., Jain, S., Halter, C. A., ... Foroud, T. (2005). Genetic screening for a single common LRRK2 mutation in familial Parkinson's disease. *Lancet*, 365(9457), 410–412. doi:10.1016/S0140-6736(05)17828-3
- Nuytemans, K., Bademci, G., Inchausti, V., Dressen, A., Kinnamon, D. D., Mehta, A., ... Vance, J. M. (2013). Whole exome sequencing of rare variants in EIF4G1 and VPS35 in Parkinson disease. *Neurology*, 80(11), 982–989. doi:10.1212/WNL.0b013e31828727d4
- Nuytemans, K., Meeus, B., Crosiers, D., Brouwers, N., Goossens, D., Engelborghs, S., ... Theuns, J. (2009). Relative contribution of simple mutations vs. copy number variations in five Parkinson disease genes in the Belgian population. *Hum Mutat*, 30(7), 1054–1061. doi:10.1002/humu.21007
- Oliveira, S. A., Scott, W. K., Martin, E. R., Nance, M. A., Watts, R. L., Hubble, J. P., ... Vance, J. M. (2003). Parkin mutations and susceptibility alleles in late-onset Parkinson's

- disease. *Ann Neurol*, 53(5), 624–629. doi:10.1002/ana.10524
- Ozelius, L. J., Senthil, G., Saunders-Pullman, R., Ohmann, E., Deligtisch, A., Tagliati, M., ... Bressman, S. B. (2006). LRRK2 G2019S as a cause of Parkinson's disease in Ashkenazi Jews. *The New England Journal of Medicine*. doi:10.1056/NEJMc055509
- Pan, F., Dong, H., Ding, H., Ye, M., Liu, W., Wu, Y., ... Ding, X. (2012). SNP rs356219 of the  $\alpha$ -synuclein (SNCA) gene is associated with Parkinson's disease in a Chinese Han population. *Parkinsonism & Related Disorders*, 18(5), 632–4. doi:10.1016/j.parkreldis.2012.01.025
- Pankratz, N., Dumitriu, A., Hetrick, K. N., Sun, M., Latourelle, J. C., Wilk, J. B., ... DeStefano, A. L. (2011). Copy number variation in familial Parkinson disease. *PLoS One*, 6(8), e20988. doi:10.1371/journal.pone.0020988
- Pankratz, N., Kissell, D. K., Pauciulo, M. W., Halter, C. A., Rudolph, A., Pfeiffer, R. F., ... Nichols, W. C. (2009). Parkin dosage mutations have greater pathogenicity in familial PD than simple sequence mutations. *Neurology*, 73(4), 279–286. doi:10.1212/WNL.0b013e3181af7a33
- Papapetropoulos, S., Adi, N., Shehadeh, L., Bishopric, N., Singer, C., Argyriou, A. A., & Chroni, E. (2008). Is the G2019S LRRK2 mutation common in all southern European populations? *Journal of Clinical Neuroscience*, 15(9), 1027–1030. doi:10.1016/j.jocn.2007.08.013
- Parkinson, I., Consortium, G., Trust, W., & Control, C. (2011). A two-stage meta-analysis identifies several new loci for Parkinson's disease. *PLoS Genetics*, 7(6), e1002142. doi:10.1371/journal.pgen.1002142
- Parkinson, J. (2002). An essay on the shaking palsy. 1817. *The Journal of Neuropsychiatry and Clinical Neurosciences*, 14, 223–236; discussion 222. doi:10.1176/appi.neuropsych.14.2.223
- Pe'er, I., Yelensky, R., Altshuler, D., & Daly, M. J. (2008). Estimation of the multiple testing burden for genomewide association studies of nearly all common variants. *Genetic Epidemiology*, 32(4), 381–385. doi:10.1002/gepi.20303
- Peltonen, L., Palotie, a., & Lange, K. (2000). Use of population isolates for mapping complex traits. *Nature Reviews. Genetics*, 1(3), 182–90. doi:10.1038/35042049

- Periquet, M., Lucking, C., Vaughan, J., Bonifati, V., Durr, A., De Michele, G., ... Brice, A. (2001). Origin of the mutations in the parkin gene in Europe: exon rearrangements are independent recurrent events, whereas point mutations may result from Founder effects. *Am J Hum Genet*, *68*(3), 617–626.
- Pfretter, R., Wszolek, Z. K., & Ebadi, M. (2012). *Parkinson's Disease* (Second Edi.). CRC Press Taylor & Francis Group.
- Pramstaller, P. P., Kis, B., Eskelson, C., Hedrich, K., Scherer, M., Schwinger, E., ... Klein, C. (2002). Phenotypic variability in a large kindred (Family LA) with deletions in the parkin gene. *Mov Disord*, *17*(2), 424–426.
- Prévôt, D., Darlix, J. L., & Ohlmann, T. (2003). Conducting the initiation of protein synthesis: The role of eIF4G. *Biology of the Cell*. doi:10.1016/S0248-4900(03)00031-5
- Prévôt, D., Décimo, D., Herbreteau, C. H., Roux, F., Garin, J., Darlix, J. L., & Ohlmann, T. (2003). Characterization of a novel RNA-binding region of eIF4GI critical for ribosomal scanning. *EMBO Journal*, *22*, 1909–1921. doi:10.1093/emboj/cdg175
- Pritchard, J. K. (2001). Are rare variants responsible for susceptibility to complex diseases? *American Journal of Human Genetics*, *69*, 124–137. doi:10.1086/321272
- Psychiatric, T., Consortium, G., & Committee, S. (2009). A framework for interpreting genome-wide association studies of psychiatric disorders. *Molecular Psychiatry*, *14*(1), 10–17. doi:10.1038/mp.2008.126
- Ramirez-Valle, F., Braunstein, S., Zavadil, J., Formenti, S. C., & Schneider, R. J. (2008). eIF4GI links nutrient sensing by mTOR to cell proliferation and inhibition of autophagy. *J Cell Biol*, *181*(2), 293–307. doi:10.1083/jcb.200710215
- Ramírez-Valle, F., Braunstein, S., Zavadil, J., Formenti, S. C., & Schneider, R. J. (2008). eIF4GI links nutrient sensing by mTOR to cell proliferation and inhibition of autophagy. *Journal of Cell Biology*, *181*, 293–307. doi:10.1083/jcb.200710215
- Riess, O., Kuhn, W., & Kruger, R. (2000). Genetic influence on the development of Parkinson's disease. *J Neurol.*, *247 Suppl* , II69–II74.
- Ripke, S., Neale, B. M., Corvin, A., Walters, J. T. R., Farh, K.-H., Holmans, P. a., ... O'Donovan, M. C. (2014). Biological insights from 108 schizophrenia-associated genetic loci. *Nature*, *511*, 421–427. doi:10.1038/nature13595



- Risch, N., & Merikangas, K. (1996). The future of genetic studies of complex human diseases. *Science (New York, N.Y.)*, *273*, 1516–1517. doi:doi:10.1126/science.273.5281.1516
- Ross, O. A., Toft, M., Whittle, A. J., Johnson, J. L., Papapetropoulos, S., Mash, D. C., ... Dickson, D. W. (2006). Lrrk2 and Lewy body disease. *Annals of Neurology*, *59*(2), 388–393. doi:10.1002/ana.20731
- Ruffmann, C., Zini, M., Goldwurm, S., Bramerio, M., Spinello, S., Rusconi, D., ... Giaccone, G. (2012). Lewy body pathology and typical Parkinson disease in a patient with a heterozygous (R275W) mutation in the Parkin gene (PARK2). *Acta Neuropathol*, *123*(6), 901–903. doi:10.1007/s00401-012-0991-7
- Scholz, S. W., Houlden, H., Schulte, C., Sharma, M., Li, A., Berg, D., ... Gasser, T. (2009). SNCA variants are associated with increased risk for multiple system atrophy. *Annals of Neurology*, *65*(5), 610–614. doi:10.1002/ana.21685
- Schork, N. J., Murray, S. S., Frazer, K. A., & Topol, E. J. (2009). Common vs. rare allele hypotheses for complex diseases. *Current Opinion in Genetics and Development*. doi:10.1016/j.gde.2009.04.010
- Schulte, E. C., Mollenhauer, B., Zimprich, A., Bereznai, B., Lichtner, P., Haubenberger, D., ... Winkelmann, J. (2012). Variants in eukaryotic translation initiation factor 4G1 in sporadic Parkinson's disease. *Neurogenetics*, *13*(3), 281–285. doi:10.1007/s10048-012-0334-9
- Sellbach, A. N., Boyle, R. S., Silburn, P. A., & Mellick, G. D. (2006). Parkinson's disease and family history. *Parkinsonism.Relat Disord.*, *12*, 399–409.
- Service, S., DeYoung, J., Karayiorgou, M., Roos, J. L., Pretorius, H., Bedoya, G., ... Freimer, N. (2006). Magnitude and distribution of linkage disequilibrium in population isolates and implications for genome-wide association studies. *Nature Genetics*, *38*, 556–560. doi:10.1038/ng1770
- Sham, P. C., & Purcell, S. M. (2014). Statistical power and significance testing in large-scale genetic studies. *Nature Reviews. Genetics*, *15*(5), 335–46. doi:10.1038/nrg3706
- Sharma, M., Ioannidis, J. P. a, Aasly, J. O., Annesi, G., Brice, A., Van Broeckhoven, C., ... Krüger, R. (2012). Large-scale replication and heterogeneity in Parkinson disease genetic loci. *Neurology*, *79*(7), 659–667. doi:10.1212/WNL.0b013e318264e353

- Shimura, H., Hattori, N., Kubo, S., Mizuno, Y., Asakawa, S., Minoshima, S., ... Suzuki, T. (2000). Familial Parkinson disease gene product, parkin, is a ubiquitin-protein ligase. *Nat Genet*, *25*(3), 302–305. doi:10.1038/77060
- Shulman, J. M., Yu, L., Buchman, A. S., Evans, D. a, Schneider, J. a, Bennett, D. a, & De Jager, P. L. (2014). Association of Parkinson disease risk loci with mild parkinsonian signs in older persons. *JAMA Neurology*, *71*(4), 429–35. doi:10.1001/jamaneurol.2013.6222
- Sidransky, E. (2006). Heterozygosity for a Mendelian disorder as a risk factor for complex disease. *Clin Genet*, *70*(4), 275–282. doi:10.1111/j.1399-0004.2006.00688.x
- Sidransky, E., Nalls, M. A., Aasly, J. O., Aharon-Peretz, J., Annesi, G., Barbosa, E. R., ... Ziegler, S. G. (2009). Multicenter analysis of glucocerebrosidase mutations in Parkinson's disease. *The New England Journal of Medicine*, *361*, 1651–1661. doi:10.1056/NEJMoa0901281
- Siitonen, A., Majounie, E., Federoff, M., Ding, J., Majamaa, K., & Singleton, A. B. (2013). Mutations in EIF4G1 are not a common cause of Parkinson's disease. *Eur J Neurol*, *20*(4), e59. doi:10.1111/ene.12051
- Singleton, A., & Hardy, J. (2011). A generalizable hypothesis for the genetic architecture of disease: Pleomorphic risk loci. *Human Molecular Genetics*, *20*, 158–162. doi:10.1093/hmg/ddr358
- Singleton, A., Myers, A., & Hardy, J. (2004). The law of mass action applied to neurodegenerative disease: a hypothesis concerning the etiology and pathogenesis of complex diseases. *Human Molecular Genetics*, *13 Spec No*, R123–R126. doi:10.1093/hmg/ddh093
- Smith, K. P. (1995). Landnam: The Settlement of Iceland in Archaeological and Historical Perspective. *World Archaeology*, *26*(3), 319–347.
- Song, Y. S., Wang, F., & Slatkin, M. (2010). General epistatic models of the risk of complex diseases. *Genetics*, *186*(4), 1467–1473. doi:10.1534/genetics.110.119008
- Spencer, C. C. A., Su, Z., Donnelly, P., & Marchini, J. (2009). Designing genome-wide association studies: Sample size, power, imputation, and the choice of genotyping chip. *PLoS Genetics*, *5*(5). doi:10.1371/journal.pgen.1000477

- Spillantini, M. G., Crowther, R. A., Jakes, R., Hasegawa, M., & Goedert, M. (1998). alpha-Synuclein in filamentous inclusions of Lewy bodies from Parkinson's disease and dementia with lewy bodies. *Proceedings of the National Academy of Sciences of the United States of America*, *95*(11), 6469–6473. doi:10.1073/Pnas.95.11.6469
- Spillantini, M. G., Schmidt, M. L., Lee, V. M., Trojanowski, J. Q., Jakes, R., & Goedert, M. (1997). Alpha-synuclein in Lewy bodies. *Nature*. doi:10.1038/42166
- Strachan, T., & P. Read, A. (2005). Molekulare Humangenetik Zitat.pdf. In *Molekulare Humangenetik* (p. 681; 730–747).
- Sudhaman, S., Behari, M., Govindappa, S. T., Muthane, U. B., Juyal, R. C., & Thelma, B. K. (2013). VPS35 and EIF4G1 mutations are rare in Parkinson's disease among Indians. *Neurobiol Aging*, *34*(10), 2442.e1–3. doi:10.1016/j.neurobiolaging.2013.04.025
- Sun, M., Latourelle, J. C., Wooten, G. F., Lew, M. F., Klein, C., Shill, H. A., ... Gusella, J. F. (2006). Influence of heterozygosity for parkin mutation on onset age in familial Parkinson disease: the GenePD study. *Arch Neurol*, *63*(6), 826–832. doi:10.1001/archneur.63.6.826
- Tan, E. K. (2006). Identification of a common genetic risk variant (LRRK2 Gly2385Arg) in Parkinson's Disease. *Annals of the Academy of Medicine Singapore*.
- Taylor, J. P., Mata, I. F., & Farrer, M. J. (2006). LRRK2: A common pathway for parkinsonism, pathogenesis and prevention? *Trends in Molecular Medicine*. doi:10.1016/j.molmed.2005.12.004
- Thorisson, G. A., & Stein, L. D. (2003). The SNP consortium website: Past, present and future. *Nucleic Acids Research*. doi:10.1093/nar/gkg052
- Tucci, A., Charlesworth, G., Sheerin, U. M., Plagnol, V., Wood, N. W., & Hardy, J. (2012). Study of the genetic variability in a Parkinson's Disease gene: EIF4G1. *Neurosci Lett*, *518*(1), 19–22. doi:10.1016/j.neulet.2012.04.033
- Ubl, A., Berg, D., Holzmann, C., Krüger, R., Berger, K., Arzberger, T., ... Riess, O. (2002). 14-3-3 protein is a component of Lewy bodies in Parkinson's disease-mutation analysis and association studies of 14-3-3 eta. *Brain Research. Molecular Brain Research*, *108*, 33–39. doi:10.1016/S0169-328X(02)00510-7
- Venter, J. C., Adams, M. D., Myers, E. W., Li, P. W., Mural, R. J., Sutton, G. G., ... Zhu, X.

- (2001). The sequence of the human genome. *Science (New York, N.Y.)*, 291(5507), 1304–1351. doi:10.1126/science.1058040
- Visscher, P. M., Brown, M. A., McCarthy, M. I., & Yang, J. (2012). Five years of GWAS discovery. *American Journal of Human Genetics*. doi:10.1016/j.ajhg.2011.11.029
- von Coelln, R., Dawson, V. L., & Dawson, T. M. (2004). Parkin-associated Parkinson's disease. *Cell Tissue Res*, 318(1), 175–184. doi:10.1007/s00441-004-0924-4
- Wain, H. M., Lush, M. J., Ducluzeau, F., Khodiyar, V. K., & Povey, S. (2004). Genew: the Human Gene Nomenclature Database, 2004 updates. *Nucleic Acids Research*, 32(Database issue), D255–D257. doi:10.1093/nar/gkh072
- Wakabayashi, K., Engelender, S., Yoshimoto, M., Tsuji, S., Ross, C. A., & Takahashi, H. (2000). Synphilin-1 is present in lewy bodies in Parkinson's disease. *Annals of Neurology*, 47, 521–523. doi:10.1002/1531-8249(200004)47:4<521::AID-ANA18>3.0.CO;2-B
- Walsh, T., McClellan, J. M., McCarthy, S. E., Addington, A. M., Pierce, S. B., Cooper, G. M., ... Sebat, J. (2008). Rare structural variants disrupt multiple genes in neurodevelopmental pathways in schizophrenia. *Science (New York, N.Y.)*, 320, 539–543. doi:10.1126/science.1155174
- Wang, K., Li, M., Hadley, D., Liu, R., Glessner, J., Grant, S. F., ... Bucan, M. (2007). PennCNV: an integrated hidden Markov model designed for high-resolution copy number variation detection in whole-genome SNP genotyping data. *Genome Res*, 17(11), 1665–1674. doi:10.1101/gr.6861907
- Wang, L., Nuytemans, K., Bademci, G., Jauregui, C., Martin, E. R., Scott, W. K., ... Zuchner, S. (2013). High-resolution survey in familial Parkinson disease genes reveals multiple independent copy number variation events in PARK2. *Hum Mutat*, 34(8), 1071–1074. doi:10.1002/humu.22344
- Wang, M., Hattori, N., Matsumine, H., Kobayashi, T., Yoshino, H., Morioka, A., ... Mizuno, Y. (1999). Polymorphism in the parkin gene in sporadic Parkinson's disease. *Ann Neurol*, 45(5), 655–658.
- Wang, S., Xu, Y.-F., Ding, X.-Y., Liu, Z.-R., Ding, Y., Jin, B., ... Ding, M.-P. (2015). Association between bone marrow stromal cell antigen 1 gene polymorphisms and the susceptibility to Parkinson's disease: a meta-analysis. *Neuroscience Letters*, 599, 120–

124.

- Wei, Z., Wang, W., Bradfield, J., Li, J., Cardinale, C., Frackelton, E., ... Hakonarson, H. (2013). Large sample size, wide variant spectrum, and advanced machine-learning technique boost risk prediction for inflammatory bowel disease. *American Journal of Human Genetics*, *92*(6), 1008–1012. doi:10.1016/j.ajhg.2013.05.002
- West, A. B., Moore, D. J., Biskup, S., Bugayenko, A., Smith, W. W., Ross, C. A., ... Dawson, T. M. (2005). Parkinson's disease-associated mutations in leucine-rich repeat kinase 2 augment kinase activity. *Proceedings of the National Academy of Sciences of the United States of America*, *102*(46), 16842–16847. doi:10.1073/pnas.0507360102
- West, A., Periquet, M., Lincoln, S., Lucking, C. B., Nicholl, D., Bonifati, V., ... Farrer, M. (2002). Complex relationship between Parkin mutations and Parkinson disease. *Am J Med Genet*, *114*(5), 584–591. doi:10.1002/ajmg.10525
- Willems, S. M., Mihaescu, R., Sijbrands, E. J. G., Van Duijn, C. M., & Janssens, a. C. J. W. (2011). A methodological perspective on genetic risk prediction studies in type 2 diabetes: Recommendations for future research. *Current Diabetes Reports*, *11*(6), 511–518. doi:10.1007/s11892-011-0235-6
- Wirdefeldt, K., Adami, H.-O., Cole, P., Trichopoulos, D., & Mandel, J. (2011). Epidemiology and etiology of Parkinson's disease: a review of the evidence. *European Journal of Epidemiology*, *26 Suppl 1*, S1–S58. doi:10.1007/s10654-011-9581-6
- Wurtman, R. J. (2013). Personalized medicine strategies for managing patients with Parkinsonism and cognitive deficits. *Metabolism: Clinical and Experimental*, *62*(SUPPL.1). doi:10.1016/j.metabol.2012.08.025
- Yu, R., Guo, J., Wang, Y., Liu, Z., Sun, Z., Su, L., ... Tang, B. (2015). The single nucleotide polymorphism Rs12817488 is associated with Parkinson's disease in the Chinese population. *Journal of Clinical Neuroscience*, *22*(6), 1002–1004.
- Zabetian, C. P., Samii, A., Mosley, A. D., Roberts, J. W., Leis, B. C., Yearout, D., ... Griffith, A. (2005). A clinic-based study of the LRRK2 gene in Parkinson disease yields new mutations. *Neurology*, *65*(5), 741–744. doi:10.1212/01.WNL.0000172630.22804.73
- Zhao, Y., Ho, P., Prakash, K. M., Foo, J. N., Liu, J. J., Au, W. L., ... Tan, E. K. (2013). Analysis of EIF4G1 in Parkinson's disease among Asians. *Neurobiol Aging*, *34*(4), 1311.e5–6. doi:10.1016/j.neurobiolaging.2012.09.003

- Zhao, Y., Ho, P., Yih, Y., Chen, C., Lee, W. L., & Tan, E. K. (2009). LRRK2 variant associated with Alzheimer's disease. *Neurobiol Aging*. doi:S0197-4580(09)00389-3 [pii]r10.1016/j.neurobiolaging.2009.11.019
- Zhou, C., Huang, Y., & Przedborski, S. (2008). Oxidative stress in Parkinson's disease: a mechanism of pathogenic and therapeutic significance. *Ann N Y Acad Sci*, 1147, 93–104. doi:10.1196/annals.1427.023
- Ziegler, A., & Inke, K. (2010). *A Statistical Approach to Genetic Epidemiology* (Second Edi.). Wiley-Blackwell.
- Zimprich, A., Biskup, S., Leitner, P., Lichtner, P., Farrer, M., Lincoln, S., ... Gasser, T. (2004). Mutations in LRRK2 cause autosomal-dominant parkinsonism with pleomorphic pathology. *Neuron*, 44, 601–607. doi:10.1016/j.neuron.2004.11.005
- Zondervan, K. T., & Cardon, L. R. (2004). The complex interplay among factors that influence allelic association. *Nature Reviews. Genetics*, 5, 89–100. doi:10.1038/nrg1314

## 9 Verzeichnisse

### 9.1 Abkürzungsverzeichnis

$\alpha$	alpha
Abb	Abbildung
ACMSD	<i>aminocarboxymuconate semialdehyde decarboxylase</i>
AD	autosomal-dominant
AR	autosomal-rezessiv
$\beta$	beta
bp	Basenpaar(e)
BST1	<i>bone marrow stromal cell antigen 1</i>
bzw.	beziehungsweise
ca.	circa
CCDC62	<i>coiled-coil domain-containing protein 62</i>
CDCV-	<i>Common Disease, Common Variant - Hypothesis</i> / häufige Erkrankungen, häufige Varianten – Hypothese
CI	<i>confidence interval</i>
cM	centiMorgans
CNV	<i>copy number variation</i> / Kopienzahlvariation
COR	<i>c-terminal of Roc</i>
dbSNP	<i>single nucleotide polymorphism database</i> / Datenbank für Einzelnukleotidpolymorphismen
del	Deletion
DGKQ	<i>diacylglycerol kinase</i>
d.h.	das heißt

DNA	<i>deoxyribonucleic acid</i> / Desoxyribonukleinsäure
dupl	Duplikation
et al.	<i>et alii</i> (und andere)
EIF4G1	<i>eukaryotic translation initiation factor 4-<math>\gamma</math> 1</i> / Eukaryotischer Initiationsfaktor 4- $\gamma$ 1
FAM47E	<i>family with sequence similarity 47 member E</i>
FGF20	<i>fibroblast growth factor 20</i>
fs	<i>frameshift</i> / Leserasterverschiebung
GA	<i>Genome Analyzer</i>
GAK	<i>cyclin G-associated kinase</i>
GBA	<i>glucosidase, beta, acid</i>
GPNMB	<i>glycoprotein NMB</i>
GWAS	<i>genom-wide association study</i> / Genomweite Assoziationsstudie
HIP1R	<i>huntingtin-interacting protein 1-related protein</i>
HRM	<i>high resolution melting</i> / hochauflösende Schmelzkurve
HT	Haplotyp
IBD	<i>identical by descent</i> / abstammungsidentisch
IBS	<i>identical by state</i> / eigenschaftidentisch
IBR	<i>in-between</i> RING Domäne
InDel	Insertion und Deletion
IPS	idiopathisches Parkinson-Syndrom
ITGA8	<i>integrin, alpha-8</i>
LD	<i>linkage disequilibrium</i> / Kopplungsungleichgewicht
<sub>L</sub> DOPA	Levodopa / <sub>L</sub> -3,4-Dihydroxyphenylalanin
LRP	<i>long-range phasing</i>



LRRK2	<i>leucin-rich repeat kinase 2</i>
MAF	<i>minor allele frequency</i> / Minor Allel-Frequenz
MAPKKK	<i>mitogen-activated protein kinase kinase kinase</i>
MAPT	<i>microtubule associated protein tau</i>
MB	Megabase
MCCC1	<i>3-methylcrotonyl-CoA carboxylase 1</i>
MMP16	<i>matrix metalloproteinase 16</i>
MNK1	<i>mitogen-activated protein kinase signal integrating kinase</i>
MPTP	<i>1-Methyl-4-phenyl-1,2,3,6-tetrahydropyridin</i>
MRV-	<i>Common Disease, Multiple Rare Variant - Hypothesis</i> / häufige Krankheiten, seltene Varianten - Hypothese
MSA	Multisystematrophie
NGS	<i>Next-Generation-Sequencing</i> / Sequenziertechnologie der nächsten Generation
NMD3	<i>NMD3 ribosome export adaptor</i>
OMIM	<i>Online Inheritance In Man</i> / Datenbank der Gene, die, wenn verändert, mit einer Krankheit assoziiert sind
OR	<i>odds ratio</i> / Quotenverhältnis
p	kurzer Arm eines Chromosoms
PABP	<i>poly(A)-binding protein</i>
PCR	<i>polymerase chain reaction</i> / Polymerasekettenreaktion
PD	<i>Parkinson's disease</i> / Parkinson Erkrankung
PM	Punktmutation
PRL	<i>pleomorphic risc loci</i> / Pleomorphe Risiko Bereiche
PS	Parkinson-Syndrom
PSP	progressive supranukleäre Blickparese

q	langer Arm eines Chromosoms
RAB7L1	<i>rab7-like 1</i>
RAI1	<i>retinoic acid-induced gene 1</i>
RING	<i>really Interesting New Gene-Box</i>
RIT2	<i>ric-like protein without CAAX motif 2</i>
ROC	<i>ras of complex proteins</i>
RR	Relatives Risiko
SCARB2	<i>scavenger receptor class B, member 2</i>
SLC41A1	<i>solute carrier family 41, member 1</i>
SNCA	<i>alpha-synuclein</i>
SNP	<i>single nucleotide polymorphism</i> / Einzelnukleotidpolymorphismus
SNpc	<i>Substantia nigra pars compacta</i>
SREBF1	<i>sterol regulatory element-binding transcription factor 1</i>
STBD1	<i>starch-binding domain-containing protein 1</i>
STK39	<i>serine/threonine protein kinase 39</i>
STX1B	<i>syntaxin 1B</i>
SYT4	<i>synaptotagmin 4</i>
<i>tagSNPs</i>	<i>tagging SNPs</i> / informative Marker
Tripl	Triplikation
UBL	<i>Ubiquitin-like-Domäne</i>
vs	versus
z.B.	zum Beispiel

**Purin- und Pyrimidinbasen**

A	Adenin
C	Cytosin
G	Guanin
T	Thymin

**Aminosäuren**

Ala	Alanin (A)
Arg	Arginin (R)
Asn	Asparagin (Asn)
Asp	Asparaginsäure (D)
Cys	Cystein (C)
Gln	Glutamin (Q)
Glu	Glutaminsäure (E)
Gly	Glycin (G)
His	Histidin (H)
Ile	Isoleucin (I)
Leu	Leucin (L)
Lys	Lysin (K)
Met	Methionin (M)
Phe	Phenylalanin (F)
Pro	Prolin (P)
Ser	Serin (S)
Thr	Threonin (T)
Trp	Tryptophan (W)

Tyr            Tyrosin (Y)

Val            Valin (V)

### **Nomenklatur der Mutationen und Gene**

Die Mutationen wurden entsprechend der internationalen Nomenklatur der Human Genome Variation Society (HGVS) und die Gennamen entsprechend der Nomenklatur des HUGO Gene Nomenclature Committee (HGNC) bezeichnet (Den Dunnen & Antonarakis, 2000; Wain, Lush, Ducluzeau, Khodiyar, & Povey, 2004). Bei der Angabe des Aminosäureaustausches wurde der Drei-Buchstaben-Code verwendet.

## 9.2 Verwendete Anglizismen

<i>An Essay on the Shaking Palsy</i>	Abhandlung über die Schüttellähmung
<i>Ankyrin repeats</i>	N-terminale Ankyrin-Domäne
<i>bias</i>	Verzerrung
<i>cold spots</i>	Bereiche in der DNA, bei denen wenig Rekombination stattfindet
<i>compound</i>	zusammengesetzt
<i>confidence interval</i>	Konfidenzintervall
<i>coverage</i>	Abdeckung, Sequenziertiefe
<i>c-terminal of Roc,</i>	COR-Domäne
<i>deCODE Genetics</i>	isländisches Pharmaunternehmen
<i>dominant negativ Effekt</i>	eine dominante Mutation die die Wirkung des Wildtyp-Allels unterdrückt, d.h. es entsteht ein strukturell verändertes, jedoch stabiles Protein. Dieses beeinträchtigt die Gesamtfunktion stärker, als das vollständige Fehlen des entsprechenden Proteins
<i>Fisher's exact test</i>	Exakter Test nach Fisher
<i>fixed effect model</i>	Modell mit festen Effekten
<i>forward</i>	vorwärts
<i>gain-of-function</i>	die Mutation verursacht eine Verstärkung der Genaktivität oder führt zu einer neuen Funktion des Gens
<i>Haploinsuffizienz</i>	das gesunde Allel kann das mutierte nicht kompensieren
<i>hot spots</i>	Bereiche in der DNA, bei denen vermehrt Rekombination stattfindet, oder vermehrt Mutationen auftreten
<i>identical by decent</i>	der Herkunft nach gleich
<i>imputation</i>	Ersetzen fehlender Werte
<i>in silico</i>	hier: computergestützt

<i>in trans</i>	auf unterschiedlichen Allelen
<i>leucin rich repeat</i>	Leucin-reiche Wiederholung von AS
<i>long-range phasing</i>	Bestimmung der Phase von Allelen/Haplotypen
<i>Loss-of-function</i>	Genmutation, die einen Funktionsverlust des betreffenden Genprodukts nach sich zieht.
<i>maternal</i>	mütterlicherseits
<i>missense Mutation</i>	sinnverändernde Mutation; Substitution eines Nukleotids, die zum Austausch einer Aminosäure führt
<i>missing heritability</i>	fehlende Erblichkeit
<i>modifier</i>	hier: genetischer Einflussfaktor
<i>odds ratio</i>	Quotenverhältnis
<i>paternal</i>	väterlicherseits
<i>phasing</i>	Bestimmung der Phase von Allelen/Haplotypen, was heißt, welcher DNA-Strang mütterlicher bzw. väterlicherseits vererbt wurde
<i>random effect model</i>	Modell mit zufälligen Effekten
<i>reverse</i>	rückwärts
<i>surrogate father</i>	Ersatzvater
<i>surrogate mother</i>	Ersatzmutter
<i>tagging SNPs</i>	<i>tagSNPs / informative Marker</i>

### 9.3 Abbildungsverzeichnis

Abbildung 1: Allelfrequenz und Effektstärke identifizierter Risikovarianten.....	4
Abbildung 2: <i>Íslendingabók</i> - Das Busch der Isländer. ....	11
Abbildung 3: <i>Phasing</i> .....	12
Abbildung 4: Ersatzelternschaft (" <i>surrogate parenthood</i> ").....	13
Abbildung 5: Übersicht des Imputationsvorgangs der Sequenzvarianten. ....	14
Abbildung 6: Zusammenfassende Darstellung der Ressourcen von deCODE Genetics.....	15
Abbildung 7: Schematische Darstellung der nigrostriatalen Projektion.....	22
Abbildung 8: Genetische Architektur der Parkinson Erkrankung .....	28
Abbildung 9: Flussdiagramm zur Literatursuche und der Extraktion relevanter Studienergebnisse .....	37
Abbildung 10: Forest-Plot zur Darstellung der Meta-Analyse .....	41

### 9.4 Tabellenverzeichnis

Tabelle 1: Bekannte autosomal-dominante PD-Gene.....	97
Tabelle 2: Bekannte autosomal-rezessive PD-Gene .....	99
Tabelle 3: Bekannte autosomal-rezessive PD-Gene mit klinisch atypischem Bild.....	100
Tabelle 4: Loci mit unsicherer PD-Assoziation.....	101
Tabelle 5: Bekannte PD-Risiko-Loci.....	103
Tabelle 6: Kumulatives Risikomodell – häufige Risikovarianten .....	104
Tabelle 7: Kumulatives Risikomodell .....	106





## 10 Lebenslauf

**Johanna Marlen Huttenlocher-Moser**, geb. Huttenlocher

Diplom-Biologin

\*22. August 1980 in Stuttgart-Bad Cannstatt

verheiratet, 1 Kind

### SCHULBILDUNG

1987-2000

Freie Waldorfschule am Kräherwald / Stuttgart,

Abitur 2000

### AKADEMISCHE AUSBILDUNG

10/2002-10/2008

#### STUDIUM

Studium der Biologie, Eberhard Karls Universität Tübingen

Hauptfach: Tierphysiologie

1. Nebenfach: Parasitologie
2. Nebenfach: Psychologie

#### DIPLOMARBEIT

Institut für Tropenmedizin

Eberhard Karls Universität Tübingen

„Charakterisierung eines neuen Antigens von *Plasmodium falciparum*“

Betreuung: PD Dr. Jürgen Kun /

Prof. Dr. Peter G. Kremsner

**PROMOTION**

04/2009 – 08/2013

Institut für Medizinische Genetik und angewandte Genomik

Eberhard Karls Universität Tübingen

„Identifikation und Charakterisierung genetischer Risikofaktoren  
der Parkinson Erkrankung in der isländischen Bevölkerung“

Betreuung: Prof. Dr. Peter Bauer (Tübingen) /

Prof. Dr. Olaf Rieß (Tübingen) /

Dr. Hreinn Stefánsson (Island)

## 11 Publikationen

Im Folgenden sind Publikationen aufgelistet, die im Rahmen der vorliegenden experimentellen Dissertation erstellt wurden.

**Huttenlocher, J.**, Stefansson, H., Steinberg, S., Helgadóttir, H. T., Sveinbjörnsdóttir, S., Riess, O., & Bauer, P. (2015). Heterozygote carriers for CNVs in PARK2 are at increased risk of Parkinson's disease. *Human Molecular Genetics*, ddv277–. doi:10.1093/hmg/ddv277

Obayashi, M., Stevanin, G., Synofzik, M., Monin, M.-L., Duyckaerts, C., Sato, N., Streichenberger, N., Vighetto, A., Desestret, V., Tesson, C., Wichmann, HE., Illig, T., **Huttenlocher, J.**, Kita, Y., Izumi, Y., Mizusawa, H., Schöls, L., Klopstock, T., Brice, A., Ishikawa, K., Dürr, A. (2014). Spinocerebellar ataxia type 36 exists in diverse populations and can be caused by a short hexanucleotide GGCCTG repeat expansion. *Journal of Neurology, Neurosurgery, and Psychiatry*, 1–10. doi:10.1136/jnnp-2014-309153

**Huttenlocher, J.**, Krüger, R., Capetian, P., Lohmann, K., Brockmann, K., Csoti, I., ... Bauer, P. (2015). EIF4G1 is neither a strong nor a common risk factor for Parkinson's disease: evidence from large European cohorts. *Journal of Medical Genetics*, 52(1), 37–41. doi:10.1136/jmedgenet-2014-102570

Steinberg, S., de Jong, S., Mattheisen, M., Costas, J., Demontis, D., Jamain, S., Pietiläinen, OPH., Lin, K., Papiol, S., **Huttenlocher, J.**, ... Stefansson, K. (2014). Common variant at 16p11.2 conferring risk of psychosis. *Molecular Psychiatry*, 19, 108–14. doi:10.1038/mp.2012.157

Jonsson, T., Stefansson, H., Steinberg, S., Jonsdóttir, I., Jonsson, P. V., Snaedal, J., Björnsson,

S., **Huttenlocher, J.**, ... Stefansson, K. (2013). Variant of TREM2 associated with the risk of Alzheimer's disease. *The New England Journal of Medicine*, 368, 107–16. doi:10.1056/NEJMoa1211103

Jonsson, T., Atwal, J. K., Steinberg, S., Snaedal, J., Jonsson, P. V., Bjornsson, S., Stefansson, H., Sulem, P., Gudbjartsson, D., Maloney, J., Hoyte, K., Gustafson, A., Liu, Y., Lu, Y., Bhangale, T., Graham, RR., **Huttenlocher, J.**, ... Stefansson, K. (2012). A mutation in APP protects against Alzheimer's disease and age-related cognitive decline. *Nature*. doi:10.1038/nature11283

Steinberg, S., de Jong, S., Andreassen, O. A., Werge, T., Børglum, A. D., Mors, O., Mortensen, PB., Gustafsson, O., Costas, J., Pietiläinen, OPH., Demontis, D., Papiol, S., **Huttenlocher, J.**, ... Stefansson, K. (2011). Common variants at VRK2 and TCF4 conferring risk of schizophrenia. *Human Molecular Genetics*, 20, 4076–4081. doi:10.1093/hmg/ddr325

Schicks, J., Synofzik, M., Pétursson, H., **Huttenlocher, J.**, Reimold, M., Schöls, L., Bauer, P. (2011). Atypical juvenile parkinsonism in a consanguineous SPG15 family. *Movement Disorders*, 26(3), 564–6. doi:10.1002/mds.23472

## 12 Anhang

### 12.1 Bekannte PD-Gene

Tabelle 1 : Bekannte autosomal-dominante PD-Gene

PD-Subtyp Genlocus OMIM	Gen Genprodukt	Genetische Varianten	Häufigkeit	Klinische Charakteristika Erkrankungsalter	Zellbiologische Funktion	Referenz
<b>PARK1</b> 4q22.1 #168601	SNCA α-Synuclein	A30P, E46K, H50Q, G51N, A53T	sehr selten	ähnlich des sporadischen PD; Lewy-Körperchen, schnelle Progression, L-DOPA-responsiv, häufig kognitive / psychiatrische Beeinträchtigung  Erkrankungsalter: 46,5 ± 10,8 Jahre; A53T (früh manifestierend) A30P (spät manifestierend)	α-Synuclein ist Hauptbestandteil der Lewy-Körperchen; vermutlich an der Regulation der synaptischen Neurotransmission beteiligt; durch Fehlfaltung des Proteins kommt es zu Fibrillenbildung und Aggregation von α-Synuclein mit nachfolgendem Nervenzelluntergang	(Appel-Cresswell et al., 2013; Kiely et al., 2013; Krüger et al., 1998; Polymeropoulos et al., 1997; Zarranz et al., 2004)
<b>PARK4</b> 4q22.1 #605543	SNCA α-Synuclein	Dupl / Tripl	sehr selten	das Ausmaß der α-Synuclein-Überexpression korreliert mit dem Schweregrad des Phänotyps der Erkrankung, rascherer und häufiger mit Demenz assoziierter Krankheitsverlauf, sowie mit dem Erkrankungsalter.	gesteigerte Expression des Proteins bei Multiplikationen	(A. B. Singleton et al., 2003)
				Erkrankungsalter: Tripl (früh manifestierend) Dupl (spät manifestierend)		

<b>PARK8</b> 12q12 #607060	<b>LRRK2</b> leucine-rich repeat kinase 2 / Dardarin	> 40 PM, 7 davon pathogen: G2019S; R1441C, R1441G, R1441H, Y1699C, I2020T, R1437H	1-2 % aller PD; 5-15 % der Patienten mit AD Form - abhängig von der ethnischen Gruppe	ähnlich des sporadischen PD; partiell $\alpha$ -Synuclein-positive Lewy-Körperchen, L-DOPA-responsiv Erkrankungsalter: ca. 60 Jahre	Phosphorylierung von $\alpha$ -Synuclein und / oder Tau; Kinase-Aktivität	(A. Di Fonzo et al., 2005; Funayama et al., 2002, 2005; Gilks et al., 2005; Mata et al., 2005; Alexander Zimprich et al., 2004)
<b>PARK17</b> 16q11.2 #614203	<b>VPS35</b> vacuolar protein sorting associated protein 35	D620N	selten (0,4 %)	ähnlich des sporadischen PD; L-DOPA-responsiv, mögliche kognitive und Verhaltensänderungen Erkrankungsalter: ca. 60 Jahre	das VPS35- Protein ist Baustein des Retromer Komplexes, dieser ist für den Rücktransport von Endosomen zum Golgi-Apparat zuständig, daher geht man von einer gestörten Wiederverwertung membrangebundener Proteine durch Mutationen in diesem Gen aus.	(Kumar, Weissbach, et al., 2012; Vilarinho-Güell et al., 2011; Wider et al., 2008; A Zimprich et al., 2011)

Tabelle 2: Bekannte autosomal-rezessive PD-Gene

PD-Subtyp Genlocus OMIM	Gen Genprodukt	Genetische Varianten	Häufigkeit	Klinische Charakteristika Erkrankungsalter	Zellbiologische Funktion	Referenz
<b>PARK2</b> 6q25-q27 #600116	<i>PRKN</i> / <i>PARK2</i> parkin	> 100 (PM, Del / Dupl)	1 % aller PD; 10-20 % der Patienten mit EO	häufig Dystonie und Dyskinesie, un- spezifische nigrale Degeneration, sel- ten Lewy-Körperchen, langsame Pro- gression, kaum kognitive Symptome, dafür schnelle motorische Fluktuation, L-DOPA-responsiv  Erkrankungsalter: 37 ± 12	die E3-Ubiquitin-Ligase ist am Parkin-PINK1-vermittelten Signalweg für die Elimination dysfunktionaler Mitochondrien (Mitophagie) in den Zellen verantwortlich. Verlust von Zellen im Gehirnstamm, keine kortikale Pathologie	(Abbas et al., 1999; Dawson & Dawson, 2010; Kitada et al., 1998; Shimura et al., 2000, 2001)
<b>PARK6</b> 1p35-p36 #605909	<i>PINK1</i> PTEN-induced kinase 1	selten Del > 40 PM: E129X, Q129fsX157, P196L, G309N, W437X, G440E, Q456X	selten mit 1-9 % nach <i>PARK2</i> die häufigste Ursache für EO	ähnliche Symptomatik wie bei <i>PARK2</i> , kaum kognitive Symptome, langsame Progression, L-DOPA-responsiv, keine Dystonie  Erkrankungsalter: 32-48	die Proteinkinase ist am Par- kin-PINK1-vermittelten Signal- weg für die Rekrutierung von Parkin zu den Mitochondrien verantwortlich. Verlust von Zellen im Gehirnstamm, keine kortikale Pathologie	(Abou-Sleiman et al., 2006; V Bonifati et al., 2005; E M Valente et al., 2001; Enza Maria Valente et al., 2004)
<b>PARK7</b> 1p36.23 #606324	<i>DJ-1</i> DJ-1	> 10 (PM, Del / Dupl) dup168-185, A39S, E64D, D149A, Q163L, L166P, M261L	sehr seltene 1 % der Patienten mit EO	ähnliche Symptomatik wie bei <i>PARK2</i> , kaum kognitive Symptome, langsame Progression, fokale Dystonie, hetero- zygote Fälle zeigen Lewy-Körperchen  Erkrankungsalter: früh	Sensormolekül für oxidativen Stress. Bei einem Funktions- verlust des Proteins, kommt es zu einer eingeschränkten mitochondrialen und einer reduzierten neuroprotektiven Funktion. Verlust von Zellen im Gehirnstamm	(V Bonifati et al., 2002; Vincenzo Bonifati et al., 2003; Kriebehl et al., 2010; van Duijn et al., 2001)



Tabelle 3: Bekannte autosomal-rezessive PD-Gene mit klinisch atypischem Bild

PD-Subtyp Genlocus OMIM	Gen Genprodukt	Genetische Varianten	Häufigkeit	Klinische Charakteristika Erkrankungsalter	Zellbiologische Funktion	Referenz
<b>PARK9</b> 1p36 #606693 Kufor-Rakeb syndrom	ATP13A2 ATPase 13A2	> 5 PM: L552fsX788, M810R, G877R, G1019fsX1021 kleine InDel: 1103insGA, del2742TT	selten	Komplexes Syndrom mit Parkinsonismus als Teil des klinischen Spektrums, die Parkinson-Symptomatik tritt jedoch im Verlauf der Krankheit in den Hintergrund  Erkrankungsalter: früh 10.-20. Lebensjahr	die ATPase ist in der lysosomalen Membran lokalisiert. Funktionsverlust des Proteins führt zu dessen Abbau im Proteasom und einer mitochondrialen Dysfunktion	( a. Di Fonzo et al., 2007; Ramirez et al., 2006)
<b>PARK14</b> 22q13.1 #612953 Dystonie- Parkinsonismus	PLA2G6 phospholipase A2 group VI	D331Y, R635Q, R7441Q, R747W	selten	Komplexes Syndrom mit Parkinsonismus als Teil des klinischen Spektrums, die Parkinson-Symptomatik tritt jedoch im Verlauf der Krankheit in den Hintergrund  Erkrankungsalter: früh 10.-20. Lebensjahr	Neurodegeneration mit Eisenablagerungen im Gehirn	(Paisan-Ruiz et al., 2009; Yoshino et al., 2010)
<b>PARK15</b> 22q12-q13 #260300 Parkinson- pyramidales Syndrom	FBXO7 F-box only protein 7	T22M, R378G, R498X	selten	Komplexes Syndrom mit Parkinsonismus als Teil des klinischen Spektrums, die Parkinson-Symptomatik tritt jedoch im Verlauf der Krankheit in den Hintergrund  Erkrankungsalter: früh 10.-20. Lebensjahr	das F-Box Protein 7 ist in das Ubiquitin-Proteasom-System involviert	( a Di Fonzo et al., 2009; Shojaaee et al., 2008)
<b>PARK19</b> 1p31.3 #615528	DNAJC6 DnaJ heat shock protein family member C6	Spleiß- Mutation c.801-2 A>G, Q734X		Komplexes Syndrom mit Parkinsonismus als Teil des klinischen Spektrums.  Erkrankungsalter: früh 10.-20. Lebensjahr	Auxilin, wird vor allem in Nervenzellkörperchen exprimiert. Es spielt eine Rolle bei der Clathrin induzierten Endozytose	(Edvardson et al., 2012)



<b>PARK20</b> 21q22.11 #615530	<b>SYNJ1</b> synaptojanin 1	<b>R258Q</b>	Komplexes Syndrom mit Parkinsonismus als Teil des klinischen Spektrums. Dystonie, kognitiver Verfall, unregelmäßige Augenbewegung und Krämpfe  Erkrankungsalter: früh	Phosphatase Aktivität. Gestörte synaptische Vesikel Verwertung	(Krebs et al., 2013; Quadri et al., 2013)
--------------------------------------	-----------------------------------	--------------	---	--	---

Tabelle 4: Loci mit unsicherer PD-Assoziation

PD-Subtyp Genlocus OMIM	Gen Genprodukt	Genetische Varianten	Häufigkeit	Klinische Charakteristika Erkrankungsalter	Zellbiologische Funktion	Referenz
<b>PARK3</b> 2p13	unbekannt			ähnlich des sporadischen PS. Dopaminerge Degeneration in der SNc, Lewy-Körperchen niedrige Peptidanz (40 %)  Erkrankungsalter: ± 59		(T Gasser et al., 1998)
<b>PARK5</b> 4p13 #613643	<i>UCHL1</i> ubiquitin-Carboxyl-terminal hydrolase L1	I93M	sehr selten, keine klare Assoziation möglich	ähnlich des sporadischen PS  Erkrankungsalter: ± 50	Beteiligung am Ubiquitin-Proteasom-Komplex; Abbau von polyubiquitinierten Proteinen	(Leroy et al., 1998)
<b>PARK10</b> 1p32 #606852	unbekannt			ähnlich des sporadischen PS  Erkrankungsalter: unsicher, spät	unbekannt	(Hicks et al., 2002)
<b>PARK11</b> 2q36-q37 #607688	<i>GIGYF2</i> GRB10-interacting GYF protein 2			ähnlich des sporadischen PS  Erkrankungsalter: spät		(Nathan Pankratz et al., 2002, 2003)

<b>PARK12</b> Xq21-q25	unbekannt		(Hicks et al., 2002; N Pankratz et al., 2003; Nathan Pankratz et al., 2002; Scott et al., 2001)
<b>PARK13</b> 2p13.1 #610297	<i>Omi/HTRA2</i> HtrA serine peptidase 2	ähnlich des sporadischen PS Erkrankungsalter: unsicher	(Simón-Sánchez & Singleton, 2008; Strauss et al., 2005)
<b>PARK16</b> 1q32	unbekannt		(Satake et al., 2009; Simón-Sánchez et al., 2009)
<b>PARK18</b> 3q27.1 #614251	<i>EIF4G1</i> eukaryotic translation- initiation faktor 4 gamma 1	ähnlich des sporadischen PS Erkrankungsalter: ± 64	(Chartier-Harlin et al., 2011) verschlechterte Stressantwort der Zelle durch veränderte Translation von mRNAs
<b>PARK21</b> 3q22.1 #614334	<i>DNAJC13</i> DnaJ heat shock protein family mem- ber C13		(Vilarinho-Güell et al., 2014)

Tabelle 5: Bekannte PD-Risiko-Loci

Genlocus OMIM	Gen Genprodukt	Genetische Varianten	Häufigkeit	Klinische Charakteristika Erkrankungsalter	Zellbiologische Funktion	Referenz
4q21	SNCA α-Synuclein	Risikofaktoren mit geringer Effektstärke	häufig	sporadisches PS Erkrankungsalter: spät manifestie- rend	erhöhte Expression des α- Synuclein Proteins durch regulatorische Polymorphis- men	(Maraganore et al., 2006; Schiesling, Kieper, Seidel, & Krüger, 2008)
17q21.31 #157140	MAPT microtubule- associated protein tau	Risikofaktoren mit geringer Effektstärke	häufig	sporadisches PS Erkrankungsalter: spät manifestie- rend	Aggregation durch Hyper- phosphorylierung von Tau	(Martin et al., 2001; M. A. Nalls et al., 2011; Scott et al., 2001)
12q12	LRRK2 leucine-rich repeat ki- nase 2 / Dardarin	Risikofaktoren mit geringer Effektstärke wie auch protective Varianten	häufig	sporadisches PS Erkrankungsalter: spät manifestie- rend		(Simón-Sánchez et al., 2009)
1q22 #606463	GBA glucosylcera midase beta	Risikofaktoren mit mittlerer Effektstärke	selten	sporadischen PS Erkrankungsalter: spät manifestie- rend	Funktionsverlust des Proteins führt zu einer Störung des lysosomalen Proteinabbaus mit daraus folgender α- Synuclein Akkumulation	(T Gasser et al., 2011; Kumar, Lohmann, & Klein, 2012)

PM=Punktmutation, Del=Deletion, Tripl=Triplikation, Dupl=Duplikation, InDel=Insertion und Deletion

## 12.2 Kumulatives Risikomodell

Tabelle 6: Kumulatives Risikomodell – häufige Risikovarianten

Risikovarianten, die für die Berechnung des kumulativen Relativen Risikos der G2019S-Mutationsträger verwendet wurden

dbSNP	Chromosom	Position	Gen	Allel (major/minor)	Risiko-Allel	MAF	OR – 1/OR	Referenz
i4000416	1	153472258	GBA	T/C	C	0,005	4,048	(Do et al., 2011)
rs708723	1	204005889	RAB7L1/PARK16	T/C	T	0,439	1,10497	(I. Parkinson et al., 2011)
rs823156	1	204031263	SLC41A1	A/G	A	0,183	1,20919	(Do et al., 2011)
rs6710823	2	135308851	ACMSD/TMEM163	G/A	A	0,19	1,38	(M. A. Nalls et al., 2011)
rs2390669	2	168800188	STK39	A/C	C	0,13	1,19	(Lill et al., 2012)
rs34016896	3	162475558	NMD3	C/T	T	0,305	1,14	(I. Parkinson et al., 2011)
rs10513789	3	184242767	MCCC1/LAMP3	T/G	T	0,201	1,24533	(Do et al., 2011)
chr4:911311	4	911311	GAK	G/C	C	0,28	1,21	(M. A. Nalls et al., 2011)
rs11248060	4	954359	DGKQ	C/T	T	0,12	1,21	(Lill et al., 2012)
rs11724635	4	15346199	BST1	A/C	A	0,45	1,14943	(M. A. Nalls et al., 2011)
rs6812193	4	77418010	SCARB2/FAM47E	C/T	C	0,365	1,10375	(Do et al., 2011)
rs356219	4	90856624	SNCA	A/G	G	0,41	1,27	(Lill et al., 2012)
rs156429	7	23272545	GNMB	A/G	A	0,403	1,11857	(I. Parkinson et al., 2011)
rs591323	8	16741462	FGF20	G/A	G	0,271	1,13122	(I. Parkinson et al., 2011)
chr8:89442157	8	89442157	MMP16	C/T	T	0,0247	1,38	(I. Parkinson et al., 2011)



rs7077361	10	15601549	ITGA8	T/C	T	0,12	1.13636	(Lill et al., 2012)
rs1491942	12	38907075	LRRK2	C/G	G	0,21	1.30	(M. A. Nalls et al., 2011)
rs12817488	12	121862247	CCDC62/HIP1R	G/A	A	0,46	1.13	(M. A. Nalls et al., 2011)
rs4889603	16	30889726	STX1B	A/G	G	0,413	1.15	(I. Parkinson et al., 2011)
rs11868035	17	17655826	SREBF1/RAI1	G/A	G	0,309	1.17509	(Do et al., 2011)
rs2942168	17	41070633	MAPT	G/A	G	0,22	1.31579	(M. A. Nalls et al., 2011)
rs4130047	18	38932233	RIT2/SYT4	T/C	C	0,313	1.161	(Do et al., 2011)
rs2823357	21	15836776	USP25	G/A	A	0,376	1.149	(Do et al., 2011)

Tabelle 7: Kumulatives Risikomodell

Sample	Effekt (logOR)-per-allele	p-Wert
G2019S-Mutationsträger mit PD vs Kontrollen	-13.25	0.64
G2019S-Mutationsträger mit PD vs G2019S-Mutationsträger ohne PD ( $\geq 70$ )	-25.85	0.53
G2019S-Mutationsträger mit PD vs PD	-26.02	0.37
G2019S-Mutationsträger mit PD vs Familiäres PD	-38.38	0.21
G2019S-Mutationsträger mit PD vs Sporadisches PD	-21.89	0.44
G2019S-Mutationsträger ohne PD ( $\geq 70$ ) vs Kontrollen	4.96	0.80
G2019S-Mutationsträger ohne PD ( $\geq 70$ ) vs PD	-7.15	0.72
G2019S-Mutationsträger ohne PD ( $\geq 70$ ) vs Familiäres PD	-16.02	0.47
G2019S-Mutationsträger ohne PD ( $\geq 70$ ) vs Sporadisches PD	-3.90	0.84
PD vs Kontrollen	12.11	9.2e-06
PD vs Familiäres PD	-9.16	0.13
PD vs Sporadisches PD	9.78	0.11
Familiäres PD vs Kontrollen	18.58	0.00033
Familiäres PD vs Sporadisches PD	9.78	0.11
Sporadisches PD vs Kontrollen	8.99	0.0044

## **Publikationen**

**EIF4G1 is neither a strong nor a common risk factor for Parkinson's disease: evidence from large European cohorts.**

**Huttenlocher J**, Krüger R, Capetian P, Lohmann K, Brockmann K, Csoti I, Klein C, Berg D, Gasser T, Bonin M, Riess O, Bauer P.

Originalarbeit publiziert in „Journal of Medical Genetics“ (2014).





## ORIGINAL ARTICLE

# *EIF4G1* is neither a strong nor a common risk factor for Parkinson's disease: evidence from large European cohorts

Johanna Huttenlocher,<sup>1</sup> Rejko Krüger,<sup>2,3,4</sup> Philipp Capetian,<sup>5</sup> Katja Lohmann,<sup>5</sup> Kathrin Brockmann,<sup>2,3</sup> Ilona Csoti,<sup>6</sup> Christine Klein,<sup>5</sup> Daniela Berg,<sup>2,3</sup> Thomas Gasser,<sup>2,3</sup> Michael Bonin,<sup>1</sup> Olaf Riess,<sup>1</sup> Peter Bauer<sup>1</sup>

► Additional material is published online only. To view please visit the journal online (<http://dx.doi.org/10.1136/jmedgenet-2014-102570>).

<sup>1</sup>Institute of Medical Genetics and Applied Genomics, University of Tübingen, Tübingen, Germany

<sup>2</sup>Center for Neurology and Hertie-Institute for Clinical Brain Research, University of Tübingen, Tübingen, Germany

<sup>3</sup>German Center for Neurodegenerative Diseases (DZNE), Tübingen, Germany

<sup>4</sup>Department of Clinical and Experimental Neuroscience, Luxembourg Center for Systems Biomedicine (LCSB), University of Luxembourg, Luxembourg

<sup>5</sup>Institute of Neurogenetics, University of Lübeck, Lübeck, Germany

<sup>6</sup>Gertrudis-Kliniken, Biskirchen, Germany

**Correspondence to**

Professor Peter Bauer, Institute of Medical Genetics and Applied Genomics, University of Tübingen, Calwerstrasse 7, Tübingen DE 72076, Germany; [peter.bauer@med.uni-tuebingen.de](mailto:peter.bauer@med.uni-tuebingen.de)

JH and RK contributed equally.

Received 4 June 2014

Revised 9 October 2014

Accepted 10 October 2014

**ABSTRACT**

**Background** Missense mutations in the eukaryotic translation initiation factor 4- $\gamma$  1 (*EIF4G1*) gene have previously been implicated in familial Parkinson's disease (PD). A large PD family with autosomal-dominant segregation showed a heterozygous missense mutation and additional patients were found to have unique sequence variants that have not been observed in controls. Subsequent studies have reported contradictory findings.

**Methods** We assessed the relevance of *EIF4G1* mutations in a European cohort of 2146 PD patients. Of these, 2051 sporadic PD patients were screened for the reported p.Ala502Val and p.Arg1205His mutations. In addition, the complete coding region of *EIF4G1* was directly sequenced in 95 familial PD patients with autosomal-dominant inheritance. Moreover, we imputed the p.Arg1205His substitution and tested for association with PD in the Icelandic population (93 698 samples).

**Results** We did not observe the presence of the p.Ala502Val substitution in our cohort; however, the p.Arg1205His mutation was identified in one sporadic PD patient. The same mutation was also found in 76 Icelandic subjects older than 65 years using haplotype imputing. Only five of these subjects reported PD symptoms (OR 1.3,  $p=0.50$ ). Thus, if causal, the p.Arg1205His *EIF4G1* mutation has a low penetrance or a late onset manifestation. A novel variant p.Arg566Cys found in a patient with familial PD did not cosegregate with PD in all three affected siblings. All further recently published *EIF4G1* mutations found in our cohort are likely to be benign polymorphisms.

**Conclusions** This is the largest genetic study of *EIF4G1* mutations in PD. Our data do not support the *EIF4G1* gene as a high-risk PD locus, neither for the familial nor the sporadic condition. Furthermore, the p.Arg1205His mutation is not significantly associated with increased risk of PD in the Icelandic population. Therefore, caution should be exercised when interpreting *EIF4G1* genotyping results in isolated patients and PD families. In summary, diagnostic testing of *EIF4G1* should not be recommended in clinical settings.

**INTRODUCTION**

Sporadic Parkinson's disease (PD) is a progressive neurodegenerative disorder of the nervous system, affecting 1%–2% of the European population older than 65 years.<sup>1</sup>

The central pathological feature of PD is a progressive degeneration of dopaminergic neurons leading to a lack of dopamine in the nigrostriatal projection as well as the presence of intraneuronal cytoplasmic inclusions termed Lewy bodies. The degeneration leads to a core set of symptoms including bradykinesia, resting tremor, muscular rigidity and postural instability.<sup>2</sup>

The aetiology of PD is not well understood; however, it has been assumed that PD is a mainly sporadic, complex disorder resulting from genetic and environmental factors.<sup>3</sup> Nevertheless, a small fraction of cases (5%–10%) seem to arise by Mendelian inheritance. During the last 16 years, investigation of families with high prevalence of PD led to the identification of 11 genes causing autosomal-recessive forms of PD (PARKIN (MIM 602544), PINK1 (MIM 608309), DJ1 (MIM 602533), ATP13A2 (MIM 610513), PLA2G6 (MIM 603604), FBX07 (MIM 605648), DNAJC6 (MIM 615528), SYNJ1 (MIM 615530)) and autosomal-dominant forms of PD (SNCA (MIM 163890), LRRK2 (MIM 609007), VPS35 (MIM 601501)).<sup>4</sup> Studying the function of these genes provided important insights into the molecular pathways involved in neuronal degeneration. It is anticipated that the identification of further genes could help unravel the underlying mechanism of the disease in familial and sporadic PD.

Recently, Farrer *et al*<sup>5</sup> performed linkage and sequencing analysis of a French family with autosomal-dominant late-onset Parkinsonism and identified linkage in a 31 cM interval between D3S1763 and D3S1580 in the chromosomal region 3q26–q28. Mutation screening in subsequent families discovered five putative disease-associated missense variants in the eukaryotic translation initiation factor 4- $\gamma$  1 (*EIF4G1*) located within this chromosomal region and therefore the PARK18 locus was assigned to this gene (OMIM #614251). In the reported French family, the arginine to histidine substitution at position 1205 of the polypeptide was predicted to be damaging to the protein function.<sup>5</sup> Subsequent studies of the *EIF4G1* gene identified previously described mutations also in healthy controls, raising some doubt on their pathogenic relevance.<sup>6–8</sup> Furthermore, no association between other putative *EIF4G1* mutations and PD was seen in a large case control sequencing study.<sup>9</sup>

**To cite:** Huttenlocher J, Krüger R, Capetian P, *et al*. *J Med Genet* Published Online First: [please include Day Month Year] doi:10.1136/jmedgenet-2014-102570

The aim of the present work was to investigate the relevance of the two most reported *EIF4G1*-associated variants (p.Ala502Val and p.Arg1205His) in a large cohort of 2051 sporadic German PD patients using high-resolution melting (HRM) analysis and subsequent Sanger sequencing. The frequency of p.Arg1205His variant was furthermore tested in the general Icelandic population by imputing the respective haplotype into 93 698 chip-typed Icelandic individuals. Moreover, we also assessed the prevalence of other *EIF4G1* mutations in familial PD by sequencing the entire coding sequence in 95 German and Icelandic individuals with a positive family history for PD.

## MATERIAL AND METHODS

### Patients and controls

A set of 2051 unrelated sporadic PD patients of European descent (Lübeck (n=672, age at onset 51.3±14.4 years, 41% women, family history known in 69%, 23% thereof positive for PD), Tübingen (n=1379, age at onset 56.1±12.0 years, 41% women, positive family history in 22.3%)) and an additional 95 patients with a reported family history of PD (Tübingen (n=80), Iceland (n=15)) were included in our study. Furthermore, the p.Arg1205His variant, identified through the whole genome sequencing project at deCODE genetics/Amgen, was imputed into the 93 698 chip-genotyped Icelandic individuals, patients and controls.

### Study subjects from Germany

Patients were recruited as consecutive outpatients at the University of Lübeck and Tübingen, both tertiary referral centres for movement disorders. The diagnosis of PD was established by a movement disorders specialist according to the UK Brain Bank Criteria.

### Study subjects from Iceland

PD patients were ascertained through an ongoing, population-based study as previously described.<sup>10–11</sup> In brief, they were identified from the Icelandic Parkinson's Disease Society, neurologists, general practitioners, and records of prescriptions for levodopa and other drugs commonly given to patients with PD. Additionally, we included individuals classified as PD patients through the Resident Assessment Instrument, used in nursing homes to judge the care dependency.<sup>12</sup> After detailed clinical examination by a neurologist specialised in movement disorders, patients were considered to have PD if they had at least two of the following cardinal signs: resting tremor, rigidity, bradykinesia or postural instability, whereas other causes of Parkinsonism were excluded. This resulted in a total of 1387 PD subjects who were available for chip-genotyping.

For the control group, we used individuals participating in different studies at deCODE, excluding those individuals diagnosed with PD or any PD symptoms. In total, 92 311 samples underwent SNP analysis.

The study was approved by the National Bioethics Committee of Iceland and the Data Protection Commission of Iceland. Written informed consent was obtained from all participants prior to the study. All personal identifiers associated with medical information and blood samples, along with the genealogy database, were encrypted using a third-party encryption system as previously described.<sup>13</sup>

### Mutation screening for the p.Arg1205His and the p.Ala502Val variants in individuals with sporadic PD

HRM curve analysis was performed to screen 2051 German patients for exon 10 and exon 24, where putative PD variants

were described.<sup>5</sup> Peripheral blood samples were collected and genomic DNA extracted using standard protocols. For amplification, we used previously reported primers<sup>5</sup> with some modifications to improve the amplicon quality (see online supplementary table S1).

PCR and HRM were carried out in a LightCycler480 instrument (Roche Diagnostics, Mannheim, Germany). Conspicuous samples identified by aberrant melting profiles different to the reference sample were subjected to direct sequencing. After reamplification from genomic DNA and purification, each PCR product was sequenced in both forward and reverse directions to identify the underlying potential variants (Big Dye Terminator Cycle Sequencing Ready Reaction Kit (Applied Biosystems, Foster City, California, USA) using an ABI3100 Genetic Analyser (Applied Biosystems)).

### Screening for novel *EIF4G1* variants in individuals with a family history of PD

In order to identify additional disease mutations, we sequenced the entire *EIF4G1* (ENSG00000114867, ENST00000424196) coding sequence (exons 1–33) in 95 PD patients with a family history of autosomal dominantly inherited PD (Germany (n=80), Iceland (n=15)). Genomic DNA was extracted from peripheral blood using standard methods. Dye-terminator Sanger sequencing was performed with the Applied Biosystems BigDye Terminator V3.1 Cycle Sequencing Kit, with Agencourt Ampure XP and Agencourt CleanSeq for cleanup of the PCR and cycle sequencing product, respectively. AmpureXP and CleanSeq bead cleaning was performed on a Zymark Sciclone ALH-500 liquid-handling robot system. Tray dilutions for PCR setup and cycle-sequencing setup were prepared on a Packard MultiprobeII HTEX liquid-handling robot system, and genomic DNA and PCR product were transferred into their respective trays using the Zymark Sciclone ALH-500. PCR and cycle sequencing reactions were performed on MJ Research PTC-225 thermal cyclers. Subsequently, electropherograms were analysed with Applied Biosystems 3730xl DNA analysers. All potentially relevant variants were examined for disease segregation in relatives of the carriers by additional sequencing.

### Genotyping and data imputation

The p.Arg1205His variant was identified in the course of the whole genome sequencing project at deCODE. A total of 2630 Icelandic individuals were sequenced on average at 15×. Subsequently, genotypes had been imputed into long-ranged phased haplotypes<sup>14</sup> of 93 698 chip-typed samples (IlluminaHumanHap and IlluminaOmniExpress arrays). The *EIF4G1* locus was sequenced at 27× and the p.Arg1205His variant was imputed from four sequenced carriers with an information score of 0.99981. To verify the imputation results, individual genotyping of the rs112176450 variant in 200 imputed individuals was done applying the Centaurus (Nanogen) platform.<sup>15</sup> Genotyping and imputation were consistent for all 200 samples.

We then performed genome-wide association testing for the p.Arg1205His variant in 2730 patients and 39 046 control subjects older than 65. Five patients and 71 controls carried the variant.

The methods used for chip-genotyping, long-range phasing, genotype imputation, genealogy-based *in silico* genotyping and association testing have recently been reported.<sup>16–17</sup>

### Microsatellite genotyping

The presence of a possible gene rearrangement within the chromosomal region harbouring the *EIF4G1* gene was assessed by reconstructing the haplotypes by means of microsatellite markers in a family carrying a putative pathogenic mutation (p.Arg566Cys). Nine previously described<sup>5</sup> microsatellite repeat markers were chosen and genotyped in all family members: D3S2314, D3S1571, D3S3609, D3S3578, Chr3\_185.633, Chr3\_185.692, D3S3583, D3S3592 and D3S1262. They cover a 4 Mb interval in chromosome 3q27.1 where the p.Arg566Cys is located. One primer of each set was 5' labelled with a fluorescent tag. PCR reactions were performed under standard conditions. To determine the fragment length, amplification products were run on a Beckman Coulter sequencer (CEQ8000) (see online supplement for primer sequences).

### Prediction of functional effects of coding variants

In silico prediction for a possible impact of non-synonymous amino acid exchanges on the protein function was performed by SIFT<sup>18</sup> (Sorting Intolerant From Tolerant; <http://sift.jcvi.org/>) and PolyPhen2<sup>19</sup> (<http://genetics.bwh.harvard.edu/pph2/>).

## RESULTS

### Screening for the p.Arg1205His and the p.Ala502Val variants in individuals with sporadic PD

The initial analysis was conducted in order to assess the frequency of the previously published p.Ala502Val and p.Arg1205His variants and screen for novel disease-causing mutations in the two hotspot regions, exon 10 and 24 of the *EIF4G1* gene. These two exons were investigated in a large cohort of 2051 patients with PD. In total, we identified six rare coding SNPs and four known polymorphisms (table 1). The p.Ala502Val variant was not found in our cohort. However, one patient carried the heterozygous p.Arg1205His mutation.

Of the remaining variants in exon 10 and exon 24, p.Met432Val,<sup>5 6 9 20</sup> p.Pro486Ser<sup>6 8 9</sup> as well as p.Gly465\_Arg467del<sup>5 7 9</sup> were previously reported while the sequence alterations p.Pro323Ser, p.Pro434Ser and

p.Arg1166Gln were not described. Whether they may affect protein function needs to be ascertained by functional analysis.

Allele frequency for the common SNP (p.Met432Val) in our HRM analysis was comparable with published frequencies in the Exom Variant Server (<http://evs.gs.washington.edu/EVS/>) and in the Icelandic population.

### Screening for novel *EIF4G1* variants in individuals with a family history of PD

Besides sporadic PD, we additionally sequenced all 33 coding exons of *EIF4G1* in 95 index patients with a positive family history of PD. We identified three non-synonymous and four silent variants (table 1). All variants have been reported before (p.Ser260Leu,<sup>6</sup> p.Arg566Cys,<sup>6</sup> p.Ile807Val<sup>5 6</sup>). Neither the p.Ala502Val nor the p.Arg1205His mutation was observed in this cohort.

The heterozygous p.Arg566Cys variant was identified in a male patient who showed a clear family burden of the disease in three first degree relatives. PolyPhen2 and SIFT analysis suggested a possible change in protein function due to this variant which is why we decided to follow-up this variant in the family of the index patient. No evidence for segregation within the family could be shown. Recombination events were excluded through microsatellite haplotype analysis (data not shown).

### Association of the p.Arg1205His haplotype in the Icelandic population

While we identified the p.Arg1205His variant only once in the 2146 PD patients and the p.Ala502Val not at all, the imputation of the p.Arg1205His variant into the long-ranged phased haplotypes in the Icelandic population allowed us to identify 76 carriers older than 65 years. Five of them reported PD symptoms without a family history of PD. Subsequent association analysis revealed no significant correlation between the mutation and the phenotype (OR=1.3, p=0.50).

## DISCUSSION

*EIF4G1* was implicated in autosomal-dominant late-onset PD in 2011.<sup>5</sup> Since then, the contribution of this gene to PD has been

**Table 1** Coding variants detected in the *EIF4G1* gene in Parkinson's disease patients

Gene	Location	Nucleotide change	Amino acid change	SIFT	PolyPhen2	dbSNP entry
EIF4G1	E10	c.780C>T	p.Ser260Leu	Tolerated	Possible damaging	
EIF4G1	E10	c.792G>A	p.Pro264Pro	Silent	Silent	
EIF4G1	E10	c.951C>T	p.Pro317Pro	Silent	Silent	
EIF4G1	E10	c.960C>T	p.Leu320Leu	Silent	Silent	
EIF4G1	E10	c.967C>T	p.Pro323Ser	Tolerated	Benign	
EIF4G1	E10	c.1294A>G	p.Met432Val	Tolerated	Benign	rs2178403
EIF4G1	E10	c.1300C>T	p.Pro434Ser	Tolerated	Benign	
EIF4G1	E10	c.1384_1392delGAAGAA	p.Gly465_Arg467del			
EIF4G1	E10	c.1456C>T	p.Pro486Ser	Tolerated	Benign	rs112545306
EIF4G1	E12	c.1696C>T	p.Arg566Cys	Damaging	Probably damaging	rs145521479
EIF4G1	E16	c.2416A>G	p.Ile807Val	Tolerated	Benign	rs62287499
EIF4G1	E20	c.2976A>G	p.Pro992Pro	Silent	Silent	rs112420733
EIF4G1	E24	c.3497G>A	p.Arg1166Gln	Tolerated	Probably damaging	
EIF4G1	E24	c.3614G>A	p.Arg1205His	Deleterious	Probably damaging	rs112176450
EIF4G1	E24	c.3654 G>C	p.Arg1218Arg	Silent	Silent	
EIF4G1	E29	c.4251C>T	p.Val1417Val	Silent	Silent	rs76779558
EIF4G1	E30	c.4383C>T	p.Phe1461Phe	Silent	Silent	rs111921843
EIF4G1	E30	c.4386C>T	p.Asp1462Asp	Silent	Silent	rs112718796

Human *EIF4G1* cDNA sequence (RefSeq Accession number NM\_198241.2) was used as the reference sequence.



controversial<sup>6–8 21 22</sup> as several non-penetrant mutation carriers have also been described. Thus, further investigation in a large cohort of PD patients was necessary to substantiate the role of *EIF4G1* in PD. We therefore conducted a comprehensive genetic analysis of the gene in 2051 patients with sporadic PD and 95 index patients with familial PD. As well, we analysed the p.Arg1205His risk haplotype in the Icelandic population.

Of particular interest was the role of the repeatedly reported p.Ala502Val and the p.Arg1205His substitutions in PD. Chartier-Harlin *et al*<sup>5</sup> could demonstrate that these variants result in an impaired oxidative stress response, downregulation of mitochondrial biogenesis and increased autophagy.<sup>23–25</sup> Thus, we focused on these mutations in our initial screening. The analysis of our 2051 sporadic PD patients revealed p.Arg1205His in one sporadic PD patient and no carriers of the p.Ala502Val variant. Beyond that, additional sequencing of the whole coding region in a cohort of 95 familial cases failed to identify either pathogenic mutation. However, we have to keep in mind that our dataset of patients with positive family history was probably too small to detect very rare variants with high risk. Thus, these results support previous smaller studies that described the mutations as very rare.<sup>6–9 26–29</sup>

In addition, we assessed the prevalence of other *EIF4G1* mutations in our combined sample from Germany and Iceland. Of the detected non-synonymous amino acid substitutions, we focused on the p.Arg566Cys variant found in a patient with three affected first degree relatives since *in silico* prediction tools suggested impaired protein function. The mutation is located close to the eIF4E binding site, insinuating a functional impact. However, although the mutation is located close to the eIF4E binding site and suggests a function impairment, segregation analysis of the family of the index patient provided no evidence of the role in PD inheritance. This can be explained by a recent study by Tucci *et al*<sup>6</sup> who analysed the EVS dataset and suggested a rather high tolerance for genetic variability in this region. The possibility that this variant has low pathogenic relevance is further supported by the observation that the p.Arg566Cys variant was observed in one European-American and one African-American control sample in the EVS (<http://evs.gs.washington.edu/EVS/>) before. Therefore, determination of their true pathogenic impact will require additional functional analysis or tests in larger cohorts and datasets.

Although we identified several variants in the *EIF4G1* gene, our study argues against *EIF4G1* mutations as a high-risk cause of PD. These findings are in line with results from Nuytemans and colleagues who showed no evidence for an overall association of the *EIF4G1* gene with PD, except for the p.Arg1205His mutation.<sup>9</sup>

Since we found p.Arg1205His only in one patient from the German cohort, we sought to further investigate the role of this mutation by performing association analysis in 2730 patients and 39 046 control subjects older than 65 in Iceland. Results did not reveal significant association with PD (OR=1.3,  $p=0.50$ ). Of note is that although the Icelandic control group excluded all individuals with reported history of PD, subtle clinical symptoms related to PD may have escaped. Nevertheless, these findings do not support a role of p.Arg1205His as disease causing variant. Instead, the data point to a need for careful interpretation of this variant and suggest a lower than predicted penetrance. This is in agreement with recent studies that found the originally described mutations also in control samples.<sup>6–8</sup>

In conclusion, we provide a large amount of data that do not support PARK18 as high-risk PD locus. Our results indicate that amino acid substitutions due to mutations in the *EIF4G1* gene

represent extremely rare events in both the German and Icelandic population. This does not exclude a possible contributing effect of the *EIF4G1* variants in PD and may indicate a genetic risk factor with reduced penetrance. As large genome-wide association studies based on common genetic variants did not lead to the identification of novel genes explaining the missing heritability in PD, attention has turned towards rare variants that do not cause disease by themselves but increase PD risk with an intermediate effect size.<sup>30 31</sup> These variants may contribute in different ways to the disease risk and their interpretation is known to be difficult.<sup>9</sup> They may explain a yet unaccounted part of the total risk of disease. This would be in accordance with the hypothesis of pleomorphic risk loci, which postulate a combination of common and rare variants at the same locus to be responsible for a person's individual risk. These pleomorphic genes can comprise highly-penetrant point mutations that cause familial forms of the disease as well as variants that confer increased susceptibility to sporadic disease.

Taken together, our results demonstrate the genetic complexity of *EIF4G1* variations in PD. Although our observations do not rule out *EIF4G1* as a genetic modifier of PD, the interpretation of genotyping results in isolated patients and PD families requires specific attention and *EIF4G1* variants may impose problems in diagnostic settings.

**Acknowledgements** We would like to thank all patients for participating in this study and M Farrer for sharing primer sequences prior to publication. We gratefully acknowledge deCODE genetics for providing access to data derived from the whole genome sequencing project.

**Contributors** All listed authors have substantially contributed to the work. RK, KL, PC, IC, DB and CK investigated PD patients and contributed samples. JH, RK, MB and PB performed genetic analyses and interpretation. JH, RK, TG, OR and PB drafted the manuscript. All authors approved the final version of this manuscript.

**Funding** This study was partly supported by an IAPP grant of the European Commission (MarkMD under grant agreement PIAP-GA-2008-230596).

**Competing interests** None.

**Ethics approval** Local ethic committees in Tübingen, Lübeck and Reykjavik.

**Provenance and peer review** Not commissioned; externally peer reviewed.

**Data sharing statement** All relevant data are included in the supplementary document.

## REFERENCES

- Bardien S, Lesage S, Brice A, Carr J. Genetic characteristics of leucine-rich repeat kinase 2 (LRRK2) associated Parkinson's disease. *Park Relat Disord* 2011;17:501–8.
- Gelb DJ, Oliver E, Gilman S. Diagnostic criteria for Parkinson disease. *Arch Neurol* 1999;56:33–9.
- Langston JW. Epidemiology versus genetics in Parkinson's disease: progress in resolving an age-old debate. *Ann Neurol* 1998;44:S45–52.
- Gasser T, Hardy J, Mizuno Y. Milestones in PD genetics. *Mov Disord* 2011;26:1042–8.
- Chartier-Harlin MC, Dachsel JC, Vilarino-Guell C, Lincoln SJ, LePrete F, Hulihan MM, Kachergus J, Milnerwood AJ, Tapia L, Song MS, Le Rhun E, Mutez E, Larvor L, Dufflot A, Vanbesien-Mailliot C, Kreisler A, Ross OA, Nishioka K, Soto-Ortolaza AI, Cobb SA, Melrose HL, Behrouz B, Keeling BH, Bacon JA, Hentati E, Williams L, Yanagiya A, Sonenberg N, Lockhart PJ, Zubair AC, Uitti RJ, Aasly JO, Krygowska-Wajs A, Opala G, Wszolek ZK, Frigerio R, Maraganore DM, Gosal D, Lynch T, Hutchinson M, Bentivoglio AR, Valente EM, Nichols WC, Pankratz N, Foroud T, Gibson RA, Hentati F, Dickson DW, Destee A, Farrer MJ. Translation initiator EIF4G1 mutations in familial Parkinson disease. *Am J Hum Genet* 2011;89:398–406.
- Tucci A, Charlesworth G, Sheerin UM, Plagnol V, Wood NW, Hardy J. Study of the genetic variability in a Parkinson's Disease gene: EIF4G1. *Neurosci Lett* 2012;518:19–22.
- Lesage S, Condroyer C, Klebe S, Lohmann E, Durif F, Damier P, Tison F, Anheim M, Honore A, Viallet F, Bonnet AM, Ouvrard-Hernandez AM, Vidailhet M, Durr A, Brice A. EIF4G1 in familial Parkinson's disease: pathogenic mutations or rare benign variants? *Neurobiol Aging* 2012;33:2233.e1–e5.
- Schulte EC, Mollenhauer B, Zimprich A, Bereznai B, Lichtner P, Haubenberger D, Pirker W, Brucke T, Molnar MJ, Peters A, Gieger C, Trenkwalder C, Winkelmann J.

- Variants in eukaryotic translation initiation factor 4G1 in sporadic Parkinson's disease. *Neurogenetics* 2012;13:281–5.
- 9 Nuytemans K, Bademci G, Inchausti V, Dressen A, Kinnamoon DD, Mehta A, Wang L, Zuchner S, Beecham GW, Martin ER, Scott WK, Vance JM. Whole exome sequencing of rare variants in EIF4G1 and VPS35 in Parkinson disease. *Neurology* 2013;80:982–9.
  - 10 Sveinbjornsdottir S, Hicks AA, Jonsson T, Petursson H, Guggmundsson G, Frigge ML, Kong A, Gulcher JR, Stefansson K. Familial aggregation of Parkinson's disease in Iceland. *N Engl J Med* 2000;343:1765–70.
  - 11 Gudmundsson KR. A clinical survey of parkinsonism in Iceland. *Acta Neurol Scand* 1967;43(Suppl 33):1–61.
  - 12 Morris JN, Hawes C, Fries BE, Phillips CD, Mor V, Katz S, Murphy K, Drugovich ML, Friedlob AS. Designing the national resident assessment instrument for nursing homes. *Gerontologist* 1990;30:293–307.
  - 13 Gulcher JR, Kristjánsson K, Gudbjartsson H, Stefansson K. Protection of privacy by third-party encryption in genetic research in Iceland. *Eur J Hum Genet* 2000;8:739–42.
  - 14 Kong A, Masson G, Frigge ML, Gylfason A, Zusmanovich P, Thorleifsson G, Olason PI, Ingason A, Steinberg S, Rafnar T, Sulem P, Mouy M, Jonsson F, Thorsteinsdottir U, Gudbjartsson DF, Stefansson H, Stefansson K. Detection of sharing by descent, long-range phasing and haplotype imputation. *Nat Genet* 2008;40:1068–75.
  - 15 Kutyavin IV, Milesi D, Belousov Y, Podyminogin M, Vorobiev A, Gorn V, Lukhtanov EA, Vermeulen NM, Mahoney W. A novel endonuclease IV post-PCR genotyping system. *Nucleic Acids Res* 2006;34:e128.
  - 16 Holm H, Gudbjartsson DF, Sulem P, Masson G, Helgadóttir HT, Zanon C, Magnusson OT, Helgason A, Saemundsdóttir J, Gylfason A, Stefansdóttir H, Gretarsdóttir S, Matthiasson SE, Thorgeirsson GM, Jonasdóttir A, Sigurdsson A, Stefansson H, Werge T, Rafnar T, Kiemeneý LA, Parvez B, Muhammad R, Roden DM, Darbar D, Thorleifsson G, Walters GB, Kong A, Thorsteinsdóttir U, Arnar DO, Stefansson K. A rare variant in MYH6 is associated with high risk of sick sinus syndrome. *Nat Genet* 2011;43:316–20.
  - 17 Sulem P, Gudbjartsson DF, Walters GB, Helgadóttir HT, Helgason A, Gudjonsson SA, Zanon C, Besenbacher S, Bjornsdóttir G, Magnusson OT, Magnusson G, Hjartarson E, Saemundsdóttir J, Gylfason A, Jonasdóttir A, Holm H, Karason A, Rafnar T, Stefansson H, Andreassen OA, Pedersen JH, Pack AI, de Visser MC, Kiemeneý LA, Geirsson AJ, Eyjólfsson GI, Olafsson I, Kong A, Masson G, Jonsson H, Thorsteinsdóttir U, Jonsdóttir I, Stefansson K. Identification of low-frequency variants associated with gout and serum uric acid levels. *Nat Genet* 2011;43:1127–30.
  - 18 Ng PC, Henikoff S. SIFT: predicting amino acid changes that affect protein function. *Nucleic Acids Res* 2003;31:3812–14.
  - 19 Adzhubei IA, Schmidt S, Peshkin L, Ramensky VE, Gerasimova A, Bork P, Kondrashov AS, Sunyaev SR. A method and server for predicting damaging missense mutations. *Nat Methods* 2010;7:248–9.
  - 20 Li K, Tang BS, Guo JF, Lou MX, Lv ZY, Liu ZH, Tian Y, Song CY, Xia K, Yan XX. Analysis of EIF4G1 in ethnic Chinese. *BMC Neurol* 2013;13:38.
  - 21 Yuan L, Song Z, Xu H, Gu S, Zhu A, Gong L, Zhao Y, Deng H. EIF4G1 Ala502Val and Arg1205His variants in Chinese patients with Parkinson disease. *Neurosci Lett* 2013;543:69–71.
  - 22 Fujioka S, Sundal C, Strongosky AJ, Castanedes MC, Rademakers R, Ross OA, Vilarino-Guell C, Farrer MJ, Wszolek ZK, Dickson DW. Sequence variants in eukaryotic translation initiation factor 4-gamma (eIF4G1) are associated with Lewy body dementia. *Acta Neuropathol* 2013;125:425–38.
  - 23 Ramirez-Valle F, Braunstein S, Zavadil J, Formenti SC, Schneider RJ. eIF4G1 links nutrient sensing by mTOR to cell proliferation and inhibition of autophagy. *J Cell Biol* 2008;181:293–307.
  - 24 LeFebvre AK, Korneeva NL, Trutschl M, Cvek U, Duzan RD, Bradley CA, Hershey JW, Rhoads RE. Translation initiation factor eIF4G-1 binds to eIF3 through the eIF3e subunit. *J Biol Chem* 2006;281:22917–32.
  - 25 Zhou C, Huang Y, Przedborski S. Oxidative stress in Parkinson's disease: a mechanism of pathogenic and therapeutic significance. *Ann N Y Acad Sci* 2008;1147:93–104.
  - 26 Siitonen A, Majounie E, Federoff M, Ding J, Majamaa K, Singleton AB. Mutations in EIF4G1 are not a common cause of Parkinson's disease. *Eur J Neurol* 2013; 20:e59.
  - 27 Sudhama S, Behari M, Govindappa ST, Muthane UB, Juyal RC, Thelma BK. VPS35 and EIF4G1 mutations are rare in Parkinson's disease among Indians. *Neurobiol Aging* 2013;34:2442.e1–3.
  - 28 Chen Y, Chen K, Song W, Chen X, Cao B, Huang R, Zhao B, Guo X, Burgunder J, Li J, Shang HF. VPS35 Asp620Asn and EIF4G1 Arg1205His mutations are rare in Parkinson disease from southwest China. *Neurobiol Aging* 2013;34:1709.e7–8.
  - 29 Zhao Y, Ho P, Prakash KM, Foo JN, Liu JJ, Au WL, Tan LC, Tan EK. Analysis of EIF4G1 in Parkinson's disease among Asians. *Neurobiol Aging* 2013;34: 1311.e5–6.
  - 30 Haines JL. The importance of rare DNA variation in neurologic disease: cautionary tale. *Neurology* 2013;80:974–5.
  - 31 Lesage S, Brice A. Role of mendelian genes in "sporadic" Parkinson's disease. *Park Relat Disord* 2012;18(Suppl 1):S66–70.

## ***EIF4G1* is neither a strong nor a common risk factor for Parkinson Disease:**

### **Evidence from large European cohorts**

Huttenlocher et al.

#### **Supplementary Material**

Supplementary table 1: *EIF4G1* HRM amplification and sequencing primers

<b>Exon</b>	<b>size</b>	<b>primer</b>	<b>Sequence 5' → 3'</b>
EIF4G1 Exon02	198 bp	2F	GGTCCCTTCCTCTTCCTGAG
		2R	GATCCAGAGGCTGGAGACTG
EIF4G1 Exon03	221 bp	3F	TGCTCCCCTTATTTCCACCAG
		3R	AGCTTTCCACTGCTCCCAC
EIF4G1 Exon04	161 bp	4F	ATTGCCTTCATTCCCTCCC
		4R	CCCCTGACTAGGGTGGAG
EIF4G1 Exon05	229 bp	5-1F	CCAGGCTATAGAGACATTGTGG
		5-1R	GGGTAGGAGATCTGGGAAGG
	230 bp	5-2F	CCTCCCGAGTGCAGAGTG
		5-2R	CCTCTGGAAAAGCAATGGAA
EIF4G1 Exon06	292 bp	6F	CCCAGAGTTGGCTTGTCTTG
		6R	GGGAAGGGTCTTCTGCTTTAG
EIF4G1 Exon07	244 bp	7F	AGGTGTCAGGCAGGCATTAG
		7R	ACCTGACCTTGATTCTCCCC
EIF4G1 Exon08	163 bp	8F	TTCCTGCTTCCCCTCATCT
		8R	GCCAGGGTATCACTCGAAAA
EIF4G1 Exon09	200 bp	9F	AAAGAAGGGTTGTGGGACTC
		9R	AGTCACTCCACCACCTCAGC
EIF4G1 Exon10	392 bp	10-1F	GATGGGGAGGAACAACAGC
		10-1R	GAAAACCTCAGATTCTGGAACCG
	387 bp	10-2F	CATCTCCCGTGAAACTGGG
		10-2R	GATGCTGTCACCTCCTTGG
	391 bp	10-3F	ACCTGAGGAACTGCTCAACG
		10-3R	CACTCTCAGCTTCTCCTGCTT
280 bp	10-4F	CATCCCCTCTGCTACTCCAG	
	10-4R	GTCATCAAGGAGGGCAAGG	
EIF4G1 Exon11	222 bp	11F	TATTTCCAGGGATTAAGGG
		11R	TGACCACTGCTCCATACCTC
EIF4G1 Exon12	218 bp	12F	GCCAAGTTGAGGTATGGAGC
		12R	CAATCATACTGCCTCCTCTGG

EIF4G1 Exon13	250 bp	13F	TTTTCTTATTAGGGCCAGAGG
		13R	CTCCTCCCAATCCATCC
EIF4G1 Exon14	234 bp	14F	CGCTCTAGCCTGCTTCTGAG
		14R	GCCACCTGCCACTTAC
EIF4G1 Exon15	323 bp	15F	CCTGGAGGCGTGTAGTAGTG
		15R	AGATGGCAGACCTGGAGAAG
EIF4G1 Exon16	334 bp	16F	TAAGTCCCCACCACCCTCTC
		16R	GGGAATAGGAATGGGCAAC
EIF4G1 Exon17	278 bp	17F	GTTAAAGGCTCTTGGAGGGC
		17R	AGCATCTGGGTAAGGGAAGG
EIF4G1 Exon18	316 bp	18F	TCTTTTTGACACAATCCCTGT
		18R	ACATATGCTTCAGCCCATCC
EIF4G1 Exon19	245 bp	19F	GCCTGCAGTTATAGGTGGGAC
		19R	ATGATGTCAGGGAAGGCAAG
EIF4G1 Exon20	299 bp	20F	CACTGCCTTGTCTTGCCTTC
		20R	GGCACTAACAGGAAGTCATCAG
EIF4G1 Exon21	232 bp	21F	GGGAAGGGGCTGCTAGG
		21R	TCTCAACTTTAGAAAAGTGAAGGG
EIF4G1 Exon22	243 bp	22F	TTTGGTTGCATATGGTGGTG
		22R	AGGCTCTGTCCTCAAGGGAG
EIF4G1 Exon23	226 bp	23F	GGGGATAAAATGCAGGAAAG
		23R	GCAACAACCAGAACCCAATC
EIF4G1 Exon24	428 bp	24F	TGCTAAGAACAAGGCCCAAC
		24R	TCCCAGCAAACAAGATTTGG
EIF4G1 Exon25	234 bp	25F	TCCAAATCTTGTTTGTCTGGG
		25R	CAGCATGATCCCTCCCTG
EIF4G1 Exon26	310 bp	26F	GGGTCTGGGTCAGACAGTG
		26R	TATGCAAGGTTCCAAGGGTC
EIF4G1 Exon27	257 bp	27F	AGAAGGAGAGTGGGGACCTG
		27R	GTATCCAACCCAGCCATTC
EIF4G1 Exon28	218 bp	28F	ATACCCTGGTTTGGAGGGAG
		28R	CAAACAGCCAAAACCACACC
EIF4G1 Exon29	206 bp	29F	TTTGGTAAGCTGGGTTGGTC
		29R	GGTTGCCAACAGGAGACAAG
EIF4G1 Exon30	232 bp	30F	CCTGTTGGCAACCTTACCTC
		30R	CCCCACCCTAAGAACTACC
EIF4G1 Exon31	155 bp	31F	CAATTCAGAATCTGGGGATCA
		31R	ACCCTTGCCCCTCCATCT
EIF4G1 Exon32	281 bp	32F	GAGGGAGGTGGGTGTCAG
		32R	CCCTGCTATCTGGCAGTCC
EIF4G1 Exon33	269 bp	33F	GGCCAGTCACCTCTAACTGC
		33R	GGGCCATCTGTGTGTCCAT

Supplementary table 2: *EIF4G1* Microsatellite repeat markers

Marker ID	Position	Primer	Sequence 5' → 3'	Label
D3S2314	183,613,071	for	AGGGTTGTTGTGGGATTTGAA	5' IRD700
		rev	GCTGTGCTCTATTGTTAAGAC	
D3S1571	185,022,606	for	ACAGTGGCTGATGCCTT	5' IRD700
		rev	CACAGGTGGGCACTACAT	
D3S3609	185,519,706	for	AGCTGGGGACCAGTCT	5' IRD700
		rev	CGAGAGTAACTTGTACGGTG	
D3S3578	185,567,338	for	GATTGGGTTACAGAGTAGCCACAC	5' IRD700
		rev	TCTGAAATGTAGACTGCCCTTTG	
Chr3_185633	185,633,367	for	GTTGTAGAAAGCCGAGATTG	5' IRD700
		rev	ACCCTTGAGCAGTATTTTCAG	
Chr3_185692	185,692,505	for	AACAAAAGCGAAACTCTGTC	5' IRD700
		rev	GAGGGAGACTTCATTTTCAC	
D3S3583	185,767,793	for	TGCAAAGTCACAGATGTCCA	5' IRD700
		rev	CGAGAGGCACCAGAGTGTT	
D3S3592	185,893,115	for	GCAGTTCTGAGTGATTTACCA	5' IRD700
		rev	TCATCTGAGGTGTCTGATTG	
D3S1262	187,706,172	for	CGGCCCTAGGATATTTCAA	5' IRD700
		rev	CCAGTTTTTATGGACGGGGT	



**Heterozygote carriers for CNVs in PARK2 are at increased risk of Parkinson's disease**

**Huttenlocher J**, Stefansson H, Steinberg S, Helgadottir HT, Sveinbjörnsdóttir S, Riess O, Bauer P, Stefansson, K.

Originalarbeit publiziert in „Human Molecular Genetics“ (2015).



## ASSOCIATION STUDIES ARTICLE

# Heterozygote carriers for CNVs in *PARK2* are at increased risk of Parkinson's disease

Johanna Huttenlocher<sup>1,2</sup>, Hreinn Stefansson<sup>1</sup>, Stacy Steinberg<sup>1</sup>, Hafdis T. Helgadottir<sup>1</sup>, Sigurlaug Sveinbjörnsdóttir<sup>3,4,5</sup>, Olaf Riess<sup>2</sup>, Peter Bauer<sup>2</sup> and Kari Stefansson<sup>1,6,\*</sup>

<sup>1</sup>deCODE Genetics/AMGEN, Reykjavik 101, Iceland, <sup>2</sup>Institute of Medical Genetics and Applied Genomics, University of Tuebingen, Tuebingen D-72076, Germany, <sup>3</sup>Department of Neurology, National University Hospital, Reykjavik 101, Iceland, <sup>4</sup>Department of Neurology, MEHT, Broomfield Hospital, Court Road, Essex CM1 7ET, UK, <sup>5</sup>Neuroscience Department, Barts and the London School of Medicine and Dentistry, Queen Mary University of London, London, UK and <sup>6</sup>Faculty of Medicine, University of Iceland, Reykjavik IS-101, Iceland

\*To whom correspondence should be addressed at: deCODE Genetics, Sturugata 8, 101 Reykjavik, Iceland. Tel: +354 5701900; Fax: +354 5701901; Email: kstefans@decode.is

## Abstract

Together with point mutations, homozygous deletions or duplications in *PARK2* are responsible for the majority of autosomal recessive juvenile Parkinsonism. It is debated, however, whether heterozygous carriers of these mutations are at increased risk of Parkinson's disease (PD). Our goal was to determine whether heterozygous carriers of copy number variants (CNVs) affecting exons of the *PARK2* gene are at risk of PD that is greater than that of non-carriers. We searched for CNVs affecting exons of *PARK2* in a sample of 105 749 genotyped Icelanders. In total, 989 carriers, including 24 diagnosed with PD, were identified. The heterozygous carriers were tested for association in a sample of 1415 PD patients and 40 474 controls  $\geq 65$  years of age. PD patients were more often heterozygous carriers of *PARK2* CNVs than controls [odds ratio (OR) = 1.69,  $P = 0.03$ ] and compound heterozygous PD patients for a CNV and a missense mutation were not found. Furthermore, we conducted a meta-analysis of studies reporting on case-control samples screened for heterozygous *PARK2* CNVs. Ten studies were included in the final analysis, with 4538 cases and 4213 controls. The pooled OR and  $P$ -value for the published and Icelandic results showed significant association between *PARK2* CNVs and risk of PD (OR = 2.11,  $P = 2.54 \times 10^{-6}$ ). Our analysis shows that heterozygous carriers of CNVs affecting exons of *PARK2* have greater risk of PD than non-carriers.

## Introduction

Parkinson's disease (PD), the second most common progressive neurodegenerative disorder, was long believed to be of non-genetic, sporadic origin. However, in recent years, the understanding of the pathogenesis of PD has changed. Research has revealed rare monogenic forms of PD that are inherited in either an autosomal dominant or an autosomal recessive manner. Moreover, several common sequence variants conferring low or modest risk of PD have been found (1).

Four genes have already been identified contributing to autosomal recessive early onset Parkinsonism, a subtype of PD. These are *PARK2* at the *PARK2* locus (2) (OMIM\*602544), *PINK1* at *PARK6* (3) (OMIM\*608309), *DJ-1* at *PARK7* (4) (OMIM\*602533) and *ATP13A2* at *PARK9* (5) (OMIM\* 610513). Mutations in *PARK2*, an E3 ubiquitin ligase (6), are responsible for 50% of familial early onset PD (7,8).

*PARK2*-linked PD was first identified in a consanguineous Japanese family by Yamamura and colleagues (2) right before the gene was mapped to chromosome 6q25.2–q27 (9). *PARK2*

Received: March 22, 2015. Revised: June 24, 2015. Accepted: July 10, 2015

© The Author 2015. Published by Oxford University Press. All rights reserved. For Permissions, please email: journals.permissions@oup.com

has a high mutation rate (10,11) and is located in an unstable region in the genome. Thus, >100 putatively pathogenic mutations have been reported worldwide (12–15). Among them are point mutations and exon rearrangements, which contribute equally to the PARK2 mutation spectrum (1,16–18). The classical view is that carriers of one risk allele are not affected although this view is debated (19). Copy number variants (CNVs) constitute a particular challenge in genetic association studies and the impact of CNVs affecting PARK2 in heterozygous carriers in large samples remains to be studied (20).

Our goal was to ascertain the frequency of PARK2 mutations in a large genotyped sample, patients as well as controls and thereby determine the risk conferred by PARK2 CNVs in heterozygous carriers. The large genotyped sample at deCODE in which genotypes are long-range-phased (LRP) (21–23) and into which variants from a large whole genome sequencing effort have been imputed (24,25) was mined. The LRP sample was most helpful in identifying and excluding possible compound heterozygous carriers of mutations in PARK2. To augment our results, a meta-analysis of reported studies focusing on heterozygous PARK2 CNV carriers was performed.

## Results

The CNV analysis of 1415 PD patients and 104 320 controls revealed several large exon rearrangements in PARK2 as well as smaller deletions and duplications affecting predominantly exon 2 (Supplementary Material, Figure S1). After visual inspection of PennCNV calls, 989 calls were confirmed and included in further analysis. In detail, 24 (1.69%) PD patients and 965 (0.92%) controls were carriers of an exon rearrangement. The control frequency remained comparable when participants younger than 65 were excluded, 373/40 474 (0.91%). This is in keeping with Kay *et al.* (20), who found heterozygous PARK2 CNVs in 0.93% carrier frequency. In addition, among the patients we detected two homozygous CNV carriers, one homozygous for an exon 2 deletion and another homozygous for a large duplication spanning the gene. In the population sample, two siblings homozygous for an exon 2 duplication were found. Both are younger than 65 and have not developed PD.

As expected, homozygous CNVs affecting PARK2 were associated with PD ( $P = 0.001$ ). More noteworthy, our data demonstrate that heterozygous CNV carriers are at risk of developing PD [odds ratio (OR) = 1.69,  $P = 0.03$ ] (Table 1).

To address whether compound heterozygous CNV carriers could explain the association, we analyzed data derived from 2261 whole genome sequenced subjects whose genotypes had been imputed into LRP haplotypes (21–23). In total, data from 1387 PD patients and 93 698 controls were analyzed. We detected non-synonymous single nucleotide polymorphism (SNPs) in the PARK2 gene at a frequency of 2.45% in the Icelandic population,

almost equally distributed between cases and controls (2.89 versus 2.44%). These results are in good agreement with other recent studies [Kay *et al.* (26) (3%), Lincoln *et al.* (27) (3–4%) and Bruggemann *et al.* (28) (3–4%)]. In more detail, eight sequence variants affecting the amino acid sequence of PARK2 were identified (Table 2). Among those, a p.Met327Ile variant is novel and here exclusively found in control subjects. The remaining seven variants were found both in cases and controls and have previously been reported elsewhere (Table 2). Variants p.Ser167Asn, p.Val380Leu and p.Asp394Asn are common benign polymorphisms (29), whereas p.Ala82Glu, p.Arg402Cys and p.Pro409Leu may confer risk of PD (30). One of the sequence variants, p.Arg275Trp, was predicted to be deleterious by both SIFT (<http://sift.jcvi.org/>) and PolyPhen2 (<http://genetics.bwh.harvard.edu/pph2/>) and had been classified as a causative or susceptibility variant in earlier studies (15,31–34). Consistent with that, under a recessive model, a suggestive association with p.Arg275Trp was detected (OR = 23.1,  $P = 0.079$ ) in our sample. A burden test of missense mutations did not reveal significant association with PD in keeping with recent studies (26).

Through imputation, we only capture variants down to frequencies of ~0.01%. To search for very rare causative variants that contribute to the risk, we Sanger sequenced PD subjects carrying CNVs in PARK2 and all patients carrying the p.Arg275Trp allele. This effort did not uncover additional known or novel PARK2 mutations.

The LRP genotypes allow for determining whether mutations in carriers of two PARK2 mutations are on the same chromosome (21,24). No PD patients were found to be compound heterozygous for the rare allele of p.Arg275Trp and a CNV affecting PARK2. Two controls, siblings older than 64, were compound heterozygous for an exon 2 duplication and an Arg275Trp mutation (Supplementary Material, Table S1).

## Meta-analysis

A comprehensive review of 130 articles reporting on CNVs in PARK2 revealed 10 studies that fulfilled the inclusion criteria for a meta-analysis. In total, data on 8751 individuals (4538 cases and 4213 controls) were used. In the combined analysis, significant association was observed between carrying one copy of PARK2 CNV and PD (OR = 2.53,  $P = 1.67 \times 10^{-5}$ ). Including the results from Iceland, the significance increased (OR = 2.11,  $P = 2.54 \times 10^{-6}$ ). Results are summarized in Table 3.

## Discussion

This study focused on the identification of heterozygous CNVs in the PARK2 gene with the purpose of evaluating their contribution to the pathogenesis of PD. Our data demonstrate that heterozygous CNV carriers are at risk of PD (OR = 1.69,  $P = 0.03$ ). Also, our comprehensive attempts to exclude compound heterozygotes

**Table 1.** Association results: OR (95% CI) of developing PD in heterozygous and homozygous PARK2 CNV carriers

	Cases w/CNV	w/out CNV	Controls w/CNV	w/out CNV	OR (95% CI)	P-value
<b>Heterozygotes</b>						
All	22	1393	965	103 355	1.69 (1.05, 2.02)	0.024
65+	22	1393	373	40 101	1.69 (1.08, 2.68)	0.03
<b>Homozygotes</b>						
All	2	1413	2	104 318	73.79 (5.35, 1011.86)	0.0011
65+	2	1413	0	40 474	Infinity	0.0011

**Table 2.** Missense mutations identified in the PARK2 gene

References	Nuc change	AA change	dbSNP entry	SIFT	Minor allele frequency	Heterozygote 65+ controls		Homozygote 65+ controls	
						P-value	OR	P-value	OR
Hattori et al. (12)	c.500G>A	p.Ser167Asn	rs1801474	Tolerated	1.747	0.178546	1.191	0.94138	0.891
–	c.981G>T	p.Met327Ile		Tolerated	0.117	0.264059	0.417	No homoz	–
Sironi et al. (30)	c.1310C>T	p.Pro437Leu	rs149953814	Probably damaging	0.071	0.323662	1.769	No homoz	–
Bardien et al. (35)	c.1204C>T	p.Arg402Cys	rs55830907	Probably damaging	0.007	0.485155	0.017	No homoz	–
Hedrich et al. (17)	c.245C>A	p.Ala82Glu	rs55774500	Tolerated	1.388	0.542428	1.096	0.576629	2.119
Wang et al. (36)	c.1138G>C	p.Val380Leu	rs1801582	Tolerated	14.834	0.624713	0.976	0.647642	0.923
Abbas et al. (15)	c.1180G>A	p.Asp394Asn	rs1801334	Tolerated	2.584	0.853262	0.979	1	1.002
Ruffmann et al. (33) and Cookson et al. (34)	c.823C>T	p.Arg275Trp	rs34424986	Damaging	0.373	0.917152	1.032	0.0790227	23.072

Human PARK2 cDNA sequence (RefSeq accession number NM\_004562.2) was used as the reference sequence.

**Table 3.** Summary of meta-analysis: OR (95% CI) of developing PD in heterozygous PARK2 CNV carriers

Study	Ethnicity	Location	Cases		Controls		OR (95% CI)	P-value
			w/ CNV	w/out CNV	w/ CNV	w/out CNV		
Kay et al. (20)			18	2073	16	1670	0.91 (0.43, 1.91)	0.8629
Pankratz et al. (37)	Caucasian	USA	11	409	0	263		0.00863
Pankratz et al. (38)	Caucasian	USA	10	386	8	848	2.74 (0.97, 8.07)	0.03897
Brooks et al. (39)	Caucasian	USA	7	243	3	273	2.27 (0.51, 13.78)	0.3397
Lesage et al. (40)	Caucasian	France	5	167	0	170		0.0607
Guerrero Camacho et al. (41)	Latin-American	Mexican-mestizo	11	52	1	119	24.76 (3.44, 1086.20)	3.758e–05
Nuytemans et al. (42)	Caucasian	Belgium	8	302	1	269	7.11 (0.94, 316.89)	0.04161
Simon-Sanchez et al. (43)	Caucasian	USA	3	273	4	272	0.75 (0.11–4.46)	1
Lincoln et al. (27)	Caucasian	USA	3	310	2	190	0.92 (0.10, 11.11)	1
Wang et al. (44)			4	243	0	104		0.3232
All			80	4458	35	4178	2.53 (1.63, 3.99)	1.662e–05
Iceland (65+)	Caucasian	Iceland	22	1393	373	40 101	1.69 (1.076, 2.684)	0.03
All			102	5851	408	44 279	2.106 (1.544, 2.873)	2.54e–6

support the notion that haploinsufficiency is a more likely explanation than compound heterozygosity through unidentified recessive mutations. Furthermore, our association results are supported by our meta-analysis of reported studies (OR = 2.11,  $P = 2.54e \times 10^{-6}$ ).

So far, most studies evaluating the risk of PARK2 mutations in heterozygous carriers focus on SNPs rather than CNVs (26,27,40,42,45). However, heterozygous CNVs in PARK2 might be even more pathogenic than heterozygous point mutations (38). A few groups have focused on CNVs. In agreement with our results, two reports indicate that heterozygous mutations in PARK2 associate with increased risk of PD (37,38), albeit others found no evidence for an association (20,46,47). One possible reason for the diverging results could be the small sample sizes of these studies. The low frequency of CNVs renders a large sample set, as we used, essential for a reliable assessment of the disease risk.

Not only association studies but also examinations of families have yielded conflicting results. Heterozygous family members of homozygous carriers have been described with mild PD signs of late onset (15,48–50), whereas another group found no typical clinical signs of the disease (47). A recent study showed that typical PARK2-associated symptoms, such as dystonia and dyskinesia, correlate with the number of mutated PARK2 alleles (51). Furthermore in comparison to homozygous carriers, characterized by an onset before the age of 40 years, some groups showed that heterozygous carriers of CNVs as well as point mutations are associated with a later onset (36,42,52–54). Still, an earlier age at

onset for PARK2 mutation carriers compared with that of individuals without mutations in PARK2 is suspected (31,53,55).

The notion that heterozygous carriers may be at increased risk of PD is also supported by neuroimaging and electrophysiological findings. Preclinical changes such as reduced fluorodopa uptake in the striatum of asymptomatic heterozygous mutation carriers were connected to development of PD later in life (32,56–58).

As shown before (46), compound heterozygous mutations in PARK2 exist and may explain the disease in a fraction of mutation carriers. However, there are few statistical methods for analyzing the role of compound heterozygous mutations in disease, especially when such mutations are rare. The major barrier is that phase information is required to determine whether both gene copies are affected and phasing rare variants is difficult. Thus, phase determination is important in carriers of two mutations as both mutations may have been inherited from the same parent. In most PD studies, phasing is impossible, parents of cases are usually not genotyped and many of the patients are sporadic cases. Thus, determining whether two or one mutations are on the same chromosome is difficult and uncertainty about the phase of strings of SNPs creates complications in genetic analysis. Most available methods can only phase a small number of SNPs effectively and become unreliable when applied to SNPs spanning many linkage disequilibrium blocks. By means of LRP haplotypes (21), we were able to determine the phase of alleles likely to confer risk of PD. Hence, to further corroborate our

results, samples from the Icelandic subjects heterozygous for the PARK2 CNVs were screened for missense mutations in the PARK2 gene. Of the uncovered coding variants, only the p.Arg275Trp mutation was considered causal (10,33,59) and the search for compound heterozygotes revealed none of our cases to carry a CNV and a p.Arg275Trp mutation in *trans*. Therefore, we conclude that the detected OR of 1.69 in our sample set is mediated through dosage of single heterozygous mutations (CNVs). Haploinsufficiency, a dominant-negative effect or gain of function, is the most likely pathogenic mechanism behind it (16,60,61).

Heterozygous PARK2 CNV carriers in the large Icelandic sample are at risk of developing PD and our meta-analysis further supports the involvement of PARK2 CNVs in heterozygous carriers. To date, most studies focusing on subjects heterozygous for CNVs and missense variants in PARK2 are limited by modest sample sizes. Still, our literature search resulted in a combined sample set of 4538 cases and 4213 controls. Tested for association, the OR was even higher than in Iceland, supporting our results. However, we have to keep in mind that sampling differs between groups. Furthermore, there has been a continuous increase in the quality of copy number analysis tools over the past years, and recent studies suggest that some earlier studies may have included false-positive calls. Publication bias is also of concern in meta-analysis (62). Most importantly, not all the meta-analysis samples were screened for missense mutations. Hence, compound heterozygous PARK2 mutation carriers might have been included in previous datasets.

With regard to genetic testing and counseling, our results do not support routine screening because the disease risk for heterozygous CNV carriers is low. However, with the increasing availability of genetic testing for familial and sporadic PD, a better understanding of the inheritance pattern and the penetrance will help to interpret the test results. In particular, those with a family history of PD and identified PARK2 mutations in a first-degree relative may benefit from genetic testing. Albeit preventive therapy is currently not available, genetic testing may allow the mitigation of uncertainty and help to interpret any initial clinical problems, enabling early intervention in mutation carriers, if necessary.

In conclusion, we provide further evidence in support of our opinion that heterozygous PARK2 CNV carriers are at increased risk of developing PD. The mutations may act as susceptibility factors or disease modifiers and might further contribute to the heritability of familial and sporadic PD. Thus, the impact of heterozygous CNVs in PARK2 should not be underestimated and should be the focus of future investigation. It is important to further elucidate their role in PD to provide an accurate diagnosis and counseling of mutation carriers.

## Materials and Methods

### Patients and controls

The ascertainment of Icelandic PD patients has been described in detail elsewhere (63–65). In brief, patients included were enrolled from two studies. The first study identified patients through genealogical data based on censuses, church books, the national registry or obituaries and information that was collected between 1953 and 1963. The second study included patients that were recruited through an ongoing, population-based study at deCODE. Since 1994, patients were identified from the Icelandic Parkinson's Disease Society, neurologists, general practitioners, the Resident Assessment Instrument used in nursing homes to estimate dependency (66) and records of prescriptions for levodopa and other drugs commonly given to patients with PD.

PD patients underwent a detailed clinical examination by a movement disorder specialist and were included in the study if they had at least two of the following cardinal signs: tremor, rigidity, bradykinesia or postural instability (67), and other causes of Parkinsonism were excluded.

### Genotyping

In total, 105 749 DNA samples were genotyped using IlluminaHumanHap and IlluminaOmniExpress arrays. Of those, 93 698 could be LRP, which is necessary for effective imputation of rare variants. Sequence variants from the whole genome sequencing project at deCODE, including 2261 subjects sequenced on average at 15×, were imputed into the large sample (24,25). Of 1415 genotyped PD samples, 1387 were imputed. The genotyped controls were population controls recruited through other studies at deCODE genetics.

All biological samples were collected after approval had been obtained from the National Bioethics Committee and the Icelandic Data Protection Authority. Written informed consent was provided by the participants. In order to protect personal information on health and disease, all samples were encrypted through a third independent party in accordance with the regulations of the Icelandic Data Protection Committee (68).

### Detection of copy number variation

The PennCNV caller, described in detail elsewhere (69), was used for searching the 105 749 genotyped subjects for CNVs. The algorithm uses signal intensity data from SNP probes to estimate the copy number of genomic regions. The approach is based on a hidden Markov model, a statistical method, which combines several components, to estimate the probability of the allelic state to be true and generate the most reliable CNV calls. The following measurements were included: the log R ratio (normalized measure of total signal intensity at each SNP), the B allele frequency (normalized measure of allelic intensity ratio at each SNP), the distance between neighboring SNPs (probability of having a copy number state change) and the population frequency of the B allele. Subsequently, the software calculates whether the intensity is above or below a given empirical threshold. Only CNVs spanning 10 or more consecutive markers were considered true after visual inspection. Frequencies of PARK2 CNVs found in samples genotyped using different Illumina genotyping chips are shown in Supplementary Material, Table S2.

### Sanger sequencing

In order to identify additional disease causing mutations that remained undetected through imputation, we sequenced the entire PARK2 (NM\_004562.2) coding sequence (exons 1–12) and flanking intronic regions (Supplementary Material, Table S3) in all PD patients heterozygous for PARK2 deletions or duplications and all patients carrying the rare allele of the identified missense mutation (rs1621268). Therefore, genomic DNA was extracted from peripheral blood using standard methods. Dye-terminator Sanger sequencing was performed with the Applied Biosystems BigDye Terminator v3.1 Cycle Sequencing Kit, with Agencourt Ampure XP and Agencourt CleanSeq for cleanup of the polymerase chain reaction (PCR) and cycle-sequencing product, respectively. AmpureXP and CleanSeq bead cleaning was performed on a Zymark Sciclone ALH-500 liquid-handling robot system. Tray dilutions for PCR setup and cycle-sequencing setup were prepared on a Packard MultiprobeII HTEX liquid-handling robot



system, and genomic DNA and PCR product were transferred into their respective trays using the Zymark Sciclone ALH-500. PCR and cycle-sequencing reactions were performed on MJ Research PTC-225 thermal cyclers. Subsequently, electropherograms were analyzed with Applied Biosystems 3730xl DNA analyzers.

### Prediction of functional effects of coding variants

*In silico* prediction for a possible change in protein function was performed by SIFT (<http://sift.jcvi.org/>) (70) and PolyPhen2 (<http://genetics.bwh.harvard.edu/pph2/>) (71).

### Literature review for meta-analysis

Published articles reporting on heterozygous PARK2 CNV carriers were identified through computer searches of the Medline database (<http://www.ncbi.nlm.nih.gov/pubmed/>) and the PD mutation database ([www.molgen.ua.ac.be/PDmutDB](http://www.molgen.ua.ac.be/PDmutDB)). Combinations of the search terms 'PD or Parkinson's disease', 'Parkin or PARK2', 'CNVs or copy number variations' and 'heterozygous' resulted in 127 articles. Furthermore, our literature list was supplemented by a manual search of references cited in published articles, resulting in three additional publications.

Of these 130 studies, 32 were excluded due to improper content, leaving 98 potential literature citations for abstract and publication review. Following this step, studies were included in our analysis when exact numbers of heterozygous CNVs were reported for cases and controls.

In total, 10 studies met the full inclusion criteria and provided sufficient information to estimate study-specific OR and 95% confidence interval (CI) by Fisher's exact test. On the whole, 8751 individuals (4538 cases and 4213 controls) were enrolled. To estimate the pooled OR and *P*-value for all non-Icelandic strata, the exact Mantel-Haenszel method was used. For the Icelandic study, which included related individuals, an empirical *P*-value was obtained based on 500 000 drop-down simulations through the Icelandic genealogy. The published and Icelandic results were combined using standard, fixed effects, inverse-variance weighted meta-analysis.

### Supplementary Material

Supplementary Material is available at HMG online.

*Conflict of Interest statement.* None declared.

### Funding

This work was supported by an European Community Framework Programme 7, People Programme and IAPP on novel genetic and phenotypic markers of Parkinson's disease and Essential Tremor (MarkMD) (under grant agreement PIAP-GA-2008-230596 MarkMD).

### References

- Gasser, T., Hardy, J. and Mizuno, Y. (2011) Milestones in PD genetics. *Mov. Disord.*, **26**, 1042–1048.
- Kitada, T., Asakawa, S., Hattori, N., Matsumine, H., Yamamura, Y., Minoshima, S., Yokochi, M., Mizuno, Y. and Shimizu, N. (1998) Mutations in the parkin gene cause autosomal recessive juvenile parkinsonism. *Nature*, **392**, 605–608.
- Valente, E.M., Abou-Sleiman, P.M., Caputo, V., Muqit, M.M., Harvey, K., Gispert, S., Ali, Z., Del Turco, D., Bentivoglio, A.R., Healy, D.G. et al. (2004) Hereditary early-onset Parkinson's disease caused by mutations in PINK1. *Science*, **304**, 1158–1160.
- Abou-Sleiman, P.M., Healy, D.G., Quinn, N., Lees, A.J. and Wood, N.W. (2003) The role of pathogenic DJ-1 mutations in Parkinson's disease. *Ann. Neurol.*, **54**, 283–286.
- Di Fonzo, A., Chien, H.F., Socal, M., Giraud, S., Tassorelli, C., Iliceto, G., Fabbrini, G., Marconi, R., Fincati, E., Abbruzzese, G. et al. (2007) ATP13A2 missense mutations in juvenile parkinsonism and young onset Parkinson disease. *Neurology*, **68**, 1557–1562.
- Shimura, H., Hattori, N., Kubo, S., Mizuno, Y., Asakawa, S., Minoshima, S., Shimizu, N., Iwai, K., Chiba, T., Tanaka, K. et al. (2000) Familial Parkinson disease gene product, parkin, is a ubiquitin-protein ligase. *Nat. Genet.*, **25**, 302–305.
- Lucking, C.B., Durr, A., Bonifati, V., Vaughan, J., De Michele, G., Gasser, T., Harhangi, B.S., Meco, G., Deneffe, P., Wood, N.W. et al. (2000) Association between early-onset Parkinson's disease and mutations in the parkin gene. *N. Engl. J. Med.*, **342**, 1560–1567.
- Gasser, T. (2009) Mendelian forms of Parkinson's disease. *Biochim. Biophys. Acta*, **1792**, 587–596.
- Matsumine, H., Saito, M., Shimoda-Matsubayashi, S., Tanaka, H., Ishikawa, A., Nakagawa-Hattori, Y., Yokochi, M., Kobayashi, T., Igarashi, S., Takano, H. et al. (1997) Localization of a gene for an autosomal recessive form of juvenile Parkinsonism to chromosome 6q25.2-27. *Am. J. Hum. Genet.*, **60**, 588–596.
- Periquet, M., Lucking, C., Vaughan, J., Bonifati, V., Durr, A., De Michele, G., Horstink, M., Farrer, M., Illarioshkin, S.N., Pollak, P. et al. (2001) Origin of the mutations in the parkin gene in Europe: exon rearrangements are independent recurrent events, whereas point mutations may result from Founder effects. *Am. J. Hum. Genet.*, **68**, 617–626.
- Mitsui, J., Takahashi, Y., Goto, J., Tomiyama, H., Ishikawa, S., Yoshino, H., Minami, N., Smith, D.I., Lesage, S., Aburatani, H. et al. (2010) Mechanisms of genomic instabilities underlying two common fragile-site-associated loci, PARK2 and DMD, in germ cell and cancer cell lines. *Am. J. Hum. Genet.*, **87**, 75–89.
- Hattori, N., Kitada, T., Matsumine, H., Asakawa, S., Yamamura, Y., Yoshino, H., Kobayashi, T., Yokochi, M., Wang, M., Yoritaka, A. et al. (1998) Molecular genetic analysis of a novel Parkin gene in Japanese families with autosomal recessive juvenile parkinsonism: evidence for variable homozygous deletions in the Parkin gene in affected individuals. *Ann. Neurol.*, **44**, 935–941.
- Hedrich, K., Djarmati, A., Schafer, N., Hering, R., Wellenbrock, C., Weiss, P.H., Hilker, R., Vieregge, P., Ozelius, L.J., Heutink, P. et al. (2004) DJ-1 (PARK7) mutations are less frequent than Parkin (PARK2) mutations in early-onset Parkinson disease. *Neurology*, **62**, 389–394.
- Hedrich, K., Eskelson, C., Wilmot, B., Marder, K., Harris, J., Garrels, J., Meija-Santana, H., Vieregge, P., Jacobs, H., Bressman, S.B. et al. (2004) Distribution, type, and origin of Parkin mutations: review and case studies. *Mov. Disord.*, **19**, 1146–1157.
- Abbas, N., Lucking, C.B., Ricard, S., Durr, A., Bonifati, V., De Michele, G., Bouley, S., Vaughan, J.R., Gasser, T., Marconi, R. et al. (1999) A wide variety of mutations in the parkin gene are responsible for autosomal recessive parkinsonism in Europe. French Parkinson's Disease Genetics Study Group and the European Consortium on Genetic Susceptibility in Parkinson's Disease. *Hum. Mol. Genet.*, **8**, 567–574.
- West, A., Periquet, M., Lincoln, S., Lucking, C.B., Nicholl, D., Bonifati, V., Rawal, N., Gasser, T., Lohmann, E., Deleuze, J.F.

- et al. (2002) Complex relationship between Parkin mutations and Parkinson disease. *Am. J. Med. Genet.*, **114**, 584–591.
17. Hedrich, K., Kann, M., Lanthaler, A.J., Dalski, A., Eskelson, C., Landt, O., Schwinger, E., Vieregge, P., Lang, A.E., Breakefield, X.O. et al. (2001) The importance of gene dosage studies: mutational analysis of the parkin gene in early-onset parkinsonism. *Hum. Mol. Genet.*, **10**, 1649–1656.
  18. Nuytemans, K., Theuns, J., Cruts, M. and Van Broeckhoven, C. (2010) Genetic etiology of Parkinson disease associated with mutations in the SNCA, PARK2, PINK1, PARK7, and LRRK2 genes: a mutation update. *Hum. Mutat.*, **31**, 763–780.
  19. Klein, C., Lohmann-Hedrich, K., Rogaeva, E., Schlossmacher, M.G. and Lang, A.E. (2007) Deciphering the role of heterozygous mutations in genes associated with parkinsonism. *Lancet Neurol.*, **6**, 652–662.
  20. Kay, D.M., Stevens, C.F., Hamza, T.H., Montimurro, J.S., Zabetian, C.P., Factor, S.A., Samii, A., Griffith, A., Roberts, J.W., Molho, E.S. et al. (2010) A comprehensive analysis of deletions, duplications, and copy number variations in PARK2. *Neurology*, **75**, 1189–1194.
  21. Kong, A., Masson, G., Frigge, M.L., Gylfason, A., Zusmanovich, P., Thorleifsson, G., Olason, P.I., Ingason, A., Steinberg, S., Rafnar, T. et al. (2008) Detection of sharing by descent, long-range phasing and haplotype imputation. *Nat. Genet.*, **40**, 1068–1075.
  22. Kong, A., Steinthorsdottir, V., Masson, G., Thorleifsson, G., Sulem, P., Besenbacher, S., Jonasdottir, A., Sigurdsson, A., Kristinsson, K.T., Frigge, M.L. et al. (2009) Parental origin of sequence variants associated with complex diseases. *Nature*, **462**, 868–874.
  23. Kong, A., Thorleifsson, G., Gudbjartsson, D.F., Masson, G., Sigurdsson, A., Jonasdottir, A., Walters, G.B., Gylfason, A., Kristinsson, K.T., Gudjonsson, S.A. et al. (2010) Fine-scale recombination rate differences between sexes, populations and individuals. *Nature*, **467**, 1099–1103.
  24. Holm, H., Gudbjartsson, D.F., Sulem, P., Masson, G., Helgadóttir, H.T., Zanon, C., Magnusson, O.T., Helgason, A., Sæmundsdóttir, J., Gylfason, A. et al. (2011) A rare variant in MYH6 is associated with high risk of sick sinus syndrome. *Nat. Genet.*, **43**, 316–320.
  25. Sulem, P., Gudbjartsson, D.F., Walters, G.B., Helgadóttir, H.T., Helgason, A., Gudjonsson, S.A., Zanon, C., Besenbacher, S., Bjornsdóttir, G., Magnusson, O.T. et al. (2011) Identification of low-frequency variants associated with gout and serum uric acid levels. *Nat. Genet.*, **43**, 1127–1130.
  26. Kay, D.M., Moran, D., Moses, L., Poorkaj, P., Zabetian, C.P., Nutt, J., Factor, S.A., Yu, C.E., Montimurro, J.S., Keefe, R.G. et al. (2007) Heterozygous parkin point mutations are as common in control subjects as in Parkinson's patients. *Ann. Neurol.*, **61**, 47–54.
  27. Lincoln, S.J., Maraganore, D.M., Lesnick, T.G., Bounds, R., de Andrade, M., Bower, J.H., Hardy, J.A. and Farrer, M.J. (2003) Parkin variants in North American Parkinson's disease: cases and controls. *Mov. Disord.*, **18**, 1306–1311.
  28. Bruggemann, N., Mitterer, M., Lanthaler, A.J., Djarmati, A., Hagenah, J., Wiegers, K., Winkler, S., Pawlack, H., Lohnau, T., Pramstaller, P.P. et al. (2009) Frequency of heterozygous Parkin mutations in healthy subjects: need for careful prospective follow-up examination of mutation carriers. *Park. Relat. Disord.*, **15**, 425–429.
  29. Lucking, C.B., Chesneau, V., Lohmann, E., Verpillat, P., Dulac, C., Bonnet, A.M., Gasparini, F., Agid, Y., Durr, A. and Brice, A. (2003) Coding polymorphisms in the parkin gene and susceptibility to Parkinson disease. *Arch. Neurol.*, **60**, 1253–1256.
  30. Sironi, F., Primignani, P., Zini, M., Tunesi, S., Ruffmann, C., Ricca, S., Brambilla, T., Antonini, A., Tesel, S., Canesi, M. et al. (2008) Parkin analysis in early onset Parkinson's disease. *Park. Relat. Disord.*, **14**, 326–333.
  31. Oliveira, S.A., Scott, W.K., Martin, E.R., Nance, M.A., Watts, R.L., Hubble, J.P., Koller, W.C., Pahwa, R., Stern, M.B., Hiner, B.C. et al. (2003) Parkin mutations and susceptibility alleles in late-onset Parkinson's disease. *Ann. Neurol.*, **53**, 624–629.
  32. Khan, N.L., Brooks, D.J., Pavese, N., Sweeney, M.G., Wood, N.W., Lees, A.J. and Piccini, P. (2002) Progression of nigrostriatal dysfunction in a parkin kindred: an [<sup>18</sup>F]dopa PET and clinical study. *Brain*, **125**, 2248–2256.
  33. Ruffmann, C., Zini, M., Goldwurm, S., Bramerio, M., Spinello, S., Rusconi, D., Gambacorta, M., Tagliavini, F., Pezzoli, G. and Giaccone, G. (2012) Lewy body pathology and typical Parkinson disease in a patient with a heterozygous (R275W) mutation in the Parkin gene (PARK2). *Acta Neuropathol.*, **123**, 901–903.
  34. Cookson, M.R., Lockhart, P.J., McLendon, C., O'Farrell, C., Schlossmacher, M. and Farrer, M.J. (2003) RING finger 1 mutations in Parkin produce altered localization of the protein. *Hum. Mol. Genet.*, **12**, 2957–2965.
  35. Bardien, S., Keyser, R., Yako, Y., Lombard, D. and Carr, J. (2009) Molecular analysis of the parkin gene in South African patients diagnosed with Parkinson's disease. *Park. Relat. Disord.*, **15**, 116–121.
  36. Wang, M., Hattori, N., Matsumine, H., Kobayashi, T., Yoshino, H., Morioka, A., Kitada, T., Asakawa, S., Minoshima, S., Shimizu, N. et al. (1999) Polymorphism in the parkin gene in sporadic Parkinson's disease. *Ann. Neurol.*, **45**, 655–658.
  37. Pankratz, N., Kissell, D.K., Pauciulo, M.W., Halter, C.A., Rudolph, A., Pfeiffer, R.F., Marder, K.S., Foroud, T. and Nichols, W.C. (2009) Parkin dosage mutations have greater pathogenicity in familial PD than simple sequence mutations. *Neurology*, **73**, 279–286.
  38. Pankratz, N., Dumitriu, A., Hetrick, K.N., Sun, M., Latourelle, J.C., Wilk, J.B., Halter, C., Doheny, K.F., Gusella, J.F., Nichols, W.C. et al. (2011) Copy number variation in familial Parkinson disease. *PLoS One*, **6**, e20988.
  39. Brooks, J., Ding, J., Simon-Sanchez, J., Paisan-Ruiz, C., Singleton, A.B. and Scholz, S.W. (2009) Parkin and PINK1 mutations in early-onset Parkinson's disease: comprehensive screening in publicly available cases and control. *J. Med. Genet.*, **46**, 375–381.
  40. Lesage, S., Lohmann, E., Tison, F., Durif, F., Durr, A. and Brice, A. (2008) Rare heterozygous parkin variants in French early-onset Parkinson disease patients and controls. *J. Med. Genet.*, **45**, 43–46.
  41. Guerrero Camacho, J.L., Monroy Jaramillo, N., Yescas Gomez, P., Rodriguez Violante, M., Boll Woehrlén, C., Alonso Vilatela, M.E. and Lopez Lopez, M. (2012) High frequency of Parkin exon rearrangements in Mexican-mestizo patients with early-onset Parkinson's disease. *Mov. Disord.*, **27**, 1047–1051.
  42. Nuytemans, K., Meeus, B., Crosiers, D., Brouwers, N., Goossens, D., Engelborghs, S., Pals, P., Pickut, B., Van den Broeck, M., Corsmit, E. et al. (2009) Relative contribution of simple mutations versus copy number variations in five Parkinson disease genes in the Belgian population. *Hum. Mutat.*, **30**, 1054–1061.
  43. Simon-Sanchez, J., Scholz, S., Matarin Mdel, M., Fung, H.C., Hernandez, D., Gibbs, J.R., Britton, A., Hardy, J. and Singleton, A. (2008) Genomewide SNP assay reveals mutations underlying Parkinson disease. *Hum. Mutat.*, **29**, 315–322.

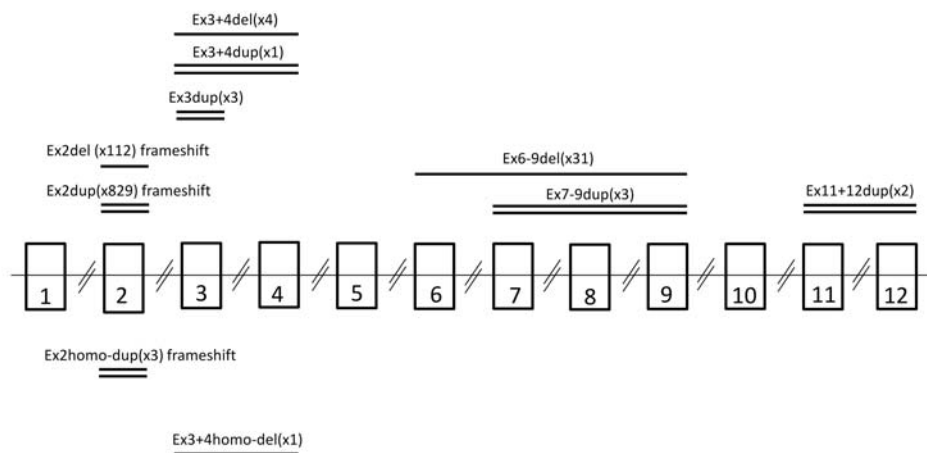


44. Wang, Y., Clark, L.N., Louis, E.D., Mejia-Santana, H., Harris, J., Cote, L.J., Waters, C., Andrews, H., Ford, B., Frucht, S. et al. (2008) Risk of Parkinson disease in carriers of parkin mutations: estimation using the kin-cohort method. *Arch. Neurol.*, **65**, 467–474.
45. Clark, L.N., Afridi, S., Karlins, E., Wang, Y., Mejia-Santana, H., Harris, J., Louis, E.D., Cote, L.J., Andrews, H., Fahn, S. et al. (2006) Case-control study of the parkin gene in early-onset Parkinson disease. *Arch. Neurol.*, **63**, 548–552.
46. Kim, S.Y., Seong, M.W., Jeon, B.S., Ko, H.S., Kim, J.Y. and Park, S.S. (2012) Phase analysis identifies compound heterozygous deletions of the PARK2 gene in patients with early-onset Parkinson disease. *Clin. Genet.*, **82**, 77–82.
47. Wang, L., Nuytemans, K., Bademci, G., Jauregui, C., Martin, E. R., Scott, W.K., Vance, J.M. and Zuchner, S. (2013) High-resolution survey in familial Parkinson disease genes reveals multiple independent copy number variation events in PARK2. *Hum. Mutat.*, **34**, 1071–1074.
48. Klein, C., Pramstaller, P.P., Kis, B., Page, C.C., Kann, M., Leung, J., Woodward, H., Castellan, C.C., Scherer, M., Vieregge, P. et al. (2000) Parkin deletions in a family with adult-onset, tremor-dominant parkinsonism: expanding the phenotype. *Ann. Neurol.*, **48**, 65–71.
49. Pramstaller, P.P., Kis, B., Eskelson, C., Hedrich, K., Scherer, M., Schwinger, E., Breakefield, X.O., Kramer, P.L., Ozelius, L.J. and Klein, C. (2002) Phenotypic variability in a large kindred (Family LA) with deletions in the parkin gene. *Mov. Disord.*, **17**, 424–426.
50. Farrer, M., Chan, P., Chen, R., Tan, L., Lincoln, S., Hernandez, D., Forno, L., Gwinn-Hardy, K., Petrucelli, L., Hussey, J. et al. (2001) Lewy bodies and parkinsonism in families with parkin mutations. *Ann. Neurol.*, **50**, 293–300.
51. Grünewald, A., Kasten, M., Ziegler, A. and Klein, C. (2013) Next-generation phenotyping using the parkin example: time to catch up with genetics. *JAMA Neurol.*, **70**, 1186–1191.
52. Hedrich, K., Marder, K., Harris, J., Kann, M., Lynch, T., Mejia-Santana, H., Pramstaller, P.P., Schwinger, E., Bressman, S.B., Fahn, S. et al. (2002) Evaluation of 50 probands with early-onset Parkinson's disease for Parkin mutations. *Neurology*, **58**, 1239–1246.
53. Sun, M., Latourelle, J.C., Wooten, G.F., Lew, M.F., Klein, C., Shill, H.A., Golbe, L.I., Mark, M.H., Racette, B.A., Perlmutter, J.S. et al. (2006) Influence of heterozygosity for parkin mutation on onset age in familial Parkinson disease: the GenePD study. *Arch. Neurol.*, **63**, 826–832.
54. Foroud, T., Uniacke, S.K., Liu, L., Pankratz, N., Rudolph, A., Halter, C., Shults, C., Marder, K., Conneally, P.M. and Nichols, W.C. (2003) Heterozygosity for a mutation in the parkin gene leads to later onset Parkinson disease. *Neurology*, **60**, 796–801.
55. Marder, K.S., Tang, M.X., Mejia-Santana, H., Rosado, L., Louis, E.D., Comella, C.L., Colcher, A., Siderowf, A.D., Jennings, D., Nance, M.A. et al. (2010) Predictors of parkin mutations in early-onset Parkinson disease: the consortium on risk for early-onset Parkinson disease study. *Arch. Neurol.*, **67**, 731–738.
56. Hiller, A., Hagenah, J.M., Djarmati, A., Hedrich, K., Reetz, K., Schneider-Gold, C., Kress, W., Munchau, A. and Klein, C. (2007) Phenotypic spectrum of PINK1-associated parkinsonism in 15 mutation carriers from 1 family. *Mov. Disord.*, **22**, 145–147.
57. Hilker, R., Klein, C., Ghaemi, M., Kis, B., Strotmann, T., Ozelius, L.J., Lenz, O., Vieregge, P., Herholz, K., Heiss, W.D. et al. (2001) Positron emission tomographic analysis of the nigrostriatal dopaminergic system in familial parkinsonism associated with mutations in the parkin gene. *Ann. Neurol.*, **49**, 367–376.
58. Khan, N.L., Scherfler, C., Graham, E., Bhatia, K.P., Quinn, N., Lees, A.J., Brooks, D.J., Wood, N.W. and Piccini, P. (2005) Dopaminergic dysfunction in unrelated, asymptomatic carriers of a single parkin mutation. *Neurology*, **64**, 134–136.
59. Klein, C., Djarmati, A., Hedrich, K., Schafer, N., Scaglione, C., Marchese, R., Kock, N., Schule, B., Hiller, A., Lohnau, T. et al. (2005) PINK1, Parkin, and DJ-1 mutations in Italian patients with early-onset parkinsonism. *Eur. J. Hum. Genet.*, **13**, 1086–1093.
60. Sidransky, E. (2006) Heterozygosity for a Mendelian disorder as a risk factor for complex disease. *Clin. Genet.*, **70**, 275–282.
61. Von Coelln, R., Dawson, V.L. and Dawson, T.M. (2004) Parkin-associated Parkinson's disease. *Cell Tissue Res.*, **318**, 175–184.
62. Colhoun, H.M., McKeigue, P.M. and Davey Smith, G. (2003) Problems of reporting genetic associations with complex outcomes. *Lancet*, **361**, 865–872.
63. Hicks, A.A., Petursson, H., Jonsson, T., Stefansson, H., Johannsdottir, H.S., Sainz, J., Frigge, M.L., Kong, A., Gulcher, J.R., Stefansson, K. et al. (2002) A susceptibility gene for late-onset idiopathic Parkinson's disease. *Ann. Neurol.*, **52**, 549–555.
64. Sveinbjornsdottir, S., Hicks, A.A., Jonsson, T., Petursson, H., Gudmundsson, G., Frigge, M.L., Kong, A., Gulcher, J.R. and Stefansson, K. (2000) Familial aggregation of Parkinson's disease in Iceland. *N. Engl. J. Med.*, **343**, 1765–1770.
65. Gudmundsson, K.R. (1967) A clinical survey of parkinsonism in Iceland. *Acta Neurol. Scand.*, **43** (Suppl. 33), 1–61.
66. Morris, J.N., Hawes, C., Fries, B.E., Phillips, C.D., Mor, V., Katz, S., Murphy, K., Drugovich, M.L. and Friedlob, A.S. (1990) Designing the national resident assessment instrument for nursing homes. *Gerontologist*, **30**, 293–307.
67. Hoehn, M.M. and Yahr, M.D. (2001) Parkinsonism: onset, progression, and mortality. *Neurology*, **57**, S11–S26.
68. Gulcher, J.R., Kristjansson, K., Gudbjartsson, H. and Stefansson, K. (2000) Protection of privacy by third-party encryption in genetic research in Iceland. *Eur. J. Hum. Genet.*, **8**, 739–742.
69. Wang, K., Li, M., Hadley, D., Liu, R., Glessner, J., Grant, S.F., Hakonarson, H. and Bucan, M. (2007) PennCNV: an integrated hidden Markov model designed for high-resolution copy number variation detection in whole-genome SNP genotyping data. *Genome Res.*, **17**, 1665–1674.
70. Ng, P.C. and Henikoff, S. (2003) SIFT: predicting amino acid changes that affect protein function. *Nucleic Acids Res.*, **31**, 3812–3814.
71. Adzhubei, I.A., Schmidt, S., Peshkin, L., Ramensky, V.E., Gerasimova, A., Bork, P., Kondrashov, A.S. and Sunyaev, S.R. (2010) A method and server for predicting damaging missense mutations. *Nat. Methods*, **7**, 248–249.

# Heterozygote carriers for CNVs in *PARK2* are at risk for PD

Johanna Huttenlocher<sup>1,2</sup>, Hreinn Stefansson<sup>1</sup>, Stacy Steinberg<sup>1</sup>, Hafdis T. Helgadottir<sup>1</sup>,  
Sigurlaug Sveinbjörnsdóttir<sup>3,4,5</sup>, Olaf Riess<sup>2</sup>, Peter Bauer<sup>2</sup>, Kari Stefansson<sup>1,6,\*</sup>

## Supplementary Material



Supplementary figure 1: Schematic representation of distribution and number of *PARK2* CNVs identified in this study. Shown are the 12 exons of the *PARK2* gene. CNVs detected in the heterozygous state are indicated above the sequence. CNVs detected in the homozygous state are shown below the sequence. The number of CNV carriers is given in parentheses.



(b)

	chr6:161889978	rs34424986	rs1801474	rs55774500	rs149953814	rs55830907	rs1801334	rs1801582	exon2 deletions	exon2 duplications	large deletions	large duplications	total number
chr6:161889978	0												93698
rs34424986	0	2											93698
rs1801474	0	16	24										93698
rs55774500	0	5	33	12									93698
rs149953814	0	0	7	3	0								93698
rs55830907	0	0	0	0	0	0							93698
rs1801334	0	18	367	161	1	1	50						93698
rs1801582	4	193	1251	669	45	12	840	2061					93698
exon2 deletions	0	0	3	2	0	0	5	17	0				104320
exon2 duplications	0	8	8	16	0	1	20	223	0	2			104320
large deletions	0	0	0	1	0	0	0	29	0	0	0		104320
large duplications	0	0	0	0	0	0	0	3	0	0	0	0	104320

(c)

	chr6:161889978	rs34424986	rs1801474	rs55774500	rs149953814	rs55830907	rs1801334	rs1801582	exon2 deletions	exon2 duplications	large deletions	large duplications	total number
chr6:161889978	0												39046
rs34424986	0	2											39046
rs1801474	0	3	11										39046
rs55774500	0	2	14	9									39046
rs149953814	0	0	4	2	0								39046
rs55830907	0	0	0	0	0	0							39046
rs1801334	0	5	173	77	0	0	27						39046
rs1801582	2	86	534	272	18	5	347	906					39046
exon2 deletions	0	0	3	0	0	0	1	8	0				40474
exon2 duplications	0	0	3	3	0	1	14	99	0	0			40474
large deletions	0	0	0	0	0	0	0	10	0	0	0		40474
large duplications	0	0	0	0	0	0	0	2	0	0	0	0	40474

**Supplementary table 2:** Frequency of *PARK2* CNVs. Comparable frequency of *PARK2* CNVs was found in samples genotyped using IlluminaHumanHap and IlluminaOmni genotyping chips.

	Genotyped subjects	<i>PARK2</i> CNV carriers	Frequency of <i>PARK2</i> CNVs	Chi-square test
IlluminaHumanHap	N = 56,846	N = 519	0.0091	p-value = 0.2618
IlluminaOmni	N = 48,887	N = 480	0.0098	
Illumina chips total	N = 105,733	N = 999	0.0094	

**Supplementary table 3: *PARK2* sequencing primers**

Exon	Forward/ Reverse	Sequenz (5' → 3')
1	fw	GGAGGCCCTGGAGGATTTAAC
	rev	CATTGACAGTTGGCACCG
2	fw	CTTCCAATTCCTTGGTCAG
	rev	TCCCCAGGCACATATTTATCT
3	fw	TTTTTTACCTTGCTCCC
	rev	TAAATATGCACCCGGTGAGG
4	fw	ACAAGCTTTTAAAGAGTTTCTTGT
	rev	AGGCAATGTGTTAGTACACA
5	fw	ACATGTCTTAAGGAGTACATTT
	rev	TCTCTAATTTCTGGCAACA
6	fw	CGGGATCCATGTGTGTA
	rev	GGGGGAGTGATGCTATTTTTAG
7	fw	TGCCTTTCCACACTGACAGGTA
	rev	TCTGTTCTTCATTAGCATTAGAGA
8	fw	TTCTAAAGAGGTGCGGTTGG
	rev	GTTTTCTTCCCATTTCAGG
9	fw	GGGTGAAATTTGCAGTCAGT
	rev	AATAATAATCCCAGCCCATGTGCA
10	fw	TTGCCAAAATGCAACCTAATG
	rev	CTCTCCATGACCTCCAGGAA
11	fw	GAACATAAAGCTCTGATCCCAGTAATAG
	rev	TGGAATCCCTGATGGGTATG
12	fw	GTTTGGGAATGCGTGTTTT
	rev	AGAATTAGAAAATGAAGGTAGACA





## Erklärung zum Anteil an gemeinschaftlichen Veröffentlichungen

Die in dieser Dissertation dargestellten und diskutierten Ergebnisse und Schlussfolgerungen sind in zwei Originalarbeiten, welche in „Journal of Medical Genetics“ (2014) und „Human Molecular Genetics“ (2015) publiziert wurden, zusammengefasst.

1. **Huttenlocher, J.**, Krüger, R., Capetian, P., Lohmann, K., Brockmann, K., Csoti, I., ... Bauer, P. (2015). EIF4G1 is neither a strong nor a common risk factor for Parkinson's disease: evidence from large European cohorts. *Journal of Medical Genetics*, 52(1), 37–41. doi:10.1136/jmedgenet-2014-102570
2. **Huttenlocher, J.**, Stefansson, H., Steinberg, S., Helgadottir, H. T., Sveinbjörnsdóttir, S., Riess, O., & Bauer, P., Stefansson, K. (2015). Heterozygote carriers for CNVs in PARK2 are at increased risk of Parkinson's disease. *Human Molecular Genetics*, ddv277–. doi:10.1093/hmg/ddv277

Die Projekte wurden von mir unter wissenschaftlicher Anleitung durch Prof. Dr. Peter Bauer und Dr. Hreinn Stefansson selbstständig bearbeitet. Folgende Beiträge zu den vorliegenden Publikationen wurden von mir ausgeführt:

1. **Huttenlocher, J.**, Krüger, R., Capetian, P., Lohmann, K., Brockmann, K., Csoti, I., ... Bauer, P. (2015). EIF4G1 is neither a strong nor a common risk factor for Parkinson's disease: evidence from large European cohorts. *Journal of Medical Genetics*, 52(1), 37–41. doi:10.1136/jmedgenet-2014-102570

Die Konzeptplanung und Organisation des Projekts wurde von mir durchgeführt. Die Hochauflösende Schmelzkurvenanalyse am LightCycler480-System des *EIF4G1*-Gens (Exon 10 und Exon 24) bei den 2.051 nicht verwandten, sporadischen Patienten europäischen Ursprungs, sowie die Validierungen auffälliger Proben durch konventionelle Sanger-Sequenzierung wurden von mir durchgeführt, ausgewertet und diskutiert.

Die Sanger-Sequenzierung des gesamten kodierenden Bereichs von 95 Patienten mit familiengebundenem PD aus Deutschland (n=80) und Island (n=15) wurde von deCODE Genetics durchgeführt. Anschließend wurden die Ergebnisse von mir eigenständig ausgewertet und diskutiert.

Die p.Arg1205His Variante wurde im Zuge des Gesamt-Genom-Sequenzierungs-Projektes bei deCODE Genetics detektiert. Die Verifizierung von 200 Proben mit imputierter rs112176450 Variante mittels direkter Genotypisierung (Nanogen) sowie die bioinformatische Analyse und Testung auf Assoziation wurde von deCODE Genetics durchgeführt. Diese Vorgänge wurden von mir koordiniert, dokumentiert und überwacht. Darüber hinaus habe ich die Genotyp-Rohdaten analysiert und die Ergebnisse der Assoziationsstudie interpretiert.

Die p.Arg566Cys Variante im Exon 12 des *EIF4G1*-Gens in einem Patienten mit familiengebundener PD wurde von mir identifiziert. Ich habe die anschließende Sanger-Sequenzierung des Exon 12 der betroffenen Familienmitglieder durchgeführt sowie die Fragmentlängenanalyse mittels Kapillarelektrophorese zur Ermittlung der Haplotypen. Die Ergebnisse wurden von mir ausgewertet und diskutiert. Darüber hinaus habe ich das gesamte Manuskript verfasst.

2. **Huttenlocher, J.**, Stefansson, H., Steinberg, S., Helgadottir, H. T., Sveinbjörnsdóttir, S., Riess, O., Bauer, P., & Stefansson, K. (2015). Heterozygote carriers for CNVs in *PARK2* are at increased risk of Parkinson's disease. *Human Molecular Genetics*, *ddv277*–. doi:10.1093/hmg/ddv277

Die Konzeptplanung und Organisation des Projekts wurde von mir durchgeführt.

Die CNVs wurden mittels des Computerprogramms PennCNV-*Caller* durch deCODE Genetics identifiziert. Anschließend wurde von mir eine visuelle Kontrolle durchgeführt. Zusätzliche CNV-Träger wurden durch eine von mir durchgeführte Stammbaumanalysen der CNV-Träger identifiziert. Die bioinformatische Analyse und Testung auf Assoziation wurden von deCODE Genetics durchgeführt und die Ergebnisse anschließend von mir analysiert und interpretiert.

Die *Missense*-Mutationen im *PARK2*-Gen wurden im Zuge des Gesamt-Genom-Sequenzierungs-Projektes detektiert. Die zusätzliche Sanger-Sequenzierung aller Patienten mit CNV oder Arg275Trp Mutation wurde von deCODE Genetics durchgeführt und die Ergebnisse von mir analysiert und interpretiert.

Alle Vorgänge wurden von mir koordiniert, dokumentiert und überwacht. Darüber hinaus habe ich die Genotyp-Rohdaten analysiert und die Ergebnisse der Testung auf Assoziation interpretiert. Statistische Auswertungen habe ich unter der Anleitung von Frau Dr. Stacy Steinberg durchgeführt.

Die Sammlung relevanter Studien für die Meta-Analyse habe ich eigenständig durchgeführt. Hierfür habe ich alle Zeitschriftenartikel der Medline Datenbank (<http://www.ncbi.nlm.nih.gov/pubmed/>) und der PD Mutation Datenbank ([www.molgen.ua.ac.be/PDmutDB](http://www.molgen.ua.ac.be/PDmutDB)) nach Artikeln durchsucht, die Daten über heterozygote *PARK2* CNV-Träger beinhalten. Die Ermittlung der studienspezifischen OR und des 95 % Konfidenzintervall mittels *Fisher's exact test* habe ich eigenständig durchgeführt, sowie die gepoolte Analyse. Die Analyse der isländischen Daten (Kontrolle der Verwandtschaft und Populationsstratifikation innerhalb der Isländischen Fall- und Kontrollgruppen) wurde durch Frau Dr. Stacy Steinberg vorgenommen. Außerdem wurde die Abbildung (Supplementary figure: Schematic representation of distribution and number of *PARK2* CNVs identified in this study) im Ergänzungsteil der Publikation von mir erstellt. Darüber hinaus habe ich das gesamte Manuskript verfasst.



## **Danksagungen**

An dieser Stelle möchte ich mich bei Herrn Prof. Dr. Olaf Rieß für die Ermöglichung dieser Promotionsarbeit im Institut für Medizinische Genetik und angewandte Genomik bedanken.

Herrn Prof. Dr. Peter Bauer gilt mein besonderer Dank für die interessante und spannende Themenstellung und seine wissenschaftliche Betreuung. Ich möchte mich für seine wertvollen Ratschläge, konstruktiven Anregungen und die Geduld bei immer neuen Fragen und Problemen bedanken.

Darüber hinaus möchte ich Prof. Dr. Olaf Rieß, Prof. Dr. Peter Bauer und Dr. Holm Graessner danken, dass sie mir den Aufenthalt in Island und die Arbeit bei deCODE Genetics ermöglicht haben. Die gesammelten Erfahrungen, sind ein großer Gewinn für meinen weiteren beruflichen Weg.

Mein aufrichtiger Dank gilt auch Herrn Dr. Hreinn Stefánsson, der mich während meines Aufenthalts in Island mit sehr viel Geduld betreute. Ich danke ihm für seine Einführung in die Thematik, seine Ideen und Diskussionen, die mir beim Erstellen der Arbeit eine große Hilfe waren.

Nicht zuletzt möchte ich mich bei ihm und seiner Familie für die beeindruckende Gastfreundschaft bedanken. Seine liebevolle Aufnahme hat sehr dazu beigetragen, dass ich mich in Island vom ersten Tag an Zuhause gefühlt habe.

Danke sage ich auch den Kooperationspartnern, für die Bereitstellung der Patientendaten sowie den Patienten und Familien für die Teilnahme an den Studien.

Ganz besonders danke ich meinen Kollegen der Arbeitsgruppe „Genomik“ Claudia Dufke, Florian Harmuth, Claudia Bauer, Sandra Junker, Juliane Gohlke, Beate Kootz, Nina Schlipf, und Marc Sturm, sowie Hjörvar Pétursson, Hafþís Helgadóttir, Ómar Gústafsson, Íris Guðjónsdóttir, Ásgeir Sigurðsson, Hrefna Jóhannsdóttir, Bragi Walters und Stacy Steinberg der Arbeitsgruppe „CNS“ bei deCODE Genetics für die freundliche und hilfsbereite Arbeits-

atmosphäre. Mein besonderer Dank gilt dabei Stacy für die wichtige Unterstützung bei der statistischen Auswertung der Daten und Ásgeir für seine Hilfe im Labor.

Hafdís, Nina und Maike Aurich standen mir als immer erreichbare und liebe Ansprechpartnerinnen, im Verlauf der Arbeit zur Seite. Ich danke Euch für die tolle Freundschaft.

Besonders möchte ich mich bei meinen Eltern und meiner Schwester bedanken, die mich auf meinem Weg immer wieder ermutigt und liebevoll unterstützt haben. Darüber hinaus möchte ich mich für die finanzielle Unterstützung meiner Eltern bedanken, die mir diesen Weg ermöglicht hat.

Ein besonderer Dank gilt Florian der meine Zeiten der überstrapazierten Nerven überstehen musste, und seiner liebevollen Betreuung unserer Tochter Pauline Lóa während meiner Zeit des Schreibens. Im Besonderen möchte ich ihm aber dafür danken, dass er während der Zeit in Island keine Mühen gescheut hat, mich so oft wie möglich zu besuchen und dafür, dass er diese wunderschöne Insel genauso wie ich als zweite Heimat in sein Herz geschlossen hat.



