

VU Research Portal

Cracking the code-ing sequence for Parkinson's disease

Jansen, I.E.

2017

document version

Publisher's PDF, also known as Version of record

[Link to publication in VU Research Portal](#)

citation for published version (APA)

Jansen, I. E. (2017). Cracking the code-ing sequence for Parkinson's disease. [PhD-Thesis - Research and graduation internal, Vrije Universiteit Amsterdam].

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal

Take down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

E-mail address:

vuresearchportal.ub@vu.nl

NEDERLANDSE SAMENVATTING

De ziekte van Parkinson (ZvP) staat vooral bekend als de bewegingsstoornis van de ouderen. Echter, personen met een relatief jonge leeftijd (beneden de 45 jaar) kunnen deze neurodegeneratieve ziekte ook krijgen. De prevalentie varieert van 0.04% in de leeftijdsgroep 40 tot 49 jaar, tot 1.9% voor personen met een leeftijd boven de 80 jaar. De ziekte wordt veroorzaakt door neurodegeneratie van de hersenen, waarbij vooral de substantia nigra aangedaan is. Deze neurodegeneratie zorgt voor bewegingsproblemen, tremoren en verstarring of expressieloze gezichtskenmerken. Additionele non-motorische symptomen includeren constipatie, slaapproblemen, verlies van reukvermogen en depressie. Deze tweede type symptomen kunnen tot wel tientallen jaren voorafgaand aan de motorische symptomen aanwezig zijn. Vaak worden de symptomen dan nog niet geassocieerd aan de ZvP. De diagnose van Parkinson is complex, wat grotendeels veroorzaakt wordt door de heterogeniteit van de beschreven symptomen, maar ook van de pathologie, aangedane mechanisme en genetica.

Het verbeteren van onze kennis over de ontwikkeling van de ziekte zal diagnoses verbeteren, maar nog belangrijker, zal ons steeds dichterbij een genezend medicijn brengen. Momenteel kan de toegepaste medicatie alleen symptomen verlichten, maar niet het ziektebeloop vertragen. Nog een stapje verder dan vertraging van het ziekteproces is het ultieme doel om een behandeling te ontwikkelen die de schade van de ziekte ongedaan kan maken. Het is van cruciaal belang de genetische risico factoren voor de ZvP in kaart te brengen. Door vast te stellen welke genetische afwijkingen bijdragen aan de ontwikkeling van Parkinson, verbreden en verdiepen we ons begrip van de biologische processen die gedereguleerd zijn. Ontdekte genen en biologische processen kunnen eventueel dienen als 'targets' van nieuwe therapieën. Ondanks dat de onderzoekswereld al meer dan 20 verschillende genetische factoren gelinkt heeft aan de ZvP is er nog niet voldoende kennis om de juiste medicijnen te ontwikkelen. Met behulp van exome sequencing data van het 'International Parkinson's Disease Genetics Consortium' heb ik in dit proefschrift onderzoek gedaan naar de genetica van Parkinson.

De eerste twee hoofdstukken zijn gericht op al bekende Parkinson genen. **Hoofdstuk 2** exploreert de betrokkenheid van *CHCHD2*, een gen die in de Aziatische populatie gelinkt is aan de ontwikkeling van de ZvP. Om te bepalen of gelijke en andere varianten in de Europese populatie eenzelfde soort effect hebben, zijn er twee IPDGC exome datasets onderzocht. Wij observeerde één 'common' variant, identiek aan een variant in de Aziatische studie, in onze genotyping-array NeuroX dataset. Ondanks dat deze variant in de originele studie een significante associatie met de ZvP laat zien, was er in onze analyses geen significant effect, wat duidt op een Aziatisch-specificiteit of invalidatie van deze specifieke variant. Hoewel er geen associatie in de Europese populatie gedetecteerd is voor de 'common' varianten, identificeerde we wel 3 zeldzame nonsynonymous varianten in 4 Parkinson patiënten. In tegenstelling is er geen enkele variant geobserveerd

A

in de gezonde individuen. Deze resultaten suggereren dat *CHCHD2* een zeldzame risico factor kan zijn in de Europese populatie. Uitgebreid sequencing onderzoek van dit gen in grotere Europese datasets moet echter in de toekomst meer duidelijkheid scheppen.

De genetische literatuur van de ZvP heeft naast de mono-genetische vormen van Parkinson ook meer dan 20 ‘common’ loci geïdentificeerd die het risico verhogen om sporadische Parkinson te krijgen. Door dezelfde IPDGC exome datasets te analyseren, hebben we in **hoofdstuk 3** getest of zeldzame coderende varianten die binnen de ‘common’ loci vallen, ook een invloed hebben op het Parkinson risico. Voor het onderzoeken van individuele zeldzame varianten is het nodig om datasets met tienduizenden personen te analyseren om voldoende statistische ‘power’ te behalen voor het detecteren van een associatie signaal. Om de kans voor een zeldzame variant associatie te verhogen terwijl het aantal geteste personen gelijk blijft, zijn de varianten per gen of gen-set (bestaande uit literatuur-gebaseerde Parkinson loci) geaggregeerd om het gezamenlijke effect op de ZvP te testen. De literatuur gebaseerde loci representeren doorgaans een geassocieerd gebied die meerdere genen bevat, waarbij de locatie van de werkelijke causale variant (en het gerelateerde gen) onbekend is. Om het meest waarschijnlijke causale gen per locus te selecteren, hebben we gebruik gemaakt van PrixFixe: een algoritme gebaseerd op onderliggende functionele overeenkomsten tussen de genen van verschillende loci. Onze analyses bevestigden de betrokkenheid van het bekende Parkinson gen *LRRK2* voor zowel ‘common’ als zeldzame variant associaties. Verder laten de genen *STBD1* en *SPATA19* ook associaties zien met de ZvP in de IPDGC exome sequencing en NeuroX datasets. Zoals verwacht, werden er bij de gen-set analyses een gezamenlijk effect van ‘common’ varianten gedetecteerd, maar ontbrak een gezamenlijk effect van zeldzame varianten. Het is complex te concluderen of de afwezigheid van een zeldzame variant gen-set associatie het gevolg is van een eventuele onjuiste selectie van causale genen door PrixFixe, of dat zeldzame varianten binnen de bekende ‘common’ risico factoren werkelijk niet bijdragen aan de etiologie van de ZvP. Dit hoofdstuk benadrukt dan ook het gebrek aan nuttige functionele priorisatie methodes, die toereikend kunnen zijn in het benoemen van causale genen binnen ‘common’ genetische factoren terwijl ze de biologische eigenschappen van de ziekte in overweging nemen.

De volgende stap omvatte de zoektocht naar nieuwe genetische factoren door te focussen op groepen van genen die een functionele eigenschap hebben die gerelateerd is aan de ZvP. Het gen *GBA* functioneert in het lysosomale systeem en mutaties met een autosomale recessieve overervingspatroon zijn geobserveerd in families met de ziekte van Gaucher, welke een zeldzame lysosomale stapelingsziekte is (LSZ). Heterozygote varianten in *GBA* zijn daarnaast ook geïdentificeerd als sterke risico factoren voor de ZvP. Naast *GBA* zijn er 53 andere genen die geassocieerd zijn met LSZ. In **hoofdstuk 4** hebben we onderzocht of varianten in de overige LSZ genen ook een invloed hebben op het Parkinson risico. Door gebruik te maken van 2 onafhankelijk whole exome sequencing datasets en de

A

NeuroX dataset hebben we het gezamenlijke effect van zeldzame varianten in LSZ genen op de ZvP bestudeerd. Hierbij hebben we vooral de nadruk gelegd op de invloed van de totale LSZ genset, om de statistische 'power' voor het detecteren van een associatie signaal te verhogen. Binnen de IPDGC exome dataset detecteren we een gezamenlijk effect van zeldzame LSZ varianten, met een voorspeld hoge impact (door het CADD algoritme), op het risico om de ZvP te ontwikkelen. Dit signaal werd gerepliceerd in zowel de PPMI exome and NeuroX datasets. De associatie blijft aanwezig in de 3 onafhankelijke datasets wanneer het reeds bekende *GBA* gen verwijderd wordt van de LSZ gen-set. Dit scheidt de suggestie dat de overige LSZ genen ook van belang zijn voor de ontwikkeling van de ZvP. Wanneer we inzoomen op de gezamenlijke associatie van zeldzame varianten binnen individuele genen in the IPDGC exome dataset, observeren we een sterk geassocieerd *GBA* gen. Dit bevestigt een goede onderzoeksopzet van onze studie. De genen *SMPD1*, *CTSD* en *SLC17A5* laten ook significante invloeden op de ZvP zien, waarvan de associatie van het eerste gen al eerder besproken is in de context van de ZvP. De resultaten suggereren verder ook dat meerdere varianten in meerdere genen binnen één individu, het zo genaamde oligogenetische effect, het risico op de ZvP verhoogt. Uit dit hoofdstuk concluderen we dat niet-*GBA* varianten binnen de LSZ genen ook bijdragen aan de ontvankelijkheid van ZvP. Toekomstig sequencing onderzoek in grotere datasets en het functioneel opvolgen van deze genen in het laboratorium zal de precieze rol van deze genen aan het licht moeten brengen.

In **hoofdstuk 5** hebben we eenzelfde manier gebruikt om genetische variatie in gen-sets te onderzoeken, dit keer gefocust op genen die betrokken zijn bij biologische processen die aangedaan zijn in de pathobiologie van de ZvP. De selectie van biologische processen is gebaseerd op veel voorkomende mechanismen in Parkinson gerelateerde literatuur of op een specifiek gen-expressie onderzoek. Het laatstgenoemde onderzoek gebruikte post-mortem breinweefsel van Parkinson patiënten om genoom-wijde gen-expressie verschillen aan te duiden tussen verschillende 'Braak stages', welke verschillende pathologische stadia beschrijven. Door ons speciaal te richten op de vroege 'Braak stages' 1 en 2, ook wel de presymptomatische fase genoemd, beoogden we te bepalen of de geobserveerde verschillen op gen-expressie niveau het gevolg van genetische defecten kon zijn. Samengevat, 3 biologische processen (*mitochondriale dysfuncties*, *renin-angiotensin signalering* en *caveolar-bemiddelende endocytose*) lieten een genetische associatie met de ZvP zien in 3 onafhankelijke exome sequencing datasets. De sterkste associatie was gedetecteerd voor *mitochondriale dysfuncties*, zowel in de ontdekkingsdataset (IPDGC) als in de replicatie datasets (PPMI en Merck). Hoewel er geen sprake was van replicatie van genen binnen dit biologische mechanisme, zijn de meeste individuele gen associaties in genen die een functie hebben in een bepaald sub proces; namelijk in de mitochondriale complex I subeenheden. Het complex I heeft een belangrijke rol in de aanmaak van ATP en is daarom van belang voor de productie van

A

reactieve zuurstofdeeltjes. Onze resultaten suggereren dat deregulatie van dit specifieke proces aangedaan is in de ZvP. Functioneel onderzoek van mitochondriale complex I in het laboratorium moet de precieze link met de ZvP onthullen.

Een laatste strategie in **hoofdstuk 6** die zich niet baseerde op biologische assumpties over de ZvP, nam elk gen uit het genoom in overweging als een mogelijk genetisch risicofactor voor de ZvP. Alle coderende gebieden van het genoom werden systematisch getest voor de aanwezigheid van zogenoemde 'loss-of-function' (LoF) varianten die de functie van het bijbehorende gen ernstig kunnen beïnvloeden. Deze LoF varianten moesten, in vergelijking tot de gezonde controle groep, een hogere allelfrequentie in de Parkinson patiënten hebben. Alleen de homozygote en compound heterozygote LoF varianten werden als interessant beschouwd omdat de patiënten groep een relatief jonge onset leeftijd hebben, welke in de literatuur vooral geassocieerd wordt met een autosomaal recessief overervingspatroon. In totaal rapporteren we 27 genen die een LoF variant bevatten met een hogere allelfrequentie in de Parkinson patiënten. *GPATCH2L*, *UHRF1BP1L*, *PTPRH*, *ARSB* en *VPS13C* zijn de meest overtuigend genetische factoren die een invloed hebben op de ontwikkeling van de ZvP omdat functionele experimenten een verandering detecteren in Parkinson-gerelateerde fenotypes. Bovendien kunnen deze genen tot een zekere hoogte gerepliceerd worden in onafhankelijke genetische datasets. De identificatie van het eerder ontdekte Parkinson gen *VPS13C* benadrukt de goede opzet van deze studie. We zijn overtuigd dat deze 5 genen veelbelovende Parkinson genen zijn en moedigen het Parkinson onderzoeksveld aan om meer gedetailleerde functionele experimenten en grotere genetische validatie studies uit te voeren.

Het onderzoek dat in dit proefschrift is verricht draagt bij aan een beter begrip van de onderliggende genetische factoren die een invloed hebben op het risico om de ZvP te ontwikkelen. Ondanks de nieuwe bevindingen, heeft het Parkinson onderzoeksveld nog een lange weg te bewandelen voordat er een genezend medicijn bestaat. De genetische onderzoeksweld zal in de nabije toekomst genoom-wijd gaan sequencen, een techniek die de voorkeur geniet omdat het mogelijk maakt elke positie van het genoom in kaart te brengen. Samen met een vergroting van de datasets zullen er additionele genetische factoren ontdekt worden. De grootste uitdaging voor het genetische onderzoeksveld is op dit moment de functionele interpretatie van de genetische bevindingen. De al bekende en nieuwe genen zullen verder onderzocht moeten worden om te bepalen via welke weg ze de ZvP veroorzaken. Het is daarom cruciaal om data van verschillende 'omics' gebieden (transcriptoom, proteoom, epigenoom en metabooloom) te integreren. In parallel, zullen grote functionele screening methodes de causale varianten en genen linken aan de pathogenese van de ZvP. Wanneer de aangedane processen als gevolg van genetische defecten correct geïdentificeerd zijn, zal dit in een ultieme situatie bij gaan dragen aan de ontwikkeling van een geneesmiddel die het Parkinson ziekteproces tot stilstand brengt of in het meest optimale geval de lichamelijke schade ongedaan kan maken.

A