Cholesterol related genetics in sporadic Alzheimer's disease

Nathalie I. V. Nilsson
Integrated Program in Neuroscience
McGill University
Montreal, Quebec
August 2020

A thesis submitted to McGill University in partial fulfillment of the requirements for the degree of Doctor of Philosophy

Abstract

The sporadic – and most common – form of Alzheimer's disease (AD) is a multifactorial disease influenced by a multitude of both modifiable and non-modifiable risk factors. Among the most important non-modifiable factors are genetics and studies suggest that, even though sporadic, AD is highly heritable. Despite recent advances in identifying loci underlying this heritability, known loci only account for about a third of the genetic variance leaving the majority to still be accounted for. The discovery of these loci is important as they can contribute to the identification of disease-causing mechanisms that in turn can lead to the development of potential drugs to treat AD. As there is still no cure for AD, such new treatment options are sorely needed.

AD is also influenced by modifiable factors and amongst them are a group of factors relating to cardiovascular risk. Of special interest for this thesis is hypercholesterolemia as it pertains to cholesterol metabolism that has been further implicated in AD. For example, midlife hypercholesterolemia associates with increased risk of developing AD later in life while the use of cholesterol lowering statins have been shown to associate with reduced risk. In a similar way, a high cholesterol diet in animal models leads to increased Aβ load in the brain, and on a cellular level, elevation of cholesterol levels in neuronal cultures affect cells directly to increase Aβ production.

Taken together, much of the genetic variance in AD remains to be discovered and the identification of new loci could identify new, or provide support to already existing, disease-causing mechanisms. To this effect, with ample evidence implicating altered cholesterol metabolism as one of the mechanisms contributing to AD, this thesis aimed at investigating genetics relating to cholesterol metabolism in AD and was assessed in three studies.

In the first study, using whole-genome sequencing data, genetic variants in the gene encoding the rate-limiting step in cholesterol synthesis, *HMGCR*, was investigated in relation to both cholesterol

metabolism and AD. We identified a genetic variant, rs72633963, that showed evidence of a protective phenotype in two Quebec based cohorts; the A allele associated with reduced Aβ plaque load in post-mortem brain tissue and with better cognition in a pre-clinical AD cohort. This was accompanied by reductions in blood total cholesterol and LDL cholesterol while HDL cholesterol remained unchanged.

The second study investigated the cellular effects of a previously discovered *HMGCR* variant, rs3846662, whose AA genotype has been shown to associate with protection in AD. It is hypothesized to act by increasing alternative splicing of the *HMGCR* transcript, resulting in a transcript lacking exon 13 (Δ13-*HMGCR*). Induced pluripotent stem cells were produced from renal epithelial cells collected from either AA or GG carriers, that were then used for differentiation into neural progenitor cells and neurons. Effects of genotype and cell type were assessed on HMGCR RNA and protein expression, HMGCR activity, intracellular TAU levels, and extracellular Aβ peptides. We found that rs3846662 genotype influenced indices of HMGCR – AA carriers had lower levels of full length *HMGCR* and increased levels of Δ13-*HMGCR*, which was accompanied by increased protein levels, but had no effect on measures of TAU or Aβ.

Findings on rs72633963 (reported here) and rs3846662 (reported in the literature) indicate that they are protective in AD and associate with reduced blood total cholesterol, leading us to hypothesize that the effects on AD is mediated through their effect on cholesterol levels. Because findings from the first study was weak, we wanted to test this in a broader sense. In the third and final study we devised a polygenic score capturing the effect of multiple genetic variants associating with blood total cholesterol levels. We found that the score associated well with total cholesterol levels (explaining ~18% of the variance) and improved prediction of hypercholesterolemia. These relationships were strongly influenced by statin use and sex, such that the best effect was observed in statin free females. The score was further assessed for associations with AD, but we could not find any significant effect on either AD risk, AD biomarkers, or AD pathology.

In conclusion, this thesis provides insights into the role of cholesterol related genetics in AD. While the literature and our findings on rs72633963 and rs3846662 support a role for cholesterol related genetics in AD, the mechanism by which they act remain to be determined. We hypothesized a mechanism through peripheral blood cholesterol levels, but our negative findings with a polygenic score would suggest otherwise. These findings highlight the complexity of the relationship between cholesterol related genetics and AD and implores more research to determine the mechanisms of cholesterol related variants.

Résumé

La forme sporadique - la plus commune – de la maladie d'Alzheimer (MA) est une maladie multifactorielle influencée par une multitude de facteurs de risque autant modifiables que non-modifiables. Parmi les facteurs non-modifiables les plus importants, on trouve la génétique. Des études suggèrent que bien qu'elle soit sporadique, la MA est hautement héritable. Malgré de récentes avancées dans l'identification de *loci* sous-jacents à cette héritabilité, les *loci* connus ne comptent que pour le tiers de la variance génétique, laissant la majorité d'entre eux inconnue. La découverte de ces *loci* est importante étant donné qu'ils peuvent contribuer à l'identification de mécanismes causatifs de la maladie qui, à leur tour, peuvent mener au développement de médicaments potentiels pour traiter la MA. Comme il n'y a toujours pas de traitement étiologique pour la MA, de telles nouvelles cibles thérapeutiques sont grandement nécessaires.

La MA est également influencée par des facteurs modifiables et parmi eux, on trouve un groupe de facteurs reliés au risque cardiovasculaire. Étant donné que le métabolisme du cholestérol a été impliqué dans la MA, l'hypercholestérolémie est d'un intérêt particulier pour cette thèse. En effet, l'hypercholestérolémie dans la quarantaine est associée à un risque accru de développer la MA plus tard dans la vie, alors que l'utilisation de statines a été associée à un risque moindre. De façon similaire, un régime riche en cholestérol dans les modèles animaux mène à une quantité accrue d'Aβ dans le cerveau et au niveau cellulaire, l'élévation des niveaux de cholestérol dans des cultures neuronales affecte les cellules directement en augmentant la production d'Aβ.

Cela suggère que l'essentiel des facteurs de risque génétique liés à la MA reste à découvrir et l'identification de nouveau *loci* pourrait permettre d'identifier de nouveaux mécanismes causatifs ou supporter des mécanismes déjà existants. À cet effet, la présence de nombreuses évidences impliquant un métabolisme anormal du cholestérol dans l'étiologie de la MA, cette thèse visait, par l'entremise de 3 études, à évaluer la composante génétique qui relie t le métabolisme du cholestérol à la MA.

Dans la première étude, des données de séquençage du génome complet ont été utilisées pour identifier les variantes génétiques impliquées dans le gène codant pour l'étape limitante de la synthèse du cholestérol, soit l'*HMGCR*. Nous avons identifié une variante génétique, le rs72633963, qui semblait démontrer un phénotype protecteur dans deux cohortes québécoises. Bien qu'il ne soit pas directement relié aux concentrations sanguines de HDL, l'allèle A est associé à une quantité moindre d'Aβ dans des tissus cérébraux post-mortem, à une meilleure cognition dans une cohorte préclinique de MA, ainsi qu'à une réduction des niveaux de cholestérol total et de LDL dans le sang.

La seconde étude examinait les effets cellulaires d'une autre variante génétique préalablement découverte, rs3846662, dont le génotype AA a été associé à un effet protecteur contre la MA. Cette variante agirait en augmentant l'épissage alternatif du transcrit de l'*HMGCR*, résultant en un transcrit manquant l'exon 13 (Δ13-*HMGCR*). Des cellules souches pluripotentes induites ont été produites à partir de cellules rénales épithéliales prélevées chez des sujets porteurs des génotypes AA ou GG. Elles ont ensuite été différenciées en cellules neurales progénitrices et en neurones. Après avoir évalué les effets du génotype AA sur l'expression d'ARN et de protéine de HMGCR, l'activité de cet enzyme, les niveaux intracellulaires de tau et les concentrations extracellulaires d'Aβ, nous avons découvert que, bien qu'il n'ait aucun effet sur les mesures de tau et d'Aβ, le génotype rs3846662 et les porteurs des génotypes AA en particulier, montrent des niveaux moindres de d'ARN messager mature de l'HMGCR mais des concentrations significativement plus élevées de l'isoforme Δ13-*HMGCR*, qui sont accompagnés de niveaux protéiques élevés.

Les résultats sur les variantes rs72633963 (rapportés ici) et rs3846662 (rapportés dans la littérature) suggèrent un effet protecteur contre la MA et associées à un cholestérol total sanguin réduit nous ont mené à formuler l'hypothèse selon laquelle les effets bénéfiques de ces variantes génétiques dans la MA sont médiés via leur effet physiologique sur les niveaux de cholestérol circulant. En dépit des résultats mitigés de la première étude, nous avons décidé d'examiner cette hypothèse dans un contexte plus large. Dans la troisième et dernière étude, nous avons établi un score polygénique capturant l'effet

de multiples variantes génétiques associées aux niveaux de cholestérol total sanguin. Nous avons découvert que ce score est fortement associé aux niveaux totaux de cholestérol circulant (expliquant ~18% de la variance) et permet d'améliorer la prédiction d'hypercholestérolémie. Ces relations sont fortement influencées par l'utilisation de statines et le sexe, de telle sorte que le meilleur effet a été observé chez les femmes non-traitées aux statines. Le score a également été utilisé pour évaluer une association formelle avec la MA. Malheureusement, aucun effet significatif sur le niveau de risque, les biomarqueurs ou la sévérité de la pathologie Alzheimer n'a été détecté dans cette étude.

En conclusion, cette thèse de doctorat fournit un aperçu sur le rôle de la génétique reliée au cholestérol dans la MA. Alors que la littérature et nos résultats concernant les variantes génétiques rs72633963 et rs3846662 sont compatibles avec un rôle direct dans le métabolisme du cholestérol dans la MA, le mécanisme par lequel elles agissent reste à déterminer. Nous avons émis l'hypothèse d'un mécanisme qui agirait via les niveaux sanguins de cholestérol, mais nos résultats négatifs en ce qui concerne le score polygénique suggère l'inverse. Ces résultats mettent de plus en évidence la complexité de la relation qui existe entre la génétique liée au cholestérol et la MA et plaident pour de plus amples recherches afin de mieux cerner les mécanismes derrière les variantes génétiques qui sont reliées au cholestérol circulant.

Acknowledgements

Most people would agree that a thesis is not a one-person endeavor, and this thesis is no exception. There are many people that has contributed in various ways to the completion of this work of which I would like to acknowledge and thank.

First and foremost, I would like to thank my supervisor Dr. Judes Poirier. By accepting me into his lab, he gave me the opportunity to start this journey towards a PhD. Throughout the years he has provided invaluable guidance on this thesis, as well as on other areas of the research world while giving me the freedom to explore my own research interests. His encouragement and support have pushed me and helped shape me as a researcher. For this contribution, I am incredibly grateful; without him this thesis would not have been possible.

As part of my advisory committee, Dr. Edith Hamel and Dr. Carl Ernst, helped shape and push this thesis forward by providing invaluable feedback and support along the years, of which I am grateful.

I further want to acknowledge the generous financial support from my supervisor Dr. Judes Poirier and from Alzheimer Society of Canada and the Canadian Consortium for Neurodegeneration in Aging that allowed me to focus whole-heartedly on the research.

I would also like to extend a thank you to PREVENT-AD participants and SToP-AD nurses whose contribution made the stem cell project possible.

I would also like to acknowledge my colleagues in the Poirier lab. Dr. Cynthia Picard introduced me to genetics and with her enthusiasm sparked an interest and pushing me to dive deeper into this area of research. She has further been a great source of knowledge, expertise, and support for both the research and in the writing of the thesis itself. Anne Labonté has contributed tremendously to this thesis with her impeccable work expanding and collecting iPSCs as well as processing and analyzing samples with ELISA. Louise Théroux introduced me to LC-MS/MS and later, very generously, came

back from retirement to run my samples for this thesis. Dr. Justin Miron as the only other PhD student in the lab for a long time; I am glad for his help in navigating through the PhD and the lab. I would also like to acknowledge Doris Dea and Josée Frappier who showed me the ropes of the lab when I first started, and Dr. Sandra Pelleieux, Marina Tedeschi-Dauar, and Isabel Sierra-Flores for feedback and support. I am glad to have worked with, and alongside all of you and appreciate all the help I have received throughout the years in the lab.

As I started the stem cell project, I was offered to learn and work in the lab of Dr. Carl Ernst and I am grateful to many members of this team. Dr. Huashan Peng taught me everything I know about stem cells and has been invaluable for the stem cell project in this thesis. I got a lot of help from Hanrong Wu who always took the time to discuss and troubleshoot my project. I further like to mention fellow students Malvin Jefri, Lilit Antonyan, and Nuwan Hettige; their presence in the lab really made me feel like part of their lab and both their professional and personal support have contributed to making this thesis a reality.

I further would probably not have made it through this endeavor without the support of friends and family. In particular I would like to thank Yvonne Yau and Zahraa Chorghay; we were brought together in our first year and I am so glad for the friendship that followed. Their support helped me navigate through PhD life, and most importantly they reminded me to have fun between all the work.

Finally, I would like to thank my family. Even though they have been far, I have always felt their love and support. Above all, Mum, Nadia, Nina, and Nicole; you are the best one can hope for and I am grateful for your love, constant support, and continued belief in me when my own faltered.

Thank you

Contribution to original knowledge

With this thesis, we are the first to perform a whole-gene analysis of *HMGCR* in relationship to AD while investigating epistatic effects with *APOE*-ε4 status. We identified two hitherto uncharacterized variants that associated with AD specifically in *APOE*-ε4 carriers and further characterized the effects of one of them – rs72633963 – in three independent cohorts with regards to phenotypes relating to cholesterol metabolism and AD. This is thus the first report of rs72633963 and its involvement in both AD and cholesterol metabolism and all the findings of this study (Chapter 2) is completely novel to the field.

Literature has reported the effect of multiple other *HMGCR* variants, one of them being rs3846662. The rs3846662 AA genotype has been shown to associate with reduced risk of AD, specifically in *APOE*-ε4 carriers, hypothesized to be due to its effect of promoting alternative splicing of the *HMGCR* transcript (Δ13-HMCGR). We established a library of induced pluripotent stem cell (iPSC) lines from renal epithelial cells (RECs) collected from either rs3846662 AA or GG carriers, which was subsequently used to differentiate into neural progenitor cells (NPCs) and neurons (Chapter 3). While rs3846662 have been studied in human cell lines before, these have either been commercially available cell lines derived from tumors or immortalized lymphocytes. This study is thus the first to report the effects of rs3846662 using unaltered somatic cells (RECs) as well as the iPSCs, and NPCs derived from them. Similarly, this is the first report of HMGCR metabolism in these cell types.

In our final study (Chapter 4), we investigated a total cholesterol polygenic score in AD. The effect of lipid polygenic scores in AD have been investigated before, but in this thesis we are the first to create an optimized score of variants associating with blood total cholesterol levels by investigating different p-value cut-offs and the effects of stratification on statin use and sex. We further evaluated this score in relation to AD. For this we used three different cohorts, and could thus assess the effect in what can be considered pre-symptomatic/pre-clinical, clinical, and pathophysiologically proved AD. This

is thus the first characterization of this score across all stages of the disease, as well as different instances of disease. We assessed AD biomarkers and post-mortem pathology, as well as AD as a biological construct.

Taken together, these findings contribute to the current knowledge of the involvement of cholesterol related genetics in AD. We have identified a new variant, rs72633963, associating with cholesterol, Aβ pathology and cognition, supporting findings on that factors contributing to reduced cholesterol levels can be beneficial in AD. Although we could not find any effects of rs3846662 on AD phenotypes in cell culture, we strengthen the hypothesis of rs3846662's effect on *HMGCR* alternative splicing by a novel approach using stem cell technology to obtain human cell lines that more closely resemble the cell types present in the human body. Finally, our findings on the total cholesterol polygenic score, confirm findings from a previous study showing no effect between a total cholesterol polygenic score and clinical AD, while also extending the analysis to include pre-clinical/pre-symptomatic and pathophysiologically proved AD, as well as AD as a biological construct.

Throughout this PhD I have also contributed to the following publications outside of the thesis:

Poirier J, **Nilsson N**, Tedeschi Dauar M, Miron J, Picard C. Chapter 6 - Pharmacogenomic approaches to the treatment of sporadic Alzheimer's disease. In: Rosenberg RN, Pascual JM, eds. *Rosenberg's Molecular and Genetic Basis of Neurological and Psychiatric Disease (Sixth Edition)*. Academic Press; 2020:81-94. doi:10.1016/B978-0-12-813955-4.00006-4

My contribution: Literature review, writing and editing of text

Miron J, Picard C, **Nilsson N**, et al. CDK5RAP2 gene and tau pathophysiology in late-onset sporadic Alzheimer's disease. *Alzheimers Dement.* 2018;14(6):787-796. doi:10.1016/j.jalz.2017.12.004

My contribution: Advised on data analysis and experimental design

Poirier J, **Nilsson N**, Labonté A, Auld D, Rosa-Neto P, the Alzheimer's Disease Neuroimaging Initiative, Breitner, JCB, Villeneuve, S, Picard C, for the PREVENT-AD research group. CSF apolipoprotein B is associated with early Tau pathology in cognitively unaffected subjects with a parental history of Alzheimer's disease. *In preparation.*

My contribution: Literature review, advised on experimental design, text editing

Contribution of Authors

Overall thesis

Nathalie Nilsson: Preparation of overall thesis, proofreading and editing

Cynthia Picard: Proofreading and editing

Judes Poirier: Proofreading and editing, providing the financial resources

Chapter 1

Nathalie Nilsson: Literature review and writing of text

Cynthia Picard: Editorial comments, Figure 1.7

Judes Poirier: Editorial comments

We obtained the rights to reprint previously published figures in this thesis but not for commercial purposes (credit attributed for each individual figure in the chapter).

Chapter 2

Nathalie Nilsson: Study design, data analysis, writing of text

Cynthia Picard: Quality control of genetic data

Anne Labonté: CSF ELISA assays (for PREVENT-AD)

Valerie Leduc: Frontal cortex *HMGCR* RNA measurements

Judes Poirier: Study design, data collection

PREVENT-AD Research group: Data collection

Alzheimer's Disease Neuroimaging Initiative: Data collection

Chapter 3

Nathalie Nilsson: Study design, data collection (development of cell lines, growth and maintenance of RECs, NPCs and neurons for analyses, qRT-PCR assays, optimization of genetic integrity assay) and analysis, writing of text

Anne Labonté: Growth and maintenance of iPSCs for analyses, ELISA assays

Louise Théroux: LC-MS/MS assay

Cynthia Picard: Genetic integrity analysis

Huashan Peng: Advice on cell procedures

Hanrong Wu: Advice on cell procedures

Carl Ernst: Study design

Judes Poirier: Study design

PREVENT-AD Research group: Data collection (genetic data)

Chapter 4

Nathalie Nilsson: Study design, preparation of genetic data, creation of the polygenic score, data analysis

Cynthia Picard: Imputation of genetic data

Anne Labonté: CSF ELISA assays (for PREVENT-AD)

Pierre-François Meyer: Statistical mentoring, advice on study design

Theresa Köbe: Advice on study design

Sylvia Villeneuve: Advice on study design, supplied PET data

Judes Poirier: Study supervision

PREVENT-AD Research group: Data collection

Alzheimer's Disease Neuroimaging Initiative: Data collection

Religious Orders Study: Data collection

Memory and Aging Project: Data collection

Abbreviations

[*Cursive gene name*]	DNA or RNA
[Print gene name]	Protein
ABC transporters	ATP binding cassette transporters
ACAT1	acetyl-CoA acetyltransferase 1
AD	sporadic Alzheimer's disease
ADNI	Alzheimer's Disease Neuroimaging Initiative
APOE	apolipoprotein E
APP	β-amyloid precursor protein
AUC	area under the curve
Aβ	amyloid-β
BMI	body mass index (weight in kg / height in m^2)
CDR	clinical dementia rating scale
CERAD	Consortium to Establish a Registry for Alzheimer's Disease
CSF	cerebrospinal fluid
CYP46A1	cholesterol 24-hydroxylase
ELISA	enzyme-linked immunosorbent assays
ext-QFP	extended Quebec founder population
FL-*HMGCR*	full length *HMGCR*
GO	gene ontology
GWA/GWAS	genome-wide association study
HDL	high-density lipoprotein
HDL-C	HDL cholesterol
HMGCR	3-hydroxy-3-methylglutaryl-CoA reductase
HNRNPA1	heterogeneous nuclear ribonucleoprotein A1
iPSC	induced pluripotent stem cell
LD	linkage disequilibrium
LDL	low-density lipoprotein
LDL-C	LDL cholesterol
LDLR	low-density lipoprotein receptor
LRP1	low-density lipoprotein receptor related protein 1
MAP2	microtubule associated protein 2
MCI	mild cognitive impairment
MMSE	mini-mental state exam
MoCA	Montreal cognitive assessment
MRI	magnetic resonance imaging

MVAL	mevalonolactone
NFTs	neurofibrillary tangles
NIA-AA	National Institute on Aging and Alzheimer's Association
NINCDS-ADRDA	National Institute of Neurological and Communicative Disorders and Stroke and the Alzheimer's Disease and Related Disorders Association
NPC	neural progenitor cell
OCT4	octamer-binding transcription factor 4
PAX6	paired box 6
PC	principle component
PCA	principle component analysis
PET	positron emission tomography
PiB	Pittsburgh compound B
PREVENT-AD	PRe-symptomatic EValuation of Experimental or Novel Treatments for Alzheimer's Disease
PSEN1	presenilin 1
PSEN2	presenilin 2
p-TAU	phosphorylated TAU
RBANS	repeatable battery for the assessment of neuropsychological status
REC	renal epithelial cell
RNA	ribonucleic acid
ROSMAP	Religious Orders Study and Rush Memory and Aging Project
SE	standard error of the mean
SNP	single nucleotide polymorphism
SORL1	sortilin related receptor 1
SOX1	SRY-box transcription factor 1
SREBF	sterol regulatory element binding factor
SSEA4	stage-specific embryonic antigen 4
SUVR	standardized uptake value ratio
SYT1	synaptotagmin
TC	blood total cholesterol
TC-PGS	total cholesterol polygenic score
TUB	β-III tubulin
UTR	untranslated region
Δ13-*HMGCR*	*HMGCR* transcript lacking exon 13

List of figures and tables

Chapter 1

Chapter 2

Chapter 3

Chapter 4

Chapter 5

Contents

Introduction

AD is a multifactorial disease influenced by a range of both non-modifiable and modifiable factors [1]. One of the most important factors is genetics, with heritability estimates of 58% to 79% [2]. A large portion of the genetic contribution to AD, however, remains unknown [3] and polygenic risk score studies indicate that low-effect loci not passing the stringent genome-wide significance level ($p > 5*10^{-8}$) still influence AD risk [4–6]. One way to identify these variants would be to increase sample sizes, which indeed has proven to be successful [7,8]. However, even with sample sizes in the hundreds of thousands the newly identified loci are unlikely to make up for the missing heritability; e.g. in the latest genome-wide association study (GWAS) meta-analysis a total of 25 gene loci (corresponding to 1000s of single nucleotide polymorphisms (SNPs)) were identified [7] which is in stark contrast with results from polygenic studies suggesting that scores composed of 100 000s of SNPs has the best prediction accuracy [6]. Further, most large-scale studies investigate main effects of SNPs, but there is evidence that epistatic effects between SNPs can explain some of the missing heritability [9]. Most notably in AD research, many loci have been shown to interact with the *APOE* genotype [10–15]. Another way of discovering these SNPs could be through hypothesis driven analyses, in which individual SNPs or a smaller number of SNPs, are selected based on previous acquired knowledge of either the SNP or the loci in which it is residing, and evaluated for risk with AD.

The mevalonate pathway, with 3-hydroxy-3-methylglutaryl-CoA reductase (HMGCR) at its core, resulting in the production of cholesterol and other non-sterol isoprenoids, have been implicated in AD; midlife hypercholesterolemia associates with increased risk in later life [16–21] and statin use has been associated with reduced risk of AD [22,23], a high cholesterol diet in animal models leads to increased Aβ load in the brain [24–29] and cellular studies have shown beneficial effects of statin treatment on Aβ pathology [30,31]. Similarly, a promoter variant (rs3761740) and an intronic variant

(rs3846662) within the *HMGCR* gene has been shown to associate with AD risk [13,32–36] and it has been suggested that these effects are mediated through modulation of HMGCR levels or activity [32,33,35,37,38].

Although AD was first described over a hundred years ago, there is no cure or preventative treatment [1]. The class of drugs that has been investigated the most are the ones targeting Aβ but they have shown to have little clinical benefit, and in some cases even detrimental effects, in randomized controlled trials [39,40]. Even though in later years the research on other non-Aβ treatments have increased, it is still clear that mechanisms contributing to AD need to be determined to generate or solidify hypothesis that could lead to sorely needed treatments.

Thus, based on the knowledge we have on the involvement of the mevalonate pathway in AD, we reasoned that some of the unknown genetic variants in AD could be identified by hypothesis driven analyses of cholesterol related variants. Identification of such variants would lend support to the hypothesis that inhibition of HMGCR, either genetically or pharmacologically, is beneficial in AD, which in turn could lead to a strategy for disease modification or prevention. Consequently, the overall objective of this thesis was to investigate whether SNPs related to cholesterol metabolism influence AD risk and/or contributes to any of the core AD pathologies. We used three different approaches, resulting in the three chapters presented here.

Objectives study 1

The main objective was to determine the involvement of SNPs in the *HMGCR* gene in AD. Specifically, we aimed to identify SNPs by examining epistatic effects with *APOE*-ε4. In an attempt to understand how any identified SNP would contribute to AD, we further aimed to characterize the SNPs relation to cholesterol metabolism (HMGCR expression levels, blood lipid profiles) and AD core pathologies (Aβ, TAU and neurodegeneration). The results are presented in Chapter 2.

Objectives study 2

The main objective was to characterize the previously identified *HMGCR* SNP, rs3846662, *in vitro* using multiple cell types derived from the same set of donors. Specifically, we aimed to develop stem cell lines for carriers or non-carriers of the protective genotype (AA) which could then be differentiated into different cell types. Across cell types, we then aimed to determine differences between carriers and non-carriers in HMGCR metabolism and AD biomarkers; e.g. levels of *HMGCR* RNA (full length and spliced variants), protein levels and activity as well as levels of intracellular TAU and extracellular Aβ-42 and Aβ-40. Results are presented in Chapter 3.

Objectives study 3

Under the hypothesis that hypercholesterolemia associates with increased AD risk and to capture the effects of multiple cholesterol related SNPs, we aimed to create a polygenic score based on SNPs associating with blood TC levels (TC-PGS) and determine its correlation with AD. Specifically, we aimed to first evaluate the influence of the TC-PGS on cholesterol levels and hypercholesterolemia, and to determine effects of sex and statin use on these relationships. Subsequently, the objectives were to determine if increased genetic load of cholesterol related alleles conferred higher risk of developing AD or any of its core pathologies. Results are presented in Chapter 4.

Chapter 1

Alzheimer's disease, genetics & cholesterol – a literature review

Alzheimer's disease (AD) is a severe neurodegenerative disease resulting in dementia and ultimately death [1]. It is the most common form of dementia [1] and, in the more typical cases, present with memory loss and executive dysfunctions [41,42]. AD and other dementias is estimated to affect about 50 million people worldwide [43] and numbers are expected to rise [44]. Based on its etiology, AD can further be divided into familial and sporadic forms. Familial AD is caused by genetic mutations and is inherited in an autosomal dominant fashion whereas the sporadic form is considered a multifactorial disease [45]. This latter form is by far the most common and is estimated to account for 99% of AD cases and typically have a later onset (> 65 years) [1]. Regardless of form, there is no cure for AD with current treatments being purely symptomatic [42]. Throughout this thesis, and from this point on, "AD" will refer to the sporadic form unless otherwise stated.

1.1 Alzheimer's disease

1.1.1 Diagnosis

AD is clinically diagnosed in a two-step process: first, the presence of dementia is established and second, whether this dementia is due to AD is determined [41]. In order to establish dementia, the presence of cognitive or behavioral symptoms is evaluated by cognitive assessments (e.g. the mini-mental state exam (MMSE) [46], the Montreal cognitive assessment (MoCA) [47], and the clinical dementia rating scale (CDR) [48]) and neuropsychological testing in combination with the individual's

history. Diagnosis is conjunctly determined by the severity of the symptoms (the ability to function at work or at usual activities is affected) and by a clear deterioration over time. A minimum of two of the following cognitive domains need to be affected: (A) ability to acquire and remember new information, (B) reasoning and handling of complex tasks, (C) visuospatial abilities, (D) language functions, and (E) personality and behavior.

A clinical AD diagnosis are made based on the presentation and can either be deemed probable or possible AD [41]. For probable AD, symptoms must have a gradual onset over months to years and there must be an obvious worsening of cognition. As mentioned above, at least two cognitive domains need to be affected. Based on the initial and most prominent cognitive deficit, AD is diagnosed as either amnestic (main domain affected is (A); impairment in learning and recall) or non-amnestic (main domain affected is either (B) reasoning and handling of complex task, (C) visuospatial abilities or (D) language functions). In addition, other causes of dementia, such as cerebrovascular disease, dementia with Lewy bodies, frontotemporal dementia, aphasia, and other neurological disease, must be excluded. Possible AD is similar to probable AD with some easing of the criteria - sudden instead of gradual symptom onset and absence of progressive decline (i.e., atypical course), or evidence of other dementing diseases can be present (i.e., etiologically mixed presentation). First after post-mortem examination of the brain, a diagnosis of definite (or pathophysiologically proven) AD can be made based on the presence of the hallmark AD pathologies: amyloid-β (Aβ) plaques and neurofibrillary tangles (NFTs) [41,49] (see 1.1.5 Pathology).

1.1.2 Definition

More than a hundred years have passed since Alois Alzheimer described the first case of AD [50] and the definition of AD has since evolved significantly with new discoveries. Alzheimer first described AD as a *pre-senile* dementia that upon autopsy, revealed the presence of then novel pathologies in the brain, now known as amyloid-β plaques and neurofibrillary tangles [50–53]. It was not until much later

that the clinical phenotype was expanded to include dementia cases of older age as it was shown that senile dementia cases also displayed high levels of Aβ plaques at autopsy [54].

In 1984, criteria for possible, probable, and definite AD were published by the National Institute of Neurological and Communicative Disorders and Stroke and the Alzheimer's Disease and Related Disorders Association (NINCDS-ADRDA) [55]. Possible and probable diagnoses were purely clinical assessments, whereas definite AD included histopathological evidence of AD [55]. The accuracy of the clinical diagnoses in predicting definite AD was evaluated and it was shown that a clinical diagnosis of probable AD had a sensitivity of 81% and a specificity of 70% [41]. This led to updated guidelines in 2011 where biomarkers were incorporated - for research purposes only - to evaluate the presence or absence of AD neuropathological change in clinical cases. Moreover, concepts like pre-clinical AD and mild cognitive impairment (MCI) were included [41,56,57]. This allowed refinement of the clinical diagnoses - at least in a research setting.

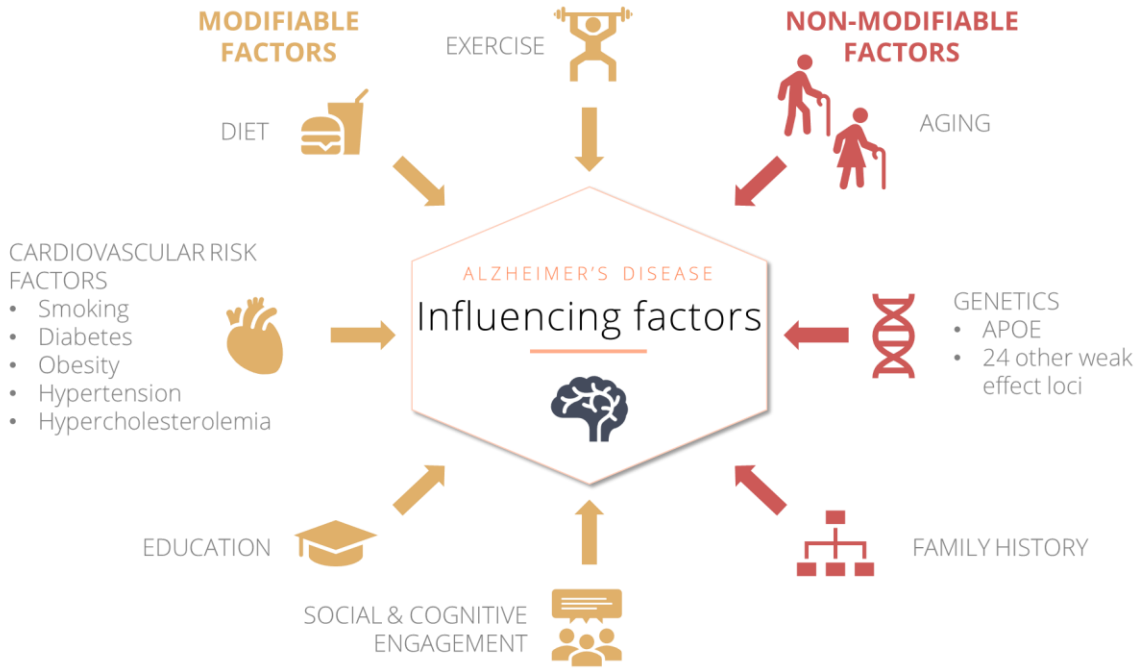
In recent years, a research framework by the National Institute on Aging and Alzheimer's Association (NIA-AA) was laid out, proposing a shift from viewing AD as a clinical syndrome to defining AD as a biological construct [58]. The objective was to create a scheme for defining and staging AD across its entire spectrum (including prodromal and preclinical phases); focusing on the diagnosis of AD *in vivo* using AD biomarkers as proxies for neuropathological change. They further proposed the AT(N) classification system as a scheme to classify individuals according to their pathology [59]. In this scheme, proxies of the hallmark pathologies Aβ (A), TAU (T) as well as neurodegeneration (N) are assessed (see 1.1.6 Biomarkers) either as continuous variables or dichotomized (e.g. Aβ/TAU positivity). These pathological categorizations could then be used in combination with clinical diagnoses of AD. The AT(N) scheme was further devised to be able to incorporate other biomarkers as research progresses; e.g. inflammation and vascular changes have been suggested to be early factors and could be integrated as AT(N)-I or AT(N)-V [58–60].

1.1.3 Etiology

While familial AD is caused by autosomal dominant mutations in β-amyloid precursor protein (*APP*), presenilin 1 (*PSEN1*), and presenilin 2 (*PSEN2*) [61], the sporadic form is considered a multifactorial disease caused by a combination of risk factors that can be divided into modifiable or non-modifiable (**Figure 1.1**) [42,61]. It should be noted that patterns of autosomal dominant inheritance (albeit with lower penetrance) of rare variants in genes such as *TREM2* and *ECE2* has been observed in cases with later onset, but it is yet unclear whether these are considered familial mutations or risk factors for sporadic AD [62,63].

Non-modifiable factors: Age, genetics, and family history all influence the risk of developing AD. As such, they contribute in assessing risk in individuals but can also aid in identifying mechanisms (e.g. by identifying genes associating with AD). Age is considered the most important factor with

Figure 1.1 Alzheimer's disease risk factors



AD is influenced by a number of factors that can be divided into modifiable (orange color) and non-modifiable (pink color).

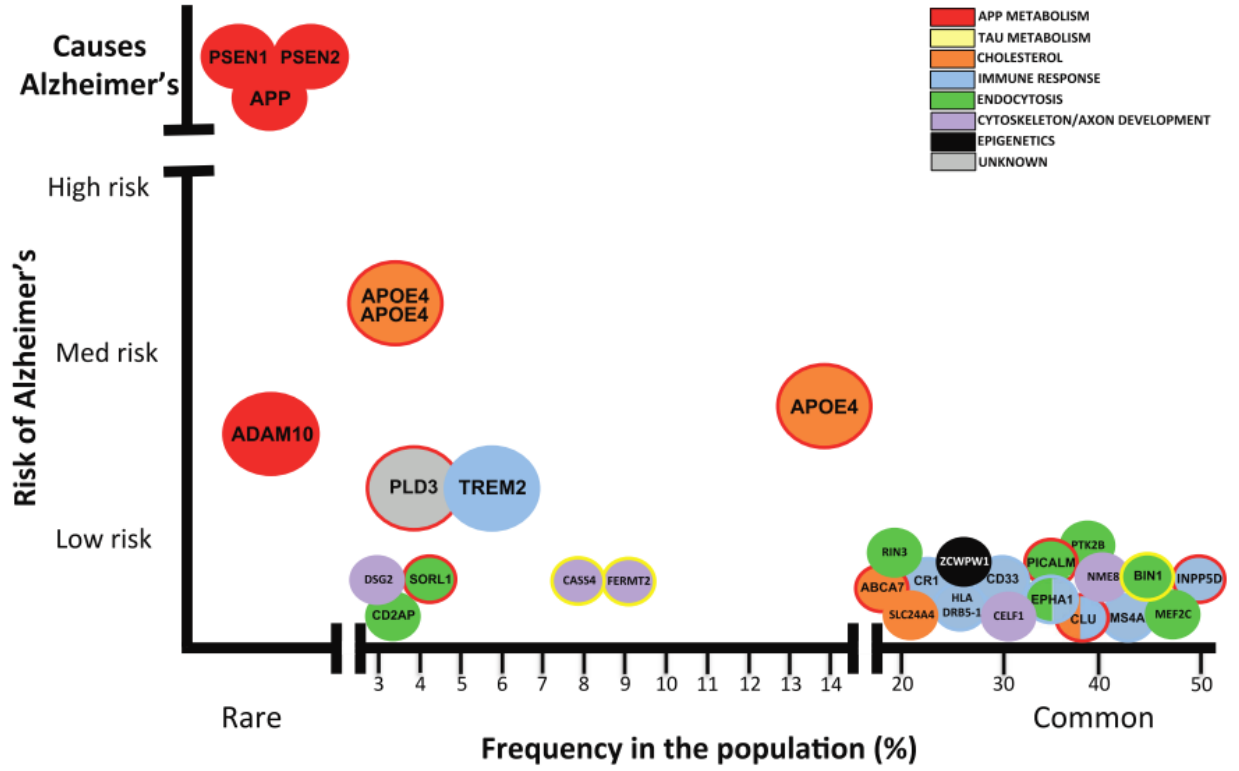
incidence rates of AD increasing from 0.13% per year at age 65-69, to 6.48% per year at age >85 years [64]. In a similar vein, the prevalence rate increases from 3% at age 65-74, to 32% at age 85 years or older [1,2]. Though it has been suggested that AD reflects an accelerated aging process, evidence indicates that it is distinct from normative aging [65].

The composite risk of genetics and environment is exemplified by the amplified risk in individuals that have first-degree relatives with AD; the lifetime risk of such individuals are 39% by the age of 96 years [66]. Risk analysis in children of AD affected parents revealed that children with two AD affected parents had a 5x higher risk of AD compared to children with no affected parents, and a 1.5x higher risk compared to children with only one affected parent [21] – highlighting a potential compounding effect.

Heritability estimates of AD range from 58% to 79% and further underscore the importance of genetics in AD [2]. The most prominent genetic loci is the one encoding apolipoprotein E (*APOE*) [1]. Two SNPs results in three different alleles; ε2, ε3 and ε4. Carries of the ε4 allele have consistently been shown to have a 3- to 12-fold increased risk compared to carriers of two ε3 alleles [67–69]. While these risk assessments are largely derived from a Caucasian population, the increased risk associated with the ε4 allele is true for other ethnicities as well, but to varying degree [70]. Genome-wide association (GWA) studies and meta-analyses thereof, as well as hypothesis-driven, gene-targeted studies, have identified many other loci (e.g. see [7,71–73]) but they either have comparably small effect size or are rare occurring variants (**Figure 1.2**) [45]. The role of genetics is discussed in detail below (see 1.2 Genetics in sporadic Alzheimer's disease).

Modifiable factors: A myriad of modifiable factors have been identified for AD and are of great interest due to the possibility for intervention. For example, education has been shown to relate negatively with risk [74–76] and can be partially, though not fully, accounted for by other factors such as cognitive reserve, socioeconomic status, and study design [1,75]. Similarly, engaging in mental, social, or productive activities have been shown to associate with reduced risk of AD [77].

Figure 1.2 Genetic risk factors in Alzheimer's disease (familial & sporadic)



The relationship between genetic variants, population frequency, and risk of AD. Figure reproduced with permission from [61,79].

Cardiovascular factors constitute another group of important risk factor. A recent meta-analysis revealed that smoking, low diastolic blood pressure, high body mass index (BMI) in midlife, and type 2 diabetes all associated with increased risk of AD [76]. Other cardiovascular factors such as obesity and hypercholesterolemia have also been associated with increased risk of AD [16,17,78]. This group of factors are attractive targets in AD research due to the availability of already approved drugs to treat these traits. Indeed, statin use and use of hypertensive medications have been linked to reduced risk of AD [76] although the clinical benefit in AD is debated [80,81]. The involvement of hypercholesterolemia is of particular interest for this thesis and is discussed in detail below (see 1.3 Cholesterol in Alzheimer's disease).

1.1.4 Epidemiology

Dementia and AD are common diseases worldwide with the World Health Organization estimating that 50 million people suffer from dementia and that there are nearly 10 million new cases each year [43]. The Public Health Agency Canada estimates that more than 419,000 Canadians over the age of 65 are affected by dementia and that there are 78,600 new cases each year [82]. In addition to its prevalence, AD cases are expected to rise as exemplified by a study of the US; from 4.7 million in 2010 to 13.8 million in 2050 [44]. Further, comparing with other diseases (such as heart disease, cancer, stroke, and HIV) where the number of deaths have decreased due to tremendous progress in finding treatments and preventions, the percentage of people dying from AD has increased by 145%, emphasizing the need for further research [1].

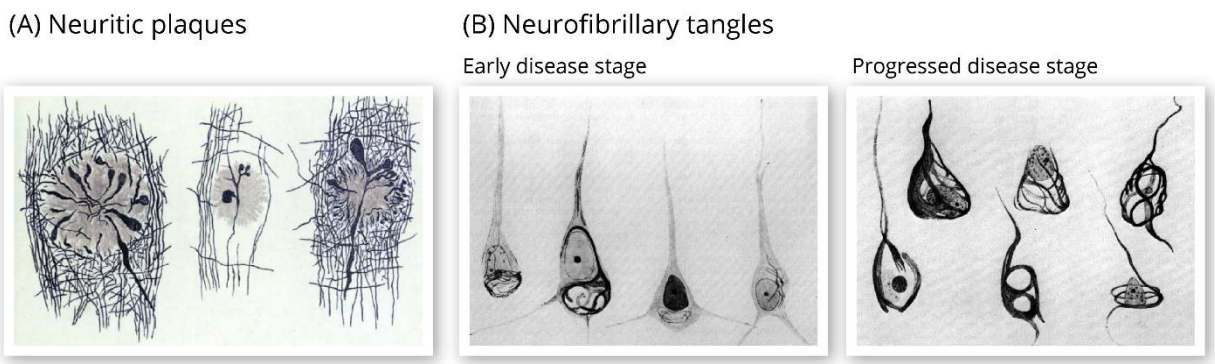
1.1.5 Pathology

As mentioned above, AD is a clinicopathological construct [41]. The hallmark pathologies include Aβ deposition, TAU accumulation, and neurodegeneration, but there are many other reported [50,83].

Amyloid depositions: Aβ is a product of sequential cleavage of APP [84] – a single-pass, transmembrane protein that can be processed in a non-amyloidogenic or amyloidogenic pathway [85]. The extracellular N-terminal domain of APP is first cleaved by either α- or β-secretase followed by transmembrane cleavage by the γ-secretase. Sequential cleavage of APP by α- and γ-secretase constitutes the non-amyloidogenic pathway and results in the release of a soluble APPsα peptide and a p3 peptide, whereas cleavage by β- and γ-secretase constitutes the amyloidogenic pathway and results in the release of a soluble APPsβ peptide and an Aβ peptide [84]. Furthermore, the γ-secretase can cleave at various sites resulting in Aβ fragments of different lengths, with the Aβ 1-40 (Aβ-40) and 1-42 (Aβ-42) being the most common ones [85–87]. These peptides can act as monomers or oligomers both of which has been shown to be detrimental to neurons, or form extracellular fibrillar structures, ultimately giving rise to deposits (a.k.a. plaques) [83,85].

There are multiple types of Aβ deposits or plaques; diffuse, stellate and focal [83]. For pathological purposes, focus has been put on diffuse and focal plaques, particularly on a subset of focal plaques referred to as senile or neuritic plaques [83,88,89]. These plaques have a core of Aβ peptides that co-localizes with microglia, and are surrounded by a corona of neuritic (TAU pathology, see below) and astrocytic components (**Figure 1.3 A**, as first depicted by Oskar Fischer in 1907 [51,52]) [83].

Figure 1.3 First depictions of neuritic plaques and neurofibrillary tangles



Early drawings of neuritic plaques by Oskar Fischer (A) and neurofibrillary tangles by Alois Alzheimer (B). Figures reproduced with permission from [52] and [90], respectively.

Numerous efforts to stage Aβ pathology have been done and are now widely used for research and diagnostic purposes. For example, Thal et al. [89] classified the progression of Aβ pathology into five phases depending on the spread of plaques. Early phases are characterized by mainly neocortical spread that then extends to the entorhinal cortex and limbic regions, and finally, spreading to brain stem nuclei and the cerebellum (**Figure 1.4 A**). The Consortium to Establish a Registry for Alzheimer's Disease (CERAD) created a standardized neuropathological assessment based on the presence of neuritic plaques, clinical diagnosis and age, resulting in classification of individuals into four groups: normal, CERAD possible AD, CERAD probable AD, and definite AD [88]. CERAD classification is part of the current clinical guidelines for determining definite AD [41,49].

TAU accumulation: TAU is an intracellular protein that binds to the cell cytoskeleton, specifically the microtubules, to aid in assembly and stabilization [91,92]. In doing so, TAU can indirectly affect cell

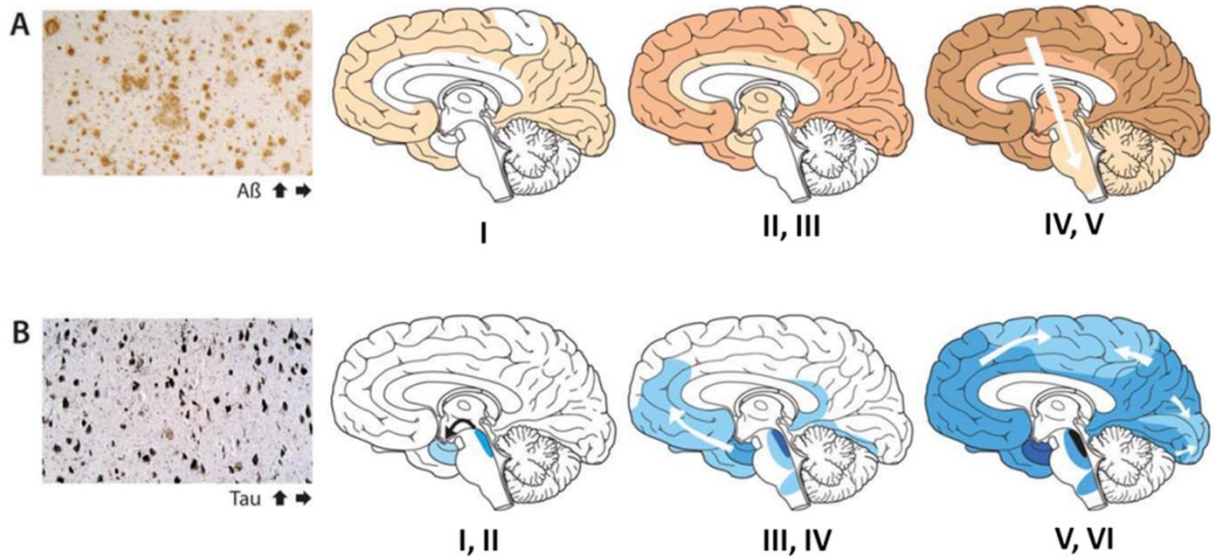
morphology, plasticity and axonal transport [92,93]. TAU harbors a large number of phosphorylation sites that when phosphorylated impairs its ability to bind microtubules [91,93]. Progressive phosphorylation of TAU at multiple sites (i.e. hyperphosphorylated TAU) makes it prone to oligomerization that can further aggregate; a phenomenon occurring in AD and other tauopathies [91]. In AD, intracellular aggregates of hyperphosphorylated TAU occurs in the cell soma giving rise to NFTs, in the dendrites resulting in neuropil threads, or in the axons resulting in the neuritic processes that are part of the neuritic plaques [83]. In later stages of the disease, when tangle bearing neurons die, NFTs becomes an extracellular structure (so called "ghost tangle") that can be degraded by astrocytes (**Figure 1.3 B**, as depicted by Alois Alzheimer) [94].

The Braak staging is widely used to assess the progression and stage of TAU accumulation [94]. The staging suggests that NFTs and neuropil threads first occur in the entorhinal region with modest involvement of CA1 and subiculum (stage I-II) that then progress to limbic regions such as the amygdala, nucleus accumbens and putamen (stage III-IV) to finally reach the neocortex and the extrapyramidal system (stage V-VI; **Figure 1.4 B**). Similar to CERAD, Braak staging is, part of the current clinical guidelines for determining definite AD [41,49].

Neurodegeneration: Loss of neurons and synapses in AD has been widely described and is hypothesized to occur mainly as a result of TAU pathology. In line with the aforementioned TAU pathology, neuronal loss in AD has been observed for the entorhinal cortex, CA1, superior temporal gyrus, supramarginal gyrus, olfactory bulb, amygdala, nucleus basalis of Meynert, substantia nigra, locus coeruleus, and raphe nuclei [83,96].

Other: Many other pathological changes have been observed in the brains of individuals with AD. Alzheimer, in his first report, noted signs of arteriosclerosis of the vascular tissues, glial changes and endothelial growths, and proliferation of vessels [50]. In fact, "pure AD" pathology (plaques and tangles only) rarely occur on its own and vascular lesions (e.g. cerebral amyloid angiopathy), Lewy bodies and TDP-43 pathologies often co-exist with the more typical plaques and tangles [83,94,97].

Figure 1.4 Pathology progression in Alzheimer's disease



Progression of Aβ (A) and TAU (B) pathology in AD according to Thal et al. [89] and Braak and Braak [94], respectively. Figure reproduced with permission from [95].

1.1.6 Biomarkers

Multiple biomarkers have been developed for AD and its pathologies, with some of them being thoroughly validated. The core pathologies of AD; Aβ, TAU and neurodegeneration have successfully been captured using cerebrospinal fluid (CSF) measurements of pathological peptides and a multitude of imaging techniques such as structural magnetic resonance imaging (MRI) and positron emission tomography (PET) [58]. These measurements have shown predictive ability for both AD status and presence of post-mortem pathologies in the brain.

Aβ pathology: In the CSF, levels of Aβ-42 and its ratio with Aβ-40 have been shown to be inversely correlated with Aβ accumulation in the brain. This is argued to be due to increased sequestration of these peptides in the brain as accumulation and plaque density increases [98–100]. The first and most common Aβ PET tracer is ^{11}C Pittsburgh compound B (PiB) that binds to fibrillar Aβ [101] and has shown to correlate well with neuropathology (both neuritic and diffuse plaques) [102]. Other Aβ PET tracers have been developed, such as the ^{18}F-Florbetapir (or ^{18}F-AV-45) that has a longer half-life

compared to PiB [103]. This tracer has been shown to correlate very well with Aβ plaque load at autopsy with a sensitivity and specificity of 92% and 100% respectively [103]. Both CSF and PET Aβ biomarkers have been used to dichotomize individuals as either having (Aβ(+)) or not having (Aβ(-)) Aβ pathology in the brain in accordance with the AT(N) system [59]. In line with this, CSF Aβ-42 and even more so the ratio of CSF p-TAU/Aβ-42 successfully predict Aβ positivity as determined by PET with area under the curve (AUC) values of 92.1% and 96.3%, respectively [104].

TAU pathology: CSF levels of phosphorylated TAU (p-TAU) and TAU strongly correlate positively with hyperphosphorylated TAU in cortical biopsies from the brain and with NFTs [99,105]. While both p-TAU and TAU correlate with NFTs, it is widely considered that p-TAU (especially TAU phosphorylated at threonines 181 or 217) is a more specific biomarker for AD, since TAU levels are increased in other neurodegenerative conditions (e.g. stroke) while p-TAU is not [105]. *In vivo* imaging of TAU accumulation is achieved with ^{18}F-Flortaucipir (or ^{18}F-AV-1451); a ligand that has high affinity for paired-helical filaments (insoluble fibers composed of aggregated hyperphosphorylated TAU) [106]. This tracer showed increased binding in Braak regions [106], and recently was shown to predict post-mortem pathological Braak stage V-VI with sensitivity ranging from 92.3% to 100% and specificity from 52.0% to 92.0% [107].

Neurodegeneration: Global neurodegeneration can be assessed by CSF TAU levels whereas more focal, regional neurodegeneration can be identified by FDG PET hypometabolism and atrophy on MRI [58]. [18F]-radiolabeled 2-fluoro-2-deoxy-d-glucose PET (FDG-PET) measures glucose metabolism and is thought to reflect synaptic loss. A meta-analysis of regions of interest affected in AD revealed reduced metabolism in the right and left angular gyri, bilateral posterior cingulate, and the right and left inferior temporal gyri [108]. MRI studies have shown that cerebral volume loss occurs with time but is more pronounced in AD and the loss occurs in a similar pattern in patients with established AD [109]. Of note, since neurodegeneration is a prominent feature of other diseases than

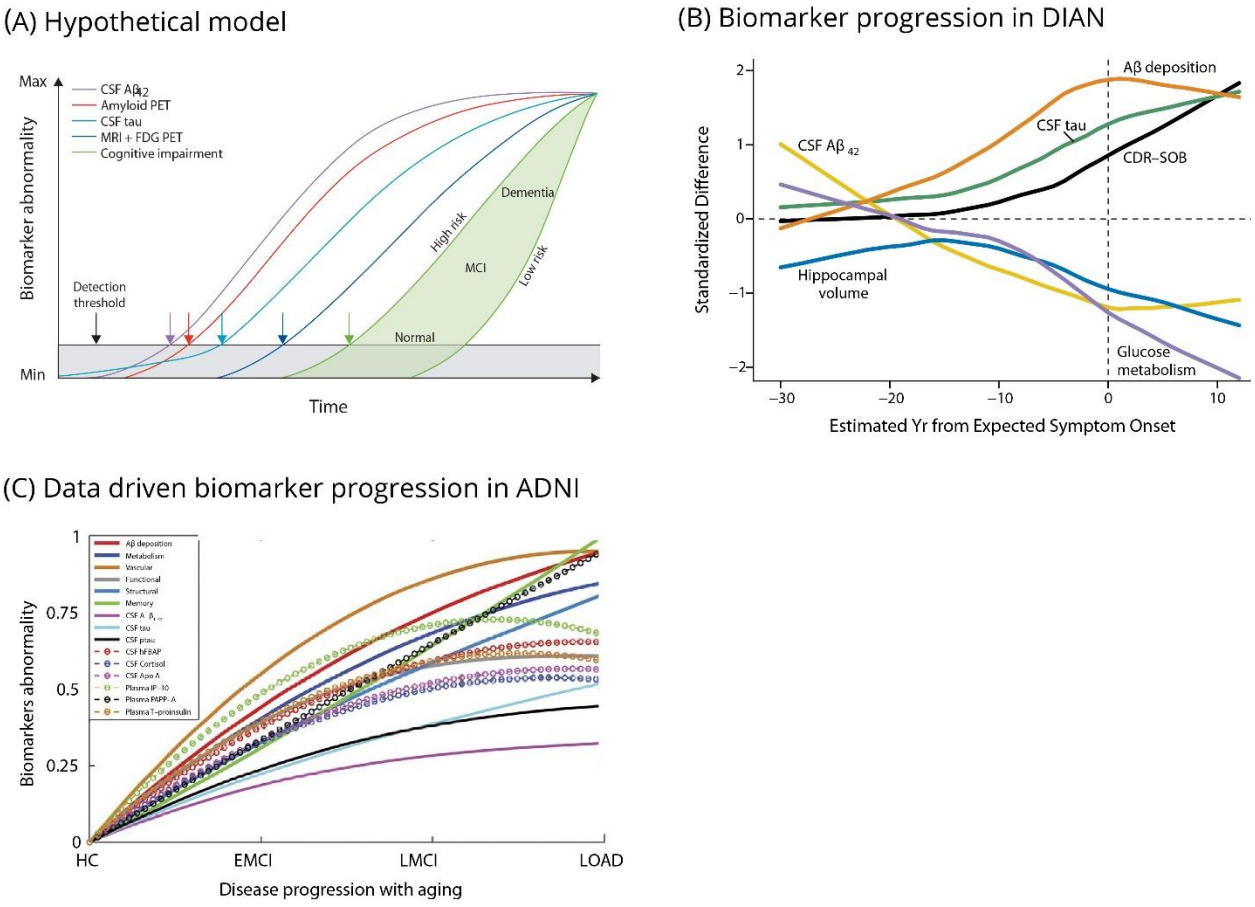
AD, alterations of the biomarkers are not specific to AD; e.g. CSF TAU is increased in AD but transient increases are also observed following a stroke or brain trauma [105].

AD status: The aforementioned biomarkers have also been evaluated for their ability to predict a diagnosis of AD. For example, CSF Aβ-42, p-TAU and TAU have demonstrated success in discriminating healthy individuals from those suffering from AD [98,110], with Aβ-42 performing best with a sensitivity of 86% and a specificity of 90% [110]. Combining multiple CSF biomarkers have been shown to yield greater sensitivity and specificity than the individual biomarkers alone. CSF TAU and Aβ-42 combined had a sensitivity of 95% and specificity of 83% to detect incipient AD in individuals with MCI, and when all three were combined (TAU and Aβ-42/p-TAU ratio) specificity was slightly increased (87%) [111]. Similarly, PET studies tracing Aβ and TAU have shown that prevalence of Aβ positivity increased with clinical diagnosis (healthy control < MCI < AD) [101,112] and that TAU PET associated with reduced cognition [106].

Temporal ordering: The sequence of the occurrence of above described biomarkers have been extensively investigated and multiple models proposed. Currently, the most discussed and accepted model was proposed by Jack et al., [113,114] wherein Aβ accumulation start early on in pre-symptomatic or prodromal stages of the disease, followed by evidence of TAU accumulation and finally neurodegeneration close to onset of disease (**Figure 1.5 A**). Using the Dominantly Inherited Alzheimer Network (DIAN) cohort, this sequence of events was confirmed in individuals with familial AD (**Figure 1.5 B**) [115].

Interestingly, a data-driven approach in late-onset cases revealed a somewhat different temporal ordering, and suggested disturbances of biomarkers other than the core Aβ, TAU and neurodegeneration (**Figure 1.5 C**) [60]. Compared to the Jack et al., model, this one suggests that Aβ accumulation is preceded by vascular dysfunction, that there are other, more sensitive proteinopathies for disease progression than CSF Aβ-42 and p-TAU (e.g. CSF hFABP, cortisol and APOA), and that biomarkers of neurodegeneration and memory consistently appear earlier. While

Figure 1.5 Temporal ordering of AD biomarkers



The temporal appearance and progression of AD biomarkers is depicted; a hypothetical model (A), actual biomarker progression in individuals with familial AD (B), and a data-driven model in sporadic AD (C). Figures reproduced with permission from [114], [115] (Copyright Massachusetts Medical Society), and [60] (Creative Commons Attribution 4.0 International Public License) for (A), (B), and (C) respectively.
Abbreviations: CDR-SOB, the Clinical Dementia Rating–Sum of Boxes; CSF, cerebrospinal fluid; FDG, fluorodeoxyglucose; EMCI, early MCI, LMCI, late MCI; LOAD, late-onset AD; MCI, mild cognitive impairment; MRI, magnetic resonance imaging; PET, positron emission tomography; Yr, year.

these models shed light on potential mechanisms and the sequence of events, it is clear that they need to be further validated.

1.2 Genetics in sporadic Alzheimer's disease

Though sporadic AD by definition is not tied to causative genetic mutations, it is nonetheless highly influenced by genetics. Heritability estimates in twin studies range from 58% to 79% [2] while the

percentage of phenotypic variance explained by genetics ranges from 24% to 53% [3,4]. The first evidence for a genetic influence on AD came from multiple concurrent studies in 1993, that implicated the *APOE* locus [67,68,116,117]. This association has since been confirmed in countless GWA studies (reviewed in [118]) and is currently recognized as the most important genetic risk factor for AD, unrivaled by any other loci [3,119]. Multiple other variants have been identified using GWA studies and meta-analyses thereof [7,8,71,72], with a recent study identifying 24 variants in addition to *APOE* [7].

1.2.1 APOE

APOE genotypes and AD risk: Two SNPs within the *APOE* gene, rs429358 and rs7412, combine to give three different alleles (ε2, ε3 and ε4), resulting in isoforms APOE2, APOE3 and APOE4 of the protein [120]. These alleles result in six possible genotypes: ε2/ε2, ε2/ε3, ε2/ε4, ε3/ε3, ε3/ε4, and ε4/ε4. While the ε2 allele associates with reduced risk of AD [121] and increased longevity [122,123], the ε4 allele associates with increased risk of AD [67,68,70,116,117]. *APOE*-ε4 is the most important genetic risk factor and is second only to age as an AD risk factor [3,119]. *APOE* associations with AD have been shown to interact with sex, ethnicity, and age [67,70,124,125]. Risk is greater for female ε4 carriers than for male ε4 carriers [124] and the effect of *APOE* genotypes decreases with age [67,70,121,126]. Furthermore, Farrer et al., [70] showed the effect of ethnicity; compared to ε3/ε3 carriers, odds ratios (ORs) for ε3/ε4 carriers ranged from 1.1 (African Americans) to 5.6 (Japanese) and the difference was even more striking for the ε4/ε4 carriers with ORs ranging from 2.2 (Hispanics) to 33.1 (Japanese).

APOE in preclinical AD: *APOE* genotype has been shown to influence AD processes in pre-clinical stages and even in young adults. Glucose metabolism in regions affected by AD are reduced in both young (20-39 years old) and middle-late life (65-80 years old) adult ε4 carriers [127,128]. ε4 carriers also have increased Aβ pathology (assessed by CSF Aβ-42 levels and PiB PET) compared to non-carriers in cognitively normal adults [129–131]. This merges with the neuropathological findings of increased neuritic plaques among ε3/ε4 [132] or ε4/ε4 carriers [133]. A few studies have also assessed the effect on TAU accumulation, but with varying results. While two studies found no association between *APOE*

genotype and CSF biomarkers of TAU pathology [129,131], two post-mortem pathological studies reported increases in neocortical levels of NFTs in ε4 carriers [132,133].

APOE mechanisms: The mechanisms by which *APOE* confers this increased risk has been researched extensively. *APOE* encodes apolipoprotein E - the main transporter of cholesterol in the brain [120]. It coordinates the redistribution of cholesterol needed for repair, growth, and maintenance in response to injury [134]. Hippocampal APOE protein levels have been shown to decrease with *APOE*-ε4 allele dose [135] and thus, it has been suggested that reduced cholesterol transport could be one of the pathways that *APOE* exerts its effect [134–136].

APOE has also been shown to co-localize with both Aβ plaques and NFTs [137] suggesting a direct effect of APOE on the core AD pathologies. *In vivo* studies using APP mouse models of AD have shown that knock-out of the *APOE* gene reduced levels of Aβ deposits in the brain [138,139]. Delivery of the different human *APOE* alleles in these mice (*APP* with *APOE* knocked-out) showed that only the ε4 allele increased Aβ deposition and Aβ-42 levels, while ε2 and ε3 alleles had no effect [140,141]. It was suggested that this is due to a reduction in Aβ clearance rather than increases in synthesis or processing [142]. Conversely, expression of the alleles together with the endogenous mouse *APOE* expression intact, revealed reduction in Aβ deposits with the ε2 allele [140]. The connection between *APOE* and TAU has been less studied, but one study suggests that neuron-specific expression of the ε4 allele, but not the ε3 allele, increases tau phosphorylation [136,143].

1.2.2 Other genetic contribution to AD

Findings from GWA studies: In an effort to identify additional SNPs and loci contributing to AD, numerous GWA studies have been performed and the field is now turning to meta-analysis of these studies to increase power [7,71,72]. More than 30 loci have been identified in this way [7,8,71,73,144,145]. Pathway analysis of implicated loci from one of the latest meta-analyses [7] revealed enrichment of genes involved in APP/Aβ metabolism, tau binding, lipid metabolism, and immune response [7]. The effect sizes of these loci are relatively small with ORs ranging from 0.80 (*SORL1*) to 2.08 (*TREM2*),

compared to the *APOE* loci where a single ε4 allele associates with ~3 times higher risk and two ε4 alleles with ~12 times higher risk (**Figure 1.2**) [7,67,70,117]. The importance of the *APOE* loci is further emphasized by a study that showed that the *APOE* genotype alone explained 25.21% of the AD genetic variance and that all other known GWAS SNPs combined, explained a mere 5.41% [3]. It is likely that with the inclusion of additional SNPs identified in later GWA studies, this percentage would increase.

Polygenic scores: The combined effect of SNPs identified by GWA studies or meta-analyses thereof has been assessed by constructing polygenic scores [146] in multiple cohorts [5,6,147–152]. Most of these studies have used the meta-analysis results from Lambert et al. [71] and excluded the *APOE* loci from the score. Polygenic scores of genome-wide significant SNPs from this study associated with increased risk of AD with an OR of 1.31 after correcting for age, sex and *APOE*-ε4 status [148]; well below the risk associated with *APOE*. However, one important aspect of polygenic scores is the selection process of SNPs to include in the score [146]. Indeed, other studies have shown that including SNPs not passing the genome-wide significance threshold improved the ability to discriminate between controls and AD with p-value thresholds of up to 0.5 best predicting AD [5,6]. Desikan et al., [147] evaluated SNPs with a p-value ≤ $1*10^{-5}$ and then used a stepwise procedure in survival analysis to reach the final set of SNPs included in the score. They showed that depending on the polygenic risk, in *APOE*-ε3/ε3 carriers, age of AD onset can vary by more than 10 years (comparing 10th decile with 1st decile). Further, the score associated with Braak stages, CERAD scores, worsening of cognition, and greater volume loss in entorhinal cortex and hippocampus [147]. Multiple studies have now assessed polygenic contribution to AD using different approaches and it is clear that these scores have an effect on AD risk prediction [5,6,148,149], cognition [149–151], hippocampal volume [5,149,152] and pathology [149].

1.2.3 Missing heritability

Although great progress has been made in identifying loci associated with AD, many remain unknown. It has been shown that known variants only account for a fraction of the total genetic variance [3,4,153]; genome wide significant SNPs were shown to explain 31% of the total genetic

variance, leaving 69% of the variance to be explained by SNPs yet to be identified [3]. This missing heritability could be due to several things.

Low effect loci: As polygenic scores reveal, while loci with weaker effect do not reach genome-wide significance in GWA studies, they nonetheless contribute to AD risk (see 1.2.2 Other genetic contribution to AD). In order to detect these loci, larger sample sizes are needed. With current advances in technology, allowing collection, processing, and analysis of ever-increasing amount of data, the most recent AD meta-analyses included hundreds of thousands individuals leading to identification of new loci [7,8,72].

Rare variants: GWA studies have mainly focused on common variants (minor allele frequency ≥ 0.05) but in the recent decade, research into rare variants have increased and a few new loci have been identified in *APP* (A673T) [154], *TREM2* (R47H) [155], *ECE2* [62], and *APOB* [156]. These, and others yet to be identified rare variants could explain some of the missing heritability.

Epistasis: There is also the suggestion that the missing heritability is due to epistasis between genetic variants [9]. Most large-scale studies have investigated main effects of SNPs, but there is evidence for epistatic effects. Multiple loci have been shown to interact with *APOE* genotype including *PICALM* [10–12], *HMGCR* [13], *BIN1* [12], *CR1* [15] and *ABCA7* [14], to mention a few. These are all interactions with APOE genotype; however, epistasis can occur between any two genotypes. One study examined multilocus patterns created from *PICALM*, *CR1*, *CLU*, *BIN1* and *APOE* genotypes and found a *PICALM-CLU* pattern to be the strongest for lower memory performance [157].

The Epistasis Project is a collaboration between seven AD research groups that first aims at replicating AD epistasis findings and secondly to reveal true risk loci [158]. Publications from this project have so far replicated interactions between inflammation markers *IL6* and *IL10* [158], iron metabolism genes *HFE* and *TF* [159], glucose metabolism genes *INS* and *PPARA* [160], estrogen related *CYP19A1* and anti-inflammatory *IL10* [161], and partially replicated interactions between noradrenergic related *DBH* gene and pro-inflammatory genes *IL6* and *IL1A* [162]. Epistasis between inflammation

markers has also been shown by other groups [163]. Of note, many of these loci only associated with AD after including their interaction, but not on their own, thus supporting the notion that epistasis can explain some of the missing heritability.

1.3 Cholesterol in Alzheimer's disease

1.3.1 Cholesterol metabolism in the brain

The brain is the most cholesterol-rich organ in the human body and accounts for about 20% of the whole body's cholesterol [164]. Most of it is found in oligodendrocyte myelin sheets (70-80%) and the rest is located in astrocytic and neuronal plasma membranes. With an intact blood brain barrier, brain cholesterol metabolism is independent from the periphery and cellular levels are maintained through a complex interchange between *de novo* synthesis, import, storage and secretion (**Figure 1.6** and **1.7**) [164].

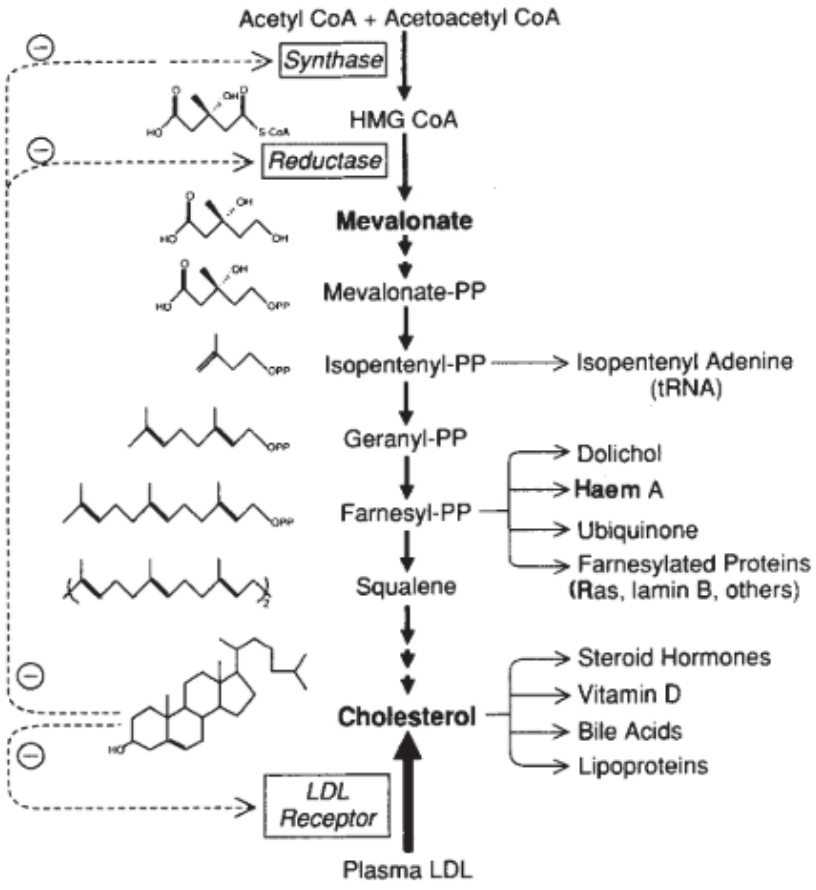
Synthesis: In the adult brain, *de novo* synthesis of cholesterol mainly occurs in glial cells although some production can also occur in neurons [164]. Synthesis is a resource-intensive process and is achieved through the mevalonate pathway (**Figure 1.6**). Briefly, acetyl-CoA is converted to HMG-CoA by HMG-CoA synthase which is then further reduced to mevalonate by HMG-CoA reductase (HMGCR) [165]. Mevalonate can be further processed in a series of enzymatic reactions to generate cholesterol as well as other non-sterol isoprenoids [164,165]. The rate-limiting step in this pathway is the conversion of HMG-CoA to mevalonate handled by the HMGCR enzyme.

Import: With low synthesis rates in neurons, the cholesterol need can be met by import from the extracellular environment. In astrocytes, cholesterol can be synthesized *de novo* or recycled after uptake from the extracellular matrix (**Figure 1.7**). ATP-binding cassette (ABC) transporters such as ABCA1, ABCA7, and ABCG1 coordinate the redistribution of the cholesterol to the plasma membrane. There they facilitate combination of the cholesterol with apolipoproteins, such as APOE. The lipoprotein lipase (LPL) hydrolyzes the particles that ultimately result in functional high-density lipoprotein (HDL) - particles similar to the ones found in the periphery [134,164,166,167]. The HDL

particles can then migrate in the extracellular space and bind to receptors such as low-density lipoprotein receptor (LDLR), low-density lipoprotein receptor related protein 1 (LRP1), and sortilin related receptor 1 (SORL1) on ependymal, glial, or neuronal cells [134,166]. The receptor-HDL complexes are endocytosed, and the cholesterol can be redistributed in the target cell (e.g. neurons) [134,164,166].

Storage: Excess cholesterol can be esterified by acetyl-CoA acetyltransferase 1 (ACAT1, a.k.a. SOAT1) which can then be stored in intracellular lipid droplets that act as storage (**Figure 1.7**) [164]. About 1% of the total cholesterol content is stored this way.

Figure 1.6 The mevalonate pathway

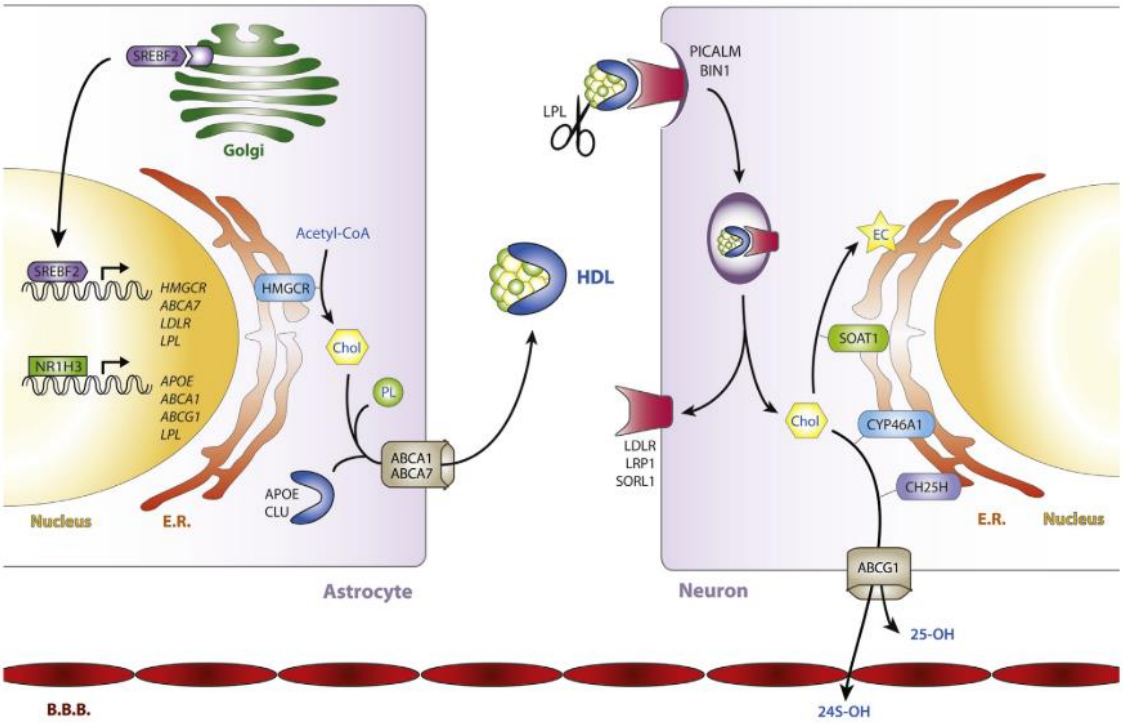


Mevalonate pathway produces cholesterol and non-sterol isoprenoids (shown on the right) and are tightly regulated by negative feedback loops. Figure reproduced with permission from [165].

Excretion: Excess cholesterol can also be excreted. In neurons, this can occur through two pathways. Either through a similar mechanism as described above for astrocytes – with the help of ABC transporters for combination with lipoproteins, or through conversion to oxysterols (**Figure 1.7**) [164]. In the latter process, cholesterol is hydroxylated by cholesterol 24-hydroxylase (CYP46A1), producing 24-hydroxycholesterol [164]. In contrast to cholesterol, 24-hydroxycholesterol can more readily cross the lipophilic plasma membranes and can thus be directly secreted across the blood brain barrier.

Homeostasis: After the bouts of cholesterol synthesis that accompany myelination during the perinatal period and adolescence, the rate of cholesterol synthesis is considerably decreased. Cellular

Figure 1.7 Cholesterol homeostasis in the brain



Schematic representation of cholesterol metabolism in the brain. Genes (italicized) and gene products are identified in black, whereas other molecules or cellular compartment are depicted in colors. Figure reproduced with permission from [166].
Abbreviations: 24S-OH, 24S-hydroxycholesterol; 25-OH, 25-hydroxycholesterol; Chol, cholesterol; EC, esterified cholesterol; E.R., endoplasmic reticulum; HDL, high-density lipoprotein; PL, phospholipids.

cholesterol levels are then regulated by feedback mechanism that balances the synthesis, import, and excretion [164,165]. Reductions in cholesterol levels are detected by sterol regulatory element binding factors (SREBFs) that act as transcription factors to upregulate expression of genes involved in the synthesis and uptake of cholesterol [165,167]. Conversely, excess cholesterol inhibits SREBFs [166]. In line with this, endogenous synthesis decreases as exogenous cholesterol are increased [135].

1.3.2 Hypercholesterolemia and statin use in AD

As one of the cardiovascular factors influencing AD, hypercholesterolemia is an attractive candidate due to the readily available drugs to treat it: statins. The effect of hypercholesterolemia (or increased serum or plasma levels of cholesterol) and statins have been investigated in population studies, randomized controlled trials, *in vivo* using animal models, and *in vitro.*

AD risk: The effect of hypercholesterolemia on AD has been extensively studied and findings were initially contradicting (**Supplemental sTable 1.1** in Appendix 1). However, it is now widely accepted that midlife (about 40-50 years of age) elevated total cholesterol (TC) levels are associated with increased risk of AD [16–21], though a few studies have reported negative findings [168–170]. In line with the former studies, early retrospective studies showed that statin use associated with lower risk of developing AD [22,23]. However randomized controlled trials with statins in AD subjects have failed to prove any beneficial effects [171–173].

TC levels later in life have yielded more contradictory results (**Supplemental sTable 1.1** in Appendix 1); TC levels have been linked to the prevalence of neuropathologically defined AD [174] but not to the clinically defined probable and possible AD diagnosis [175,176]. The directionality of effect is also debated with prospective studies investigating the incidence of AD showing TC levels to have no effect [177], reducing risk [175,176], or increasing risk [178]. Other studies have looked at differences in cholesterol levels in AD, compared to healthy controls, and found levels to be both increased [174,179,180] and decreased [181].

Findings of TC levels in AD are thus somewhat contradictory. Anstey et al., [182] performed a systemic review and meta-analysis of 18 prospective studies with a total sample size of 14,311 participants followed for 4.8 – 29 years. They concluded that high midlife TC levels and increased AD risk was indeed a consistent finding, whereas meta-analysis of late-life TC levels did not reveal any significant associations. These findings have subsequently been confirmed in an even larger study [183].

AD pathology: Multiple studies have shown associations between both midlife and late life high cholesterol and increased amyloid load in the brain [184–186]. Mild hypercholesterolemia in midlife associated with increased Aβ pathology post-mortem [184] and cross-sectionally in a pre-clinical AD cohort (PREVENT-AD) cholesterol levels associated with Aβ load in the brain as assessed by PET and CSF Aβ-42 levels [185]. Similar findings have been observed with late life cholesterol levels in AD cohorts; increased cholesterol levels associate with increased Aβ load as assessed by CSF Aβ-42 [187], brain levels of Aβ-42 [180], and brain neuritic plaque density [188]. Cholesterol has also been shown to associate with hypometabolism in brain regions affected by AD, including precuneus, parietotemporal, and prefrontal areas, as assessed by FDG-PET [189] but seems to have no effect on TAU pathology as no correlations were found between cholesterol and biomarkers of TAU pathology (CSF p-TAU and PET) [185] or NFT density [188].

1.3.3 Cholesterol and statins in in vitro and in vivo AD models

Cholesterol and the effect of statins have been extensively studied using *in vitro* and *in vivo* models of AD.

In vivo: In the early 90's, Sparks et al., first showed that a high cholesterol diet in rabbits induced Aβ pathology in the brain [24–26]. Follow-up studies found that the diet also increased plasma levels of cholesterol but not triglycerides [190,191], while the effect on brain cholesterol levels were inconsistent (no change or increased) [191,192]. Other changes were also evident; for example, increased intraneuronal APOE immunoreactivity that preceded the increase in Aβ [190,192,193], increased ventricular volume [194,195], increased microglia immunoreactivity [26,192], impaired behavior in a

classical conditioning task [196], and reduced markers of the cholinergic system [192]. The effect of a high cholesterol diet on Aβ pathology scaled with duration such that over time, AD pathology became more widespread and affected behavior [24,196]. In fact, even a low-level cholesterol diet administrated over a long period of time can result in Aβ deposition in the brain [197].

Due to the role of APOE in lipid homeostasis, several studies have also investigated the effects of a lipid diet in APOE mouse models. Many APOE mouse models exists and each differ in their expression patterns and phenotypes [198]. Further confounding the literature is the use of different controls, making studies difficult to compare. For example, one set of targeted replacement mice (replacing the mouse *APOE* locus with human *APOE*-ε3 and *APOE*-ε4 alleles) show ε4 mice having increased serum cholesterol levels but decreased brain cholesterol levels compared to wild type and ε3 mice on normal diets [199]. Conversely, another set show no difference in plasma cholesterol levels between ε3 and ε4 mice [200,201]. The latter set of mice have been characterized with regards to AD, and ε4 mice have been shown to display multiple AD phenotypes such as Aβ deposition and TAU hyperphosphorylation (reviewed in [198]). In addition, although cholesterol metabolism is unaltered on a normal diet, a high fat/high cholesterol diet induces a 5-fold increase in blood TC (compared to 1.5-fold in wild type mice). Thus investigating the interaction between *APOE* genotype and high cholesterol diet in these mice, there was a non-significant increase in Aβ immunoreactivity in ε4 animals compared to ε3 animals, which was further increased with the diet to a significant level [202].

There are other animal models of AD that are based on the mutations in the APP gene causing the familial form of AD. In these models, diet-induced hypercholesterolemia has been shown to increase and accelerate the accumulation of Aβ [27–29]. In a more extensive study, it was further shown that hypercholesterolemia impaired memory, reduced levels of synaptic markers, and also accelerated tau pathology [29].

In vitro: Modelling the effect of cholesterol levels on AD pathological processes *in vitro* is difficult in large part due to its complex regulation. Cellular cholesterol levels are maintained by intracellular

synthesis through the mevalonate pathway and extracellular delivery from cholesterol containing lipoprotein particles (**Figures 1.6** and **1.7**) [135]. Blocking either supply will cause compensation through the other. In addition, cholesterol is not the sole end-product of the mevalonate pathway with other non-sterol isoprenoids also being produced. These are used for post-translational modification of numerous proteins affecting processes such as protein trafficking and signaling, cell motility, cytoskeletal structure, and membrane transport [203].

Nevertheless, the effects of statins on Aβ and TAU in cell culture systems have been mainly beneficial. For example, statins seem to have a robust effect on Aβ pathology as several studies have shown that statin treatment reduce Aβ production [30,204–207] and promote the non-amyloidogenic pathway [31]. A similar beneficial effect has been shown on cell plasticity with statin treatment stimulating neurite outgrowth [208]. In contrast to these beneficial effects, other studies showed that statins increase levels of TAU and APP, and promote cell death [206,209,210]. These discrepancies were addressed in a recent study which concluded that the effects of statins were cell type specific and dose dependent [204].

Due to statins' effect on both cholesterol and non-sterol isoprenoids, it is debated which of these two are mediating the effects. Some of the literature's discrepancies could thus be explained by the effect of statins on the actual cholesterol and non-sterol isoprenoid levels. Two early studies showed that cholesterol reductions with intact non-sterol isoprenoid levels (achieved by statin treatment supplemented with mevalonate) led to reduced amyloidogenic processing and lowered intra- and extracellular levels of both Aβ-40 and Aβ-42, suggesting a cholesterol specific effect [30,31]. On the other hand, blockage of both cholesterol and non-sterol isoprenoid production was shown to increase p-TAU levels and promote cell death [209]. These changes were reversed by mevalonate or non-sterol isoprenoid supplementation, but not with cholesterol. These findings suggest that lowering cholesterol while keeping levels of non-sterol isoprenoids intact, would be beneficial on a cellular level.

In summary, both *in vivo* and *in vitro* studies support a role for cholesterol and statins in AD. However, their relationships are complex. Particularly evident from *in vitro* statin studies, there are non-cholesterol pathways that needs to be taken into consideration to tease apart cholesterol specific effects from effects associated with non-sterol isoprenoids.

1.3.4 Cholesterol related genetics in AD

Cholesterol metabolism in the etiology of AD has also been implicated in findings from genetic studies. Most notably, and as discussed above, APOE is the main lipid transporter in the brain and is also the most important genetic risk factor for AD (see 1.2.1 APOE and 1.3.1 Cholesterol metabolism in the brain). Genetic studies have however revealed several other loci involved in both AD and cholesterol or lipid regulation.

Evidence from GWA studies: Numerous of the loci implicated in AD from GWA studies are involved or connected to lipid metabolism [211]. In fact, in a recent GWAS meta-analysis, a gene ontology (GO) pathway analysis detected nine pathways in four clusters of GO terms of which one pertained to lipid metabolism (the other ones being APP metabolism/Aβ formation, tau protein binding, and immune response) [7]. Interestingly, out of the nine GO terms, five belonged to the lipid metabolism cluster: protein-lipid complex assembly, protein-lipid complex, reverse cholesterol transport, protein-lipid complex subunit, and plasma lipoprotein particle assembly [7]. Similar findings with an earlier GWAS meta-analysis [71] was used to create pathway specific polygenic scores [212]. The lipid polygenic score was found to correlate with increased Aβ pathology (CSF Aβ-42/Aβ-40 ratio and PiB PET) but not with TAU pathology (CSF p-TAU) or neurodegeneration (CSF TAU). However, after removing *APOE* from the score (leaving *CLU* and *ABCA7*) only the CSF Aβ-42/Aβ-40 association remained nominally significant [212].

Evidence from genetic variants influencing blood cholesterol: An attempt to investigate the link between the genetics underlying blood cholesterol levels and AD has also been done. Proitsi et al., [213] investigated polygenic scores based on results from blood cholesterol GWA studies.

Specifically, polygenic scores for TC, low-density lipoprotein cholesterol (LDL-C), HDL cholesterol (HDL-C) and TG levels were constructed and evaluated for association with AD risk. Each score associated significantly with its own trait explaining between 1.83% (LDL-C) and 4.34% (TG) of the variance but did not associate with AD risk.

Evidence from hypothesis-driven studies: In addition to GWAS identified loci, hypothesis-driven studies based on the empirical evidence of a link between cholesterol and AD have investigated multiple other loci. In a study of 28 cholesterol related genes, significant associations in four genes (*HMGCS2, FDPS, NPC2* and *ABCG1*) were detected that could be replicated in at least one additional cohort [214]. Another study investigated 17 cholesterol related genes for associations with pathology and related proxies and found consistent evidence for a protective effect of a SNP in the *SREBF2* loci. The rs2269657 T allele associated negatively with Aβ and TAU pathology as well as indices of neurodegeneration [166]. Rs2269657 was further shown to associate with reduced *SREBF2* RNA expression levels. *SREBF2* encodes a transcription factor that regulates many genes involved in the cholesterol synthesis (*HMGCR*), uptake (*LDLR*) and excretion (*ABCA7*, **Figure 1.7**).

Other single gene-targeted studies have further investigated SNPs in other cholesterol related genes such as *CYP46* [215,216] (enzyme converting cholesterol to 24S-hydroxycholesterol for elimination from the brain), *PCSK9* [217] (involved in controlling cholesterol uptake) and *LRP* [218] (a major receptor for APOE).

HMGCR; the target of statins: Multiple studies have focused on the *HMGCR* loci which encodes the 3-hydroxy-3-methylglutaryl-CoA reductase. This is due to its role in the rate-limiting step in cholesterol synthesis and because it is the target of statins [165]. For example, our lab as well as others have found a protective effect of the intronic *HMGCR* SNP rs3846662; the AA genotype associated with reduced risk and delaying disease onset by nearly four year [13,32–34]. Although findings are inconsistent, there is preliminary evidence for *HMGCR*'s interactions with both *APOE* genotype and sex, suggesting a protective effect specifically in female *APOE*-ε4 carriers [13,34]. Interestingly, the A

allele associates with alternative splicing of *HMGCR*, resulting in exclusion of exon 13 (Δ13-*HMGCR*) [32,33,37,38]; a transcript that if expressed alone (without the full length transcript (FL-*HMGCR*)) results in a protein completely devoid of activity [32].

Conversely, rs3761740, a promoter SNP, was shown to associate with increased risk of AD and increased rate of cognitive decline (A allele), specifically in *APOE*-ε4 non-carriers [35,36]. When co-transfected with *SREBF2* in cell lines, the A allele associated positively with HMGCR activity [35]. These findings are corroborated by another study reporting an association between rs5909, a 3-prime UTR SNP but in complete linkage with the rs3761740 SNP, and AD [219].

Taken together, these findings support the same hypothesis put forward for the protective effect of statins; rs3846662 associates with protection, hypothesized to be due to its association with the production of a protein with reduced activity while rs3761740 in the promoter associates with increased risk, possibly due to its positive effect on HMGCR activity.

1.3.5 Other evidence for the involvement of cholesterol in the etiology of AD

Although associations between hypercholesterolemia and TAU pathology have been sparse, there is evidence for a direct effect of altered cholesterol metabolism. The most striking evidence comes from another disease, Niemann-Pick type C (NPD-C). NPD-C is an autosomal recessive disease caused by mutations in either of the NPC intracellular cholesterol transporter 1 or 2 genes (*NPC1* and *NPC2*, respectively) resulting in maldistribution of cholesterol and other lipids both in the periphery and in the brain [220]. Similar to AD, neurons of NPD-C cases display intracellular accumulation of TAU and NFTs but are free from Aβ pathology [221]. Moreover, accumulation of cholesterol and NFTs seemed to correlate on a regional and cellular level in NPD-C cases [222]. A similar study was done but using brains from both AD individuals and NPD-C cases [223] and confirmed that cholesterol levels were higher in tangle bearing neurons than in adjacent tangle-free neurons.

These initial NPD-C findings suggest that cholesterol can directly affect TAU pathology without acting through amyloid. However, a more recent study investigated CSF levels of multiple Aβ

peptides (Aβ-38, -40 and -42), total TAU and p-TAU in a small group of NPD-C patients and controls and found increased levels of all Aβ species as well as an increased ratio of Aβ-42/Aβ-40 in NPD-C patients [224]. Similarly, levels of CSF TAU were elevated but no difference in p-TAU levels was detected. This study suggest that Aβ metabolism may be altered in NPD-C such that CSF levels are increased without having Aβ accumulation in the brain.

While these findings support an effect of lipids in Aβ processing and TAU pathology, the relationship between the processes need to be further investigated to draw conclusions of causality and the independency of mechanisms leading to each type of pathology.

1.4 Summary

To summarize, the sporadic form of AD is a multifactorial disease that nevertheless is highly heritable. Most of the genetic variance contributing to AD is however unknown and identification of these genetic variants can contribute to understanding the etiology. The already identified AD variants are involved in processes such as APP metabolism/Aβ formation, tau protein binding, immune response, and most notably lipid metabolism. Among lipids, cholesterol is of special interest due to its many implications in AD. In the periphery, increased blood TC levels in midlife have been shown to associate with increased risk of developing AD possibly by increasing Aβ load in the brain and on a cellular level, increasing cholesterol levels can directly promote the production of Aβ peptides in neuronal cell types. Contrary, randomized controlled trials of statins in AD, have failed to show any clinical benefits and statins effect cellularly have been conflicting. Thus, in this thesis we aimed to identify some of the unknown genetic variants by specifically investigating genetic variants relating to cholesterol metabolism and by doing so, provide support for the role of cholesterol in AD (for specific rationale and objectives, see Introduction).

Chapter 2

Effects of genetic variants in the HMGCR locus on Alzheimer's disease

As an important regulator of cholesterol and the target of statins, the *HMGCR* gene is an interesting locus for evaluation in the context of AD. Multiple variants in the *HMGCR* gene locus have been investigated in relation to AD. Most notably two variants have been identified to both associate with HMGCR activity and risk of developing AD. With the event of whole genome sequencing, it is now possible to analyze an increasing number of variants and in this chapter, we thus set out to investigate the full *HMGCR* gene locus using such a data set. Due to the influence of APOE on cholesterol metabolism, we further wanted to investigate the interaction between *HMGCR* variants and *APOE*-ε4 status. We thus identified a variant, rs72633963, that associated with AD risk only in *APOE*-ε4 carriers, thus confirming the interaction between the *HMGCR* and *APOE* loci. Although our initial analysis indicated the A allele to associate with increased risk, using two Quebec based cohorts, we found evidence for protection in these cohorts; the A allele associated with reduced Aβ pathology and improved cognition specifically in *APOE*-ε4 carriers. These findings add to the literature on *HMGCR* variants and their role in AD.

2.1 Introduction

The protein 3-Hydroxy-3-Methylglutaryl-Coenzyme A Reductase (HMGCR) has been implicated in AD. The protein is encoded by the *HMGCR* gene and it is an enzyme that constitutes the rate-limiting step in cholesterol and isoprenoid synthesis [165,167]. The leading hypothesis is that reducing its activity, either pharmacologically by the use of statins [22,23] or genetically [13,32], is protective in AD.

Pharmacological inhibition of HMGCR has proven to be beneficial in AD in retrospective studies; statin use associated with a 0.29 relative risk of AD [22] and the prevalence of AD in statin users was 69.6% lower than in the total population [23]. Further, inhibition of HMGCR by statins *in vitro* reduces the production of Aβ [30,204,205] and increases processing of APP through the non-amyloidogenic pathway [31]. Similarly, *in vivo* studies have shown that statin treatment results in decreased Aβ in the brain [30,225].

Multiple *HMGCR* genetic variants have been investigated with regards to AD and to the regulation of *HMGCR* itself. Our lab and others have established a protective effect of the rs3846662 SNP in AD; the A allele has been associated with reduced risk of AD and delaying disease onset by nearly 4 years [13,32–34]. There is also evidence for interactions with sex and *APOE*-ε4 status such that protection is particularly occurring in female *APOE*-ε4 carriers [13]. This SNP is intronic and its A allele has been suggested to promote the alternative splicing of *HMGCR* resulting in the Δ13-*HMGCR* transcript [32,33,37,38]. Furthermore, expression of Δ13-*HMGCR* without the full length transcript results in a protein completely devoid of activity [32].

Rs3761740, a promoter SNP, was shown to associate with increased risk of AD and increased rate of cognitive decline (A allele), specifically in *APOE*-ε4 non-carriers [36]. There was no effect on HMGCR activity as assessed by a luciferase assay. Another group examined this in a Swedish cohort and found a compounded effect of the A allele and *APOE*-ε4 status; compared to non-carriers of both *APOE*-ε4 and the A allele, *APOE*-ε4 carriers that did not carry the A allele had 4.57 times greater risk of AD and carriers of both *APOE*-ε4 and the A allele had a 6.21 times increased risk of AD [35]. In this study, the A

allele associated positively with HMGCR activity when co-transfecting cell lines with *SREBF2*, a transcription factor controlling the expression of *HMGCR*. Another study reported the association of rs5909, a 3-prime UTR SNP, with AD [219]. It was noted that this variant is in complete linkage with rs3761740 and could thus represent the same effect.

Alterations of HMGCR activity influence blood lipid metabolism. For example, statin use is well-known to cause reductions in cholesterol levels [165,167] and all three SNPs (rs3846662, rs3761740 and rs5909) associates with cholesterol levels [32,226,227]. Multiple studies have shown associations between mid-life hypercholesterolemia and AD [16–20] and cardiovascular risk factors are recognized as important risk factors also for AD [1].

Taken together these findings suggest that genetic variants within the *HMGCR* gene locus can regulate *HMGCR* expression levels and/or activity that in turn affects lipid metabolism, resulting in altered risk for AD. In this study we used a whole-gene approach to identify SNPs in the entire *HMGCR* gene locus that associates with AD. Further we wanted to evaluate the effect of any identified SNPs on HMGCR levels, lipid metabolism, as well as AD pathology and biomarkers.

2.2 Material and methods

2.2.1 Cohort descriptions

PREVENT-AD: The Pre-symptomatic Evaluation of Novel or Experimental Treatments for Alzheimer's Disease (openpreventad.loris.ca/) cohort, based at the Centre for Studies on the Prevention of AD in Montreal, Canada (StoP-AD, douglas.research.mcgill.ca/stop-ad-centre) is a longitudinal study of older, healthy individuals (55+) with a parental or multiple-sibling history of AD [228]. Data for all variables were obtained from data release 3.0 (2016-11-30). Each participant and study partner provided written informed consent. All procedures were approved by the McGill University Faculty of Medicine Institutional Review Board and complied with ethical principles of the Declaration of Helsinki.

ADNI: Data used in the preparation of this article were obtained from the Alzheimer's Disease Neuroimaging Initiative (ADNI) database (adni.loni.usc.edu). The ADNI was launched in 2003 as a public-private partnership, led by Principal Investigator Michael W. Weiner, MD in the U.S. The primary goal of ADNI has been to test whether serial MRI, PET, other biological markers, and clinical and neuropsychological assessment can be combined to measure the progression of MCI and early AD. For up-to-date information, see www.adni-info.org. For this study, a subset of ADNI consisting of individuals with whole-genome sequencing data were used. All data were downloaded on December 3, 2015.

Ext-QFP: The extended Quebec Founder Population is a cohort of Quebecoise individuals that are (QFP) or are not descendants (extended) of the French settlers who colonized Nouvelle France between 1608 and 1759 [229–231]. Brain tissue was donated to, and obtained from, the Douglas-Bell Canada Brain Bank (Montreal, Canada). Collection of tissue and analyses thereof, conformed to the Code of Ethics of the World Medical Association and was approved by the Ethics Board of the Douglas Hospital Research Centre. Each participant, or his or her legal guardian, signed an informed consent form.

2.2.2 Genetic data

Quality control, principle component analysis (PCA), whole-gene analysis, and SNP extraction were done with the PLINK tool set (www.cog-genomics.org/plink/1.9/) [232,233].

Genotyping: PREVENT-AD was genotyped using the Illumina Infinium Omni2.5M-8 array (Illumina, San Diego, CA, USA). In ADNI we used whole-genome sequencing data downloaded from adni.loni.usc.edu. Briefly, genomic DNA samples derived from blood were sequenced at Illumina on the Illumina HiSeq2000, details are published elsewhere [234]. Ext-QFP was genotyped using DNA extracted from brain tissue or blood lymphocytes with the Illumina HumanHap 550k Beadchip (Illumina, San Diego, CA, USA), details are published elsewhere [235].

Quality control: Genetic data were filtered to exclude sex mismatches, sample and SNP missingness > 5%, sample heterozygosity and Hardy-Weinberg disequilibrium ($p < 0.001$).

Population stratification (PCA): ADNI is a much more ethnically diverse cohort compared to PREVENT-AD and Ext-QFP that almost exclusively have European ancestry (data not shown). In order to account for population stratification, ADNI individuals was evaluated for European ancestry by a genetic PCA using the 1000 Genomes as a reference panel [236].

Imputation: rs72633963 were imputed for PREVENT-AD and ext-QFP using the Sanger Imputation Service [237] (imputation.sanger.ac.uk/). Briefly, quality controlled genetic data was uploaded and pre-phased with SHAPEIT2 [238] and imputed with positional Burrows-Wheeler transform [239] using the 1000 Genomes cohort [236,240] as a reference panel.

2.2.3 Whole-gene exploratory analysis in ADNI

The ADNI data set was used to extract all SNPs within the *HMGCR* gene locus. Briefly, the *HMGCR* gene and an upstream gene encoding an antisense RNA (*CTD2235C13.2*) that overlaps with *HMGCR*, were selected and SNPs within the genes and in flanking regions of 1000 bp were extracted. Dominant effects of each SNP (carrier *vs* non-carriers of minor allele) were evaluated for association with last visit AD status using a logistic regression model corrected for age and sex, and stratified for *APOE*-ε4 status.

2.2.4 Dependent variables

***HMGCR* RNA**: The quantification of frontal cortex *HMGCR* RNA levels (ext-QFP) is described elsewhere [241]. Briefly, RNA was extracted from ext-QFP frontal cortex samples and levels of *HMGCR* mRNA were determined using Taqman-qPCR. ADNI blood *HMGCR* mRNA levels were extracted from a micro-array data set ("ADNI_Gene_Expression_Profile.csv", 2015-04-27) and has been described elsewhere [234]. Briefly, total RNA was extracted from peripheral blood samples and analyzed with the Affymetrix Human Genome U219 Array (Affymetrix, Santa Clara, CA, USA).

Lipids: Lipid levels were assessed as part of standard blood screening labs. In ADNI, fasting blood samples were used to determine TC levels. In PREVENT-AD, non-fasting plasma levels of TC as well as LDL-C and HDL-C were assessed.

Aβ-42, p-TAU and TAU in the CSF: In both PREVENT-AD and ADNI, CSF was obtained by lumbar puncture following an overnight fast. Levels of Aβ-42, p-TAU (phosphorylated at threonine 181) and total TAU were then measured by the Innotest® enzyme-linked immunosorbent assays (ELISA, Fujirebio) [242] and the Roche Elecsys CSF immunoassays (data file UPENNBIOMK9_04_19_17.csv) [243,244], for PREVENT-AD and ADNI respectively. Of note, the Elecsys Aβ-42 CSF immunoassay is currently under development for investigational use only and has an upper technical limit of 1700 pg/ml. Values above this limit are based on extrapolation of the calibration curve, and the performance of these values has not been formally established. These are still included in this study. Baseline levels were used for both cohorts.

Cognition and AD diagnosis: Cognition in PREVENT-AD was assessed with the Repeatable Battery for the Assessment of Neuropsychological Status (RBANS), that assess neurocognitive status over five domains: immediate memory, visuospatial/constructional, language, attention, and delayed memory [245]. Values for baseline was used. In ext-QFP, individuals were diagnosed according to the Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition criteria. For a subset of autopsied brains from the ext-QFP where neuropathological examination was possible, definite AD was diagnosed according to NINCDS-ADRDA criteria [246]. In ADNI, individuals were diagnosed according to the later version of the same criteria [41].

AD pathology: AD pathology was assessed in a subset of the ext-QFP cohort. Neuritic plaque and NFT densities were quantified in hippocampal CA1, subiculum, parasubiculum, fusiform gyrus and frontal and parietal cortices. Details on quantification has been described elsewhere [247].

2.2.5 Analyses and statistics

All analyses were performed in R [248]. Data was handled using the "tidyverse" package [249] and plotted with the "ggplot2" package [250]. Data was analyzed using the "car" package [251]. A complete list of data sets, software, and R packages used is available in **Supplemental sTable 2.1** in Appendix 2.

Descriptives: Cohort characteristics were summarized using the "psych" package [252] and differences between cohorts were assessed with χ^2-tests (categorical variables), t-tests (continuous variables comparing two cohorts) or ANOVAs (continuous variables comparing all three cohorts). In the case of significant differences between the three cohorts, pairwise χ^2-tests (using the "rcompanion" package [253]) or TukeyHSD post-hoc analysis was used to determine which cohorts differed.

Continuous dependent variables: For each dependent variable, possible covariates were evaluated with a linear regression and included in further analysis if $p \leq 0.05$. Outliers were identified and removed if a value was outside ± 2 SDs of the mean. Data were then analyzed with a two-factor (*APOE*-ε4 status and rs72633963 A status) ANOVA or ANCOVA (depending on inclusion of covariates) using the "car" package [251] and checked for normality with the Shapiro-Wilk test. In the case of not fulfilling the normality assumption (Shapiro-Wilk test $p \leq 0.01$), the dependent variable was transformed (log10 or square root). If, after transformation, normality assumption was still violated, the data was stratified for the *APOE*-ε4 status and evaluated using the non-parametric Mann-Whitney test. Post hoc pairwise comparisons were evaluated with the "emmeans" package [254] using the Tukey method if $p \leq 0.100$.

Binary dependent variables: Variables such as AD status were evaluated using logistic regressions corrected for age and sex, with stratification for *APOE*-ε4 status.

Survival analyses: Effect of rs72633963 on conversion rate (ADNI) and age of onset (ext-QFP) was assessed using Kaplan-Meier analyses with a log-rank test [255] after stratification for *APOE*-ε4 status, using the "survival" package [256,257]. In ADNI, a subset was selected as follows: individuals were included if being either healthy or having an MCI diagnosis at baseline while also having data for the

48 month visit and having not "reverted" (e.g. gone from AD to MCI, or from MCI to healthy during follow up). For both ext-QFP and ADNI, the event was defined as a clinical diagnosis of AD and the time of event was months after baseline in ADNI and age of onset in ext-QFP.

2.3 Results

2.3.1 Cohort characteristics

HMGCR SNPs were evaluated for associations with AD in three different cohorts; PREVENT-AD, ADNI and ext-QFP. Cohorts were compared on proportions of females, *APOE*-ε4 carriers and rs72633963 A carriers, as well as age and blood TC levels (**Table 2.1**). Percentage females differed significantly between all cohorts with PREVENT-AD having the highest percentage (70.9%) followed by ext-QFP (53.8%), followed by ADNI (42.5%, ps ≤ 0.012). The cohorts also differed in age with PREVENT-AD (62.4 years) being significantly younger than ADNI (73.6 years) that were younger than ext-QFP (78.2 years, ps ≤ 0.001). The proportion of *APOE*-ε4 carriers were different between cohorts (p = 0.042) with post-hoc analyses trending for a significantly higher proportion in ext-QFP compared to both PREVENT-AD and ADNI (ps ≤ 0.075). TC levels were available for PREVENT-

Table 2.1 Cohort characteristics

	PREVENT-AD	ADNI	ext-QFP	p	Post-hoc
N (healthy/MCI/AD)	110/0/0	253/443/45	47/0/105		
Females [%]	70.9 (4.4)	42.5 (1.8)	53.8 (4.0)	< 0.001[a]	PREVENT-AD > QFP > ADNI
Age [years][#]	62.4 (0.52)[d]	73.6 (0.26)	78.2 (0.71)	< 0.001[b]	PREVENT-AD < ADNI < QFP
***APOE*-ε4 carriers** [%]	35.5 (4.6)	41.2 (1.8)	50.0 (4.0)	0.042[a]	QFP > PREVENT-AD & ADNI
rs72633963 A carriers [%]	21.8 (4.0)	22.0 (1.5)	18.1 (3.1)	0.552[a]	
TC levels [mM]	5.4 (0.086)[d]	5.0 (0.037)[e]	NA	< 0.001[c]	

Data are presented as mean (standard error of mean).
Abbreviations: MCI, mild cognitive impairment; AD, sporadic AD; TC, total cholesterol.
[#] Mean age was calculated for baseline in PREVENT-AD and ADNI and for age of death in ext-QFP.
[a] Calculated with χ^2 test
[b] Calculated with ANOVA
[c] Calculated with Welch two sample t-test
[d] missing data for 1 individual
[e] missing data for 8 individuals

AD and ADNI and were found to be significantly higher in PREVENT-AD (5.4 ± 0.086 mM) compared to ADNI (5.0 ± 0.037 mM, $p < 0.001$). The number of rs72633963 A carriers did not differ between cohorts and the proportions ranged between 18.1% - 22.0% ($p = 0.552$).

2.3.2 HMGCR SNPs and AD risk in ADNI

The *HMGCR* gene locus was extracted from whole-genome sequencing data from ADNI. To evaluate the effect of each SNP on AD status a logistic regression stratified for *APOE*-ε4 status and corrected for age and sex was performed. Due to sample size, dominant effects of the minor alleles were evaluated (i.e. carriers vs non-carriers). It revealed two SNPs – rs72633963 and rs61303403 – significantly associating with AD in *APOE*-ε4 carriers (**Table 2.2**); carriers of the minor alleles (A and T, respectively) were 2.57-2.70 (ps = 0.04) times more likely to develop AD than non-carriers in addition to their increased risk due to *APOE*-ε4.

The two significant SNPs are located upstream of *HMGCR* in the *CTD2235C13.2* gene encoding an antisense RNA (**Figure 2.1**). BLASTing of the exonic sequences reveal no alignment except for with *HMGCR* (data not shown). We evaluated the linkage disequilibrium (LD) between the two variants and found them to be highly linked with an $r^2 = 0.885$. Since rs72633963 is located in an exon and rs61303403 is intronic, we chose to proceed with rs72633963.

Table 2.2 HMGCR SNPs and AD risk

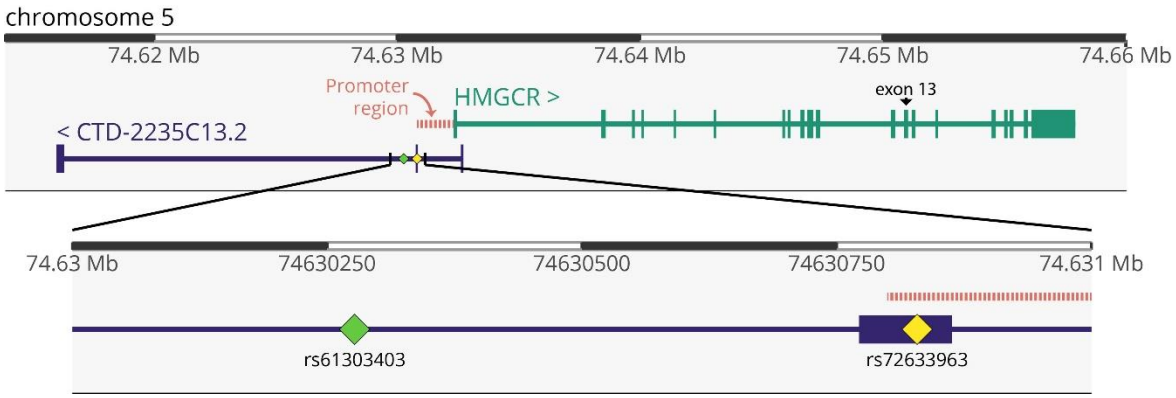
		***APOE*-ε4 non-carrier**				***APOE*-ε4 carrier**			
	Allele	**N**	**OR**	**T statistic**	**P**	**N**	**OR**	**T statistic**	**P**
ADNI - discovery									
rs72633963	A	210	1.01	0.021	0.983	180	2.57	2.056	0.040
rs61303403	T	210	1.20	0.494	0.621	180	2.70	2.051	0.040
ext-QFP - replication									
rs72633963	A	74	0.75	-0.468	0.639	78	0.32	-1.372	0.170

Abbreviations: OR, odds ratio (relating to mentioned allele); P, nominal p-value.

2.3.3 Rs72633963 and cholesterol metabolism

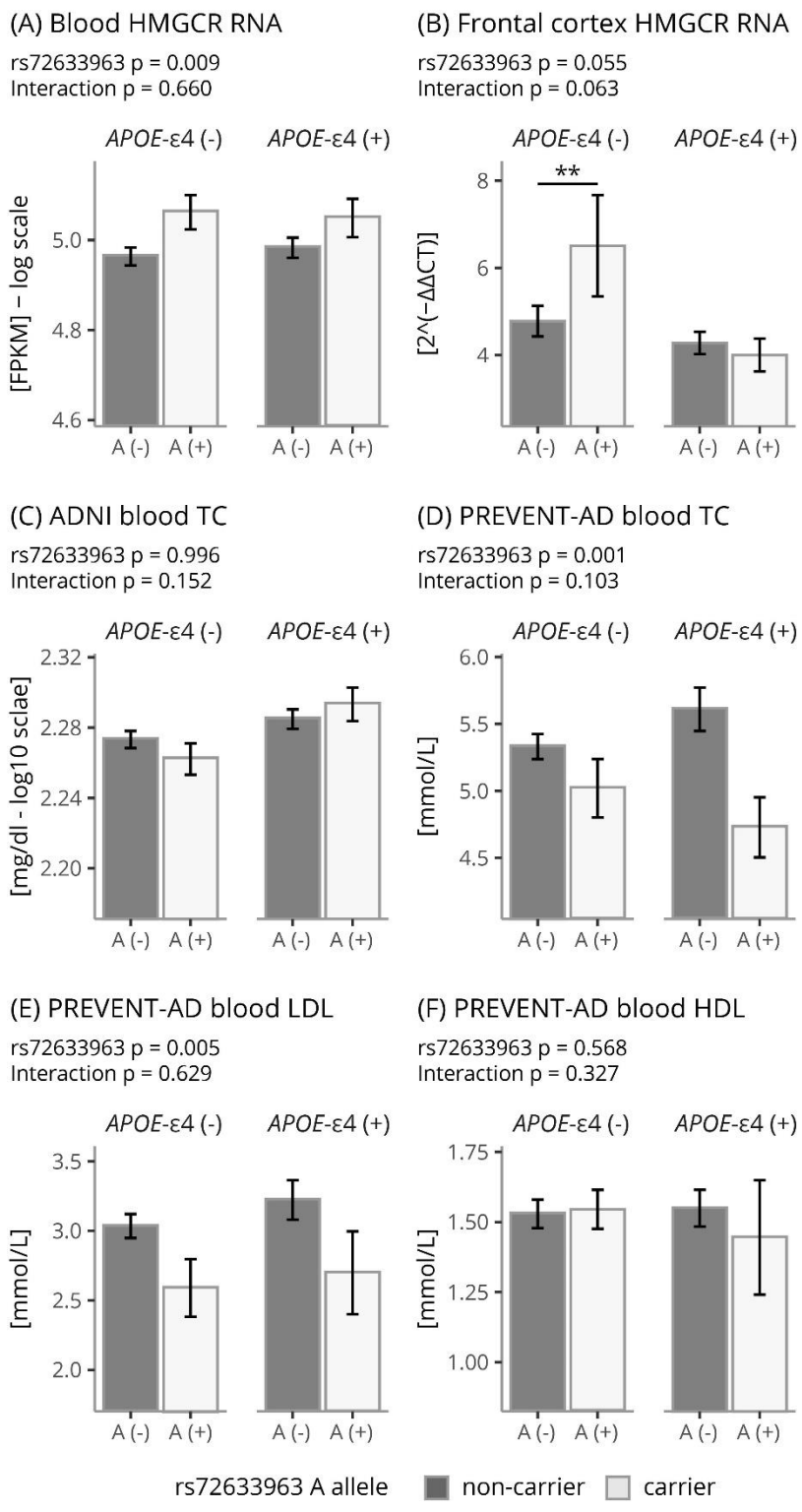
***HMGCR* expression**: Due to HMGCRs role in cholesterol metabolism and our hypothesis that SNPs might exert their effect by influencing expression of *HMGCR*, we examined the effect of rs72633963 on lipids and *HMGCR* expression in the blood and brain (**Figure 2.2 A** and **B**). Blood *HMGCR* RNA expression levels were available in ADNI. A two-factor ANCOVA corrected for sex, age and statin use, revealed a significant main effect of rs72633963 ($F(1, 661) = 6.84$, $p = 0.009$) with carriers of the A allele on average having higher *HMGCR* expression than non-carriers (**Figure 2.2 A**). In ext-QFP, *HMGCR* RNA expression levels were examined in the frontal cortex. A two-factor ANCOVA corrected for age and RNA integrity number found trends for significance for both the interaction term ($F(1, 85) = 3.538$, $p = 0.063$) and the main effect of rs72633963 ($F(1, 85) = 3.775$, $p = 0.055$; **Figure 2.2 B**). Post-hoc analysis revealed a significant effect of rs72633963 A allele in *APOE*-ε4 non-carriers (RQ values 6.51 ± 1.16 and 4.78 ± 0.35, for A carriers and non-carriers, respectively; $p = 0.009$) and no effect in *APOE*-ε4 carriers (RQ values 4.00 ± 0.38 and 4.28 ± 0.26 in A carriers and non-carriers, respectively; $p = 0.964$).

Figure 2.1 Azheimer's disease associated SNPs in the HMGCR locus



Larger blocks (**|**) indicate exons while thinner blocks (**–**) indicate introns. Broken blocks (**---**) indicate promoter with flanking regions.
Abbreviations: Mb, position in mega base pair; SNP, single nucleotide polymorphism.

Figure 2.2 Associations of rs72633963 and cholesterol metabolism



The effect of rs72633963 A allele status was assessed on cholesterol metabolism stratified for *APOE*-ε4 status; Blood HMGCR RNA levels in ADNI (A), frontal cortex HMGCR RNA levels in ext-QFP (B), blood TC levels in ADNI (C) and in PREVENT-AD (D), and LDL-C (E) and HDL-C (F) levels in PREVENT-AD. Error bars represent SE. For statistical details see 2.2.5 Analyses and statistics and 2.3 Results.
Abbreviations: (-)/(+), allele non-carriers/carriers; FPKM, fragments per kilobase of exon model per million reads mapped; SE, standard error of the mean; TC, total cholesterol.
** p ≤ 0.01 Tukey corrected post-hoc analysis

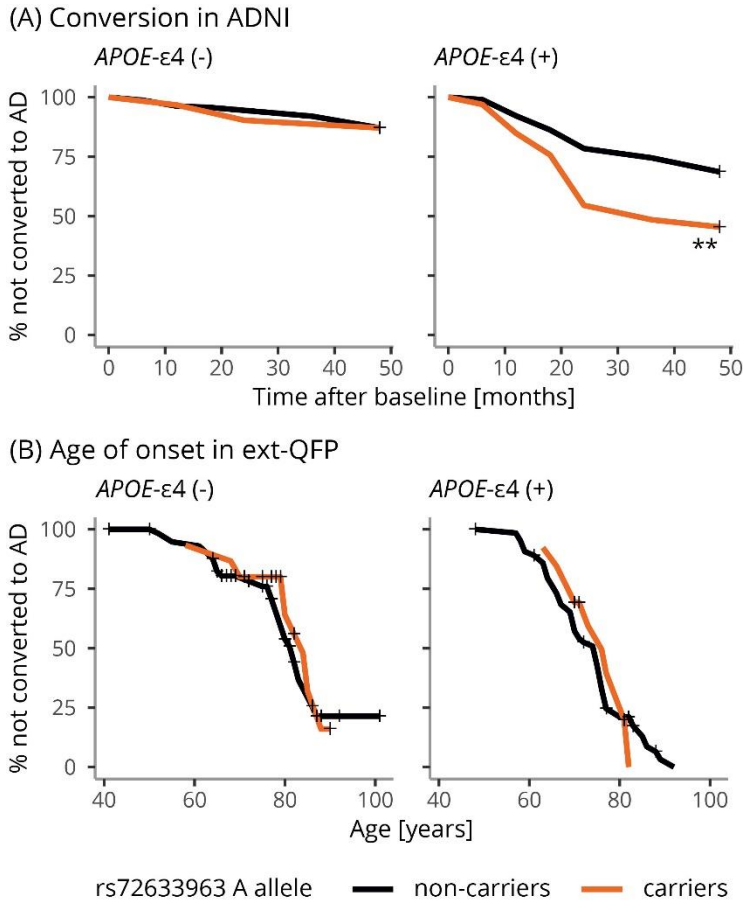
Blood lipid levels: The effect of rs72633963 was assessed on blood lipids using ADNI and PREVENT-AD. In ADNI, there was no significant effects of rs72633963 on TC levels as a main effect ($F(1, 724) < 0.001$, $p = 0.996$) or *APOE*-ε4 interaction ($F(1, 724) = 2.053$, $p = 0.152$) after correcting for sex, age and statin use (**Figure 2.2 C**). In PREVENT-AD, the effect of rs72633963 was evaluated on levels of TC, LDL-C and HDL-C at baseline (**Figure 2.2 D-F**). rs72633963 A allele significantly associated with decreased TC levels ($F(1, 100) = 11.392$, $p = 0.001$, corrected for sex) and decreased LDL-C levels ($F(1, 93) = 8.352$, $p = 0.005$, corrected for age and sex) while having no detectable effect on levels of HDL-C ($F(1, 99) = 0.328$, $p = 0.568$, corrected for sex). No interaction effects with *APOE*-ε4 status were detected for either of the PREVENT-AD lipid analyses ($Fs \leq 2.730$, $ps \geq 0.103$).

2.3.4 Rs72633963 and cognition and AD status

Conversion rate and age of onset: We next evaluated whether rs72633963 have an effect on conversion rate in ADNI or age of onset in ext-QFP (**Figure 2.3**). In ADNI, Kaplan-Meier survival analysis, stratified for *APOE*-ε4 status, revealed an accelerated conversion rate in rs72633963 A carriers compared to non-carriers in *APOE*-ε4 carriers ($\chi^2(1) = 6.7$, $p = 0.01$) whereas there was no effect in ε4 non-carriers ($\chi^2(1) = 0.0$, $p = 1.00$; **Figure 2.3 A**). In ext-QFP, we found no effect of rs72633963 A allele on age of onset in either ε4 non-carriers ($\chi^2(1) = 0.6$, $p = 0.445$) or ε4 carriers ($\chi^2(1) = 0.5$, $p = 0.468$; **Figure 2.3 B**). There was no difference in the proportion of females or age in the different strata (**Supplemental sTable 2.2** in Appendix 2).

Early cognitive decline: Rs72633963 was further assessed for effects on early cognitive decline in PREVENT-AD (**Table 2.3**). Cognition was assessed with the RBANS, that assess neurocognitive status over five domains: immediate memory, visuospatial/constructional, language, attention, and delayed memory [245]. We found a significant main effect of rs72633963 on the immediate memory index ($F(1, 97) = 4.228$, $p = 0.042$) after correcting for age and sex such that A carriers had higher scores than non-carriers, indicating better cognition. No significant effect of rs72633963 on the total score or the other indices could be detected.

Figure 2.3 rs72633963 associations with AD conversion rate and age of onset



Kaplan-Meier survival analyses stratified for *APOE*-ε4 status were used to determine differences in conversion rate in ADNI (A) and age of onset in ext-QFP (B) between rs72633963 A allele carriers and non-carriers. In ADNI, conversion from healthy/MCI to clinical AD was assessed in months after baseline whereas age of onset of clinical AD was used in ext-QFP. For statistical details see 2.2.5 Analyses and statistics and 2.3 Results.
Abbreviations: (-)/(+), allele non-carriers/carriers.
** $p \leq 0.01$

AD risk in ext-QFP: Rs72633963 was evaluated for AD risk in ext-QFP (**Table 2.2**). Contrary to our findings in ADNI, rs72633963 did not associate with AD, with ORs in A allele carriers of 0.32 and 0.75 in *APOE*-ε4 carriers and non-carriers, respectively (ps ≥ 0.170).

2.3.5 Rs72633963 and AD pathology

The effect of rs72633963 on AD pathology and proxies thereof were further assessed in all three cohorts.

Amyloid pathology: In ext-QFP, rs72633963 A allele carriers displayed lower levels of neuritic plaques compared to non-carriers ($F(1, 88) = 4.244$, $p = 0.042$) after correcting for sex (**Figure 2.4 C**).

Table 2.3 rs72633963 associations with cognition in PREVENT-AD

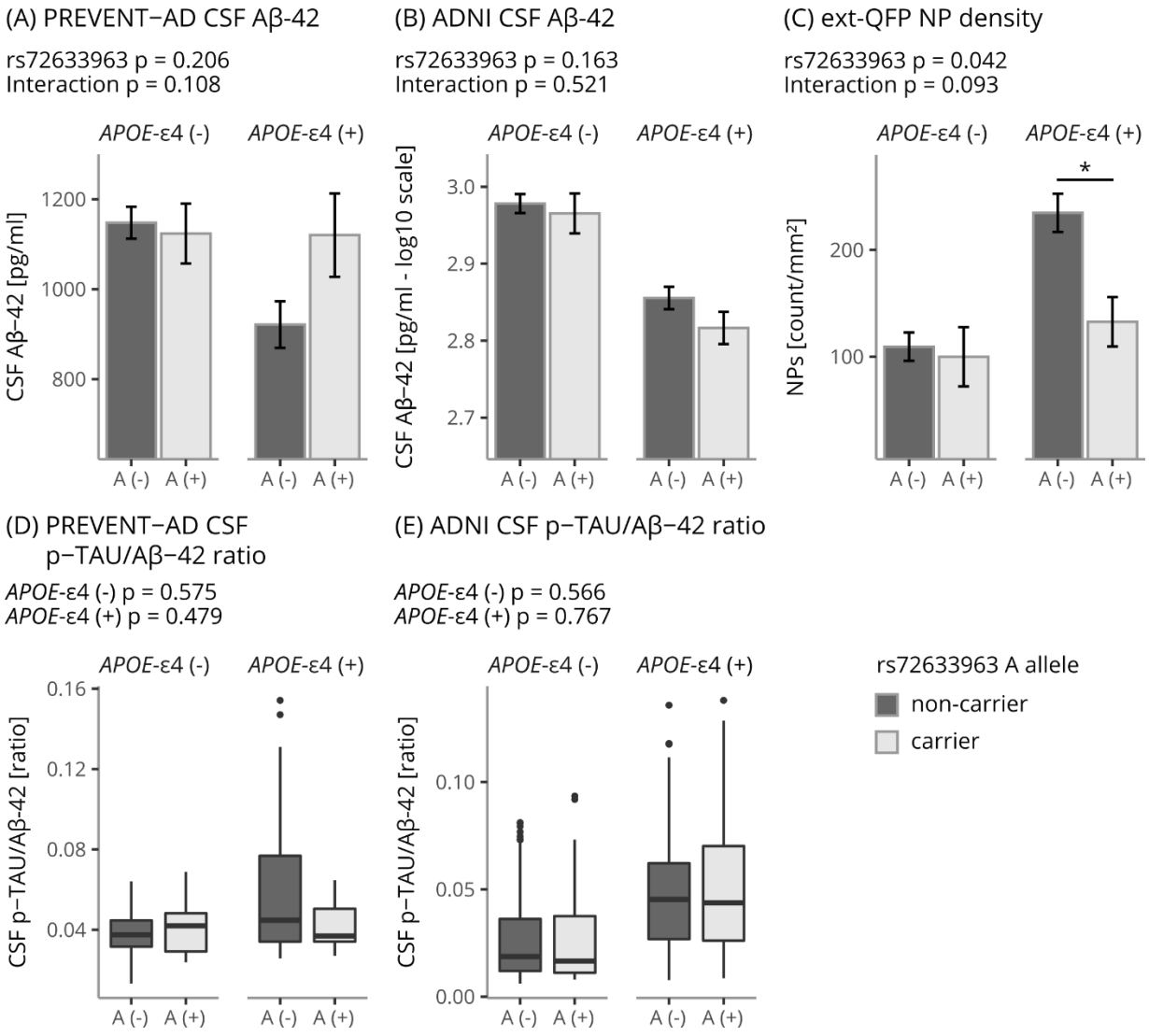
RBANS indexes	*APOE*-ε4 non-carrier				*APOE*-ε4 carrier					
	A non-carrier		A carrier		A non-carrier		A carrier			
	N	Mean (SE)	N	Mean (SE)	N	Mean (SE)	N	Mean (SE)	p(int)[#]	p(main)[$]
Total scale	51	98.88 (1.12)	14	100.36 (1.75)	29	100.52 (2.01)	8	105.75 (2.94)	0.505	0.137
Attention	51	104.78 (1.71)	14	103.29 (3.49)	30	104.60 (2.81)	8	111.00 (5.89)	0.284	0.502
Delayed memory	50	100.72 (1.01)	15	102.53 (2.18)	27	101.85 (1.42)	8	97.75 (4.75)	0.109	0.532
Immediate memory	52	102.29 (1.26)	14	104.21 (1.79)	29	101.07 (1.60)	8	108.25 (4.21)	0.345	0.043
Language	52	102.06 (1.36)	14	99.93 (2.00)	28	101.21 (1.39)	8	107.25 (5.08)	0.118	0.413
Visuospatial constructional	52	93.75 (1.42)	14	98.86 (2.37)	29	97.76 (2.34)	8	98.25 (5.71)	0.297	0.359

Abbreviations: A, rs72633963 A allele; RBANS, repeatable battery for the assessment of neuropsychological status; SE, standard error of the mean.
[#] p-value for the *APOE*-ε4 status * rs72633963 A status interaction.
[$] p-value for the main effect of rs72633963 A status.

There was a trend for an interaction effect ($F(1, 88) = 2.891$, $p = 0.093$) and post-hoc analysis revealed that the reduction was specific for *APOE*-ε4 carriers ($p = 0.013$) and not ε4 non-carriers ($p = 0.790$).

In PREVENT-AD and ADNI CSF Aβ-42 levels and the ratio of p-TAU/Aβ-42 were used as a proxy for amyloid pathology (**Figure 2.4 A**, **B**, **D**, and **E**). The latter has been shown to correlate very well with amyloid load in the brain as measured by PET [104]. In PREVENT-AD, we found a trend for an interaction effect ($F(1, 93) = 2.641$, $p = 0.108$) on Aβ-42, and examining the means show lower levels in A carriers in both *APOE*-ε4 non-carriers and carriers, but a greater difference in ε4 carriers (921.26 vs 1123.69 pg/ml in A non-carriers and carriers, respectively). There was no main effect of rs72633963 (**Supplemental sTable 2.3** in Appendix 2). For Aβ-42 in ADNI, and for p-TAU/Aβ-42 ratio in both cohorts, neither rs72633963 main effects nor interaction terms were significant (**Supplemental sTables 2.3** and **2.4** in Appendix 2).

Figure 2.4 rs72633963 associations with amyloid pathology



Amyloid pathology was assessed by CSF levels of Aβ-42 in PREVENT-AD (A) and ADNI (B), by density of neuritic plaques in the brain in ext-QFP (C), and by CSF p-TAU/Aβ-42 ratio in PREVENT-AD (D) and ADNI (E). Bar plots indicates ANOVA/ANCOVA analyses, boxplot indicates Mann-Whitney Wilcoxon test. Error bars represent SE. For statistical details see 2.2.5 Analyses and statistics and 2.3 Results.

Abbreviations: (-)/(+), allele non-carrier/carrier; CSF, cerebrospinal fluid; NP, neuritic plaque; SE, standard error of the mean.

* p ≤ 0.05 Tukey corrected post-hoc analysis

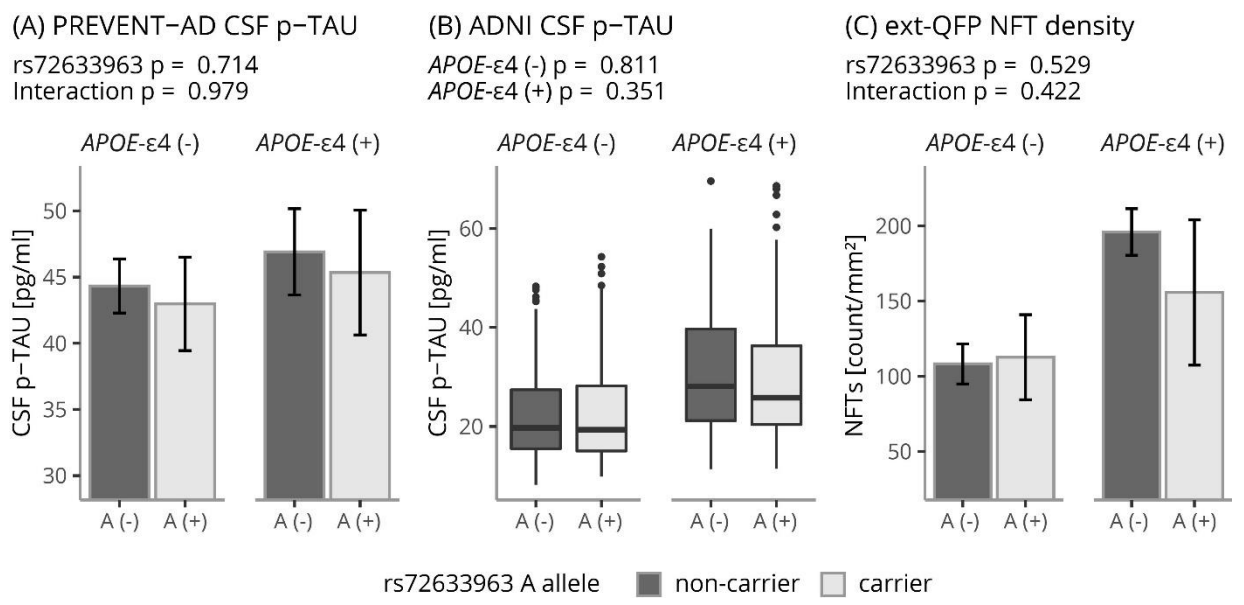
TAU pathology: NFTs and proxies of TAU pathology (CSF p-TAU levels) were assessed for associations with rs72633963 (**Figure 2.5**). We found no significant rs72633963 main or interaction effects in either cohort (Fs ≤ 0.652, ps ≥ 0.422; **Supplemental sTables 2.2** and **2.3** in Appendix 2).

Neurodegeneration: CSF TAU levels in PREVENT-AD and ADNI were used as proxies for neurodegeneration (**Figure 2.6**). We found no significant rs72633963 main or interaction effects in either cohort (Fs ≤ 0.485, ps ≥ 0.487, **Supplemental sTable 2.3** in Appendix 2).

2.4 Conclusion

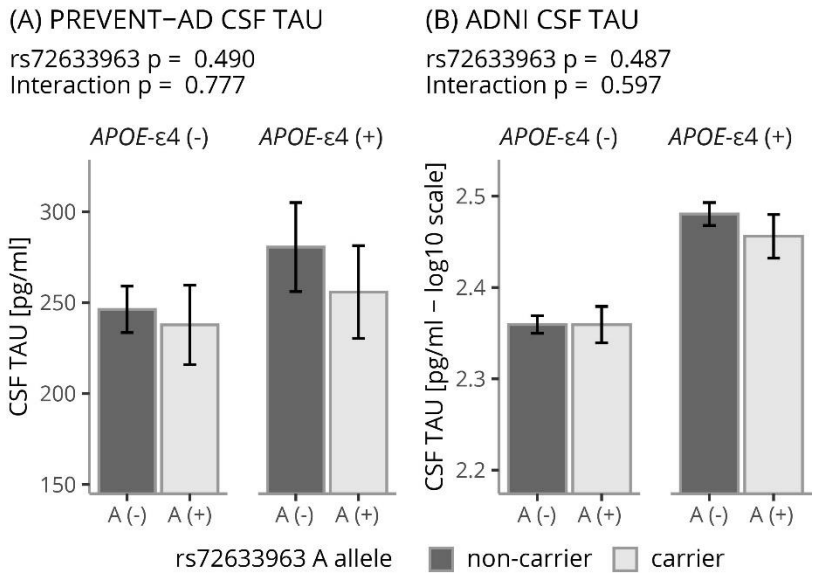
We identified a SNP in the *HMGCR* gene locus, rs72633963, that associates with increased risk in *APOE*-ε4 carriers in ADNI (**Table 2.2**). The SNP was further evaluated in two Quebec based cohorts: PREVENT-AD and ext-QFP. In these cohorts, the rs72633963 A allele associated with reduced Aβ

Figure 2.5 rs72633963 associations with TAU pathology



TAU pathology was assessed by CSF levels of p-TAU in PREVENT-AD (A) and ADNI (B), and by NFT density in ext-QFP (C). Bar plots indicates ANOVA/ANCOVA analyses, boxplot indicates Mann-Whitney Wilcoxon test. Error bars represent SE. For statistical details see 2.2.5 Analyses and statistics and 2.3 Results.
Abbreviations: (-)/(+), allele non-carrier/carrier; CSF, cerebrospinal fluid; NFTs, neurofibrillary tangles; SE, standard error of the mean.

Figure 2.6 rs72633963 associations with CSF TAU



Neurodegeneration was assessed by CSF levels of TAU in PREVENT-AD (A) and ADNI (B). Error bars represent SE. For statistical details see 2.2.5 Analyses and statistics and 2.3 Results.
Abbreviations: (-)/(+), allele non-carrier/carrier; CSF, cerebrospinal fluid; SE, standard error of the mean.

pathology in *APOE*-ε4 carriers (**Figure 2.4**) and improved cognition (**Table 2.3**). While not reaching significance, rs72633963 A allele trended for increased CSF Aβ-42 in *APOE*-ε4 carriers in PREVENT-AD (**Figure 2.4**) and the OR for AD risk in ext-QFP indicated a protective effect (**Table 2.2**). These changes were accompanied by increased levels of *HMGCR* expression in the blood (ADNI) and brain (ext-QFP), as well as decreased blood cholesterol and LDL-C levels (PREVENT-AD, **Figure 2.2**). Taken together, our findings suggest a risk profile in ADNI with rs72633963 A allele associating with increased risk n *APOE*-ε4 carriers, whereas in ext-QFP and PREVENT-AD results suggest a protective profile with the A allele associating with reduced cholesterol levels overall and reduced amyloid pathology in *APOE*-ε4 carriers.

The discrepancies between ADNI and the two Quebec based cohorts could be explained by a few factors. First, the cohorts differ in the proportion of females (**Table 2.1**) such that both PREVENT-AD and ext-QFP have higher proportions than ADNI. Sex could be interacting with SNPs in the *HMGCR* gene. In fact, Leduc et al., [13] showed using QFP, that rs3846662 AA genotype was protective in females, but not in males. Similarly, in a study of individuals with familial hypercholesterolemia

they showed that the AA genotype associated with higher blood levels of *HMGCR* RNA in females, whereas the opposite was observed in men [258]. In this study we did not examine interactions with sex, and it is thus possible that there is a sex**APOE*-ε4 status*rs72633963 interaction effect on both HMGCR RNA levels and AD risk.

Secondly, what could set PREVENT-AD and ext-QFP apart from ADNI, is the population structure, i.e. differences in allele frequencies between cases and controls due to systematic ancestry differences [259]. As both PREVENT-AD and ext-QFP are Quebec based cohorts and the fact that a large proportion of ext-QFP is part of a founder population, their population structure could differ. While we performed a PCA to limit ADNI to individuals with European ancestry, the population stratification within Europeans can still vastly differ and affect results. For example, in the actual QFP ~90% of descendants had French origin [260] which is unlikely to be the case for ADNI. In order to address this, we performed a genetic PCA also in ext-QFP and found that the population structure indeed looks very different in the two cohorts (**Supplemental sFigure 2.1** in Appendix 2). Although ext-QFP is mainly of French descent, it is also clear that there is a great deal of stratification within the cohort, probably reflecting the smaller founding effects within the cohort described before [229,231]. Further filtering of ADNI and ext-QFP would need to be done to evaluate the effect of population stratification.

Further, the discrepancies in risk and Kaplan-Meier analyses in ADNI and ext-QFP could also be due to ADNI being a clinical cohort and part of ext-QFP being pathologically confirmed AD cases. Sensitivity and specificity rates of clinical AD diagnosis are 81% and 70% respectively [41] suggesting that a large portion of clinically diagnosed AD cases will not receive a diagnosis of definite AD upon neuropathological examination. Thus, the ext-QFP cohort might more correctly represent neuropathological determined AD than ADNI. In an effort to address this, we refined control and AD status categorizing individuals as having or not having amyloid pathology using CSF level proxies of p-TAU and Aβ-42 (similar to [104]). Controls were thus defined as amyloid negative and free from

cognitive impairment and AD as amyloid positive with a clinical diagnosis of AD. Indeed, when doing this, the significance between rs72633963 A allele status and AD is lost in ADNI (**Supplemental sTable 2.5** in Appendix 2).

Thus, looking only at the findings from ext-QFP and PREVENT-AD, we see a reduction in amyloid pathology with the rs72633963 A allele. This is accompanied by a reduction in blood TC and LDL-C levels (**Figure 2.2 D** and **E**). This begs the question if the effect of rs72633963 on Aβ pathology are mediated by its effect on cholesterol levels. Although the interaction term for TC levels did not reach a significance level of $p \leq 0.05$, there was a trend ($p = 0.103$) suggesting that the effect of rs72633963 was stronger in *APOE*-ε4 carriers (**Figure 2.2 D**). Further, we analyzed *HMGCR* levels in the frontal cortex of a subset of ext-QFP individuals. This analysis revealed a trend for significant interaction ($p = 0.063$, **Figure 2.2 B**) and contrary to the Aβ pathology findings, post-hoc analysis revealed a significant effect in *APOE*-ε4 negative individuals. Our results thus suggest that rs72633963 might act through altering peripheral cholesterol metabolism rather than central. However, the measures of cholesterol metabolism (*HMGCR* RNA in ext-QFP and blood TC/LDL-C in PREVENT-AD) are not the same, and further analysis where similar measures are analyzed in the different tissues are needed in order to establish this.

In summary, we have found a genetic variant, rs72633963 (A allele), within the *HMGCR* gene that associates with reduced amyloid pathology and improved cognition in two Quebec based cohorts. We further show that this effect is true in *APOE*-ε4 carriers and not in *APOE*-ε4 non-carriers, and that this is possibly mediated by reductions in blood TC and/or LDL-C levels.

Chapter 3

The effect of rs3846662 genetic variant on HMGCR metabolism and Alzheimer's disease biomarkers in induced pluripotent stem cells

While the functional implications and causality of rs72633963 remains to be established, there is evidence for causality of another *HMGCR* variant: rs3846662. The AA genotype of this SNP has been associated with protection in AD and is thought to do so by promoting alternative splicing of *HMGCR* resulting in reduced activity of the HMGCR protein. While these studies have used commercially available cell lines derived from tumors or immortalized lymphocytes, the effect of this variant has not been established in somatic cells or in non-cancerous human brain cells. Establishing causality in cell lines as well as in cell types relevant to the natural state of cells in the body is important both for understanding the mechanisms by which genetic variants act but also to determine cell type specific effects. In this chapter we thus collected somatic renal epithelial cells that were converted into induced pluripotent stem cells that were subsequently used for differentiation into neural progenitor cells as well as neurons – cells relevant to the AD pathophysiological process. We were thus able to examine the effects of rs3846662 across multiple cell types from the same set of donors. While we confirmed the effect of rs3846662 on HMGCR metabolism we were underpowered to detect any potential effects on AD processes.

3.1 Introduction

Recently, our group described the protective effect of the rs3846662 SNP in the *HMGCR* gene on AD risk [13]. Specifically, the rs3846662 AA genotype was shown to delay AD onset by nearly four years and completely attenuate the increased risk seen in *APOE*-ε4 carriers [13]. These findings are supported by a number of other studies showing either increased risk associated with the A allele or decreased risk associated with the G allele [33,34,261].

Only a few studies have investigated the functional implications of rs3846662, where they have shown that it associates with the production of the alternatively spliced Δ13-*HMGCR* transcript in the liver of healthy subjects as well as in several cell lines [33,262]. The mechanism by which rs3846662 promotes alternative splicing was suggested by Yu et al., [263] that found that rs3846662 A allele alters a binding motif, resulting in the preferential binding of heterogeneous nuclear ribonucleoprotein A1 (HNRNPA1), which in turn was shown to increase expression levels of Δ13-*HMGCR* while having no effect on FL-*HMGCR*. FL-HMGCR peptides form functional tetramer structures and it has been shown that the insertion of Δ13-HMGCR compromises its resulting enzymatic activity [38]. Indeed, expression of Δ13-HMGCR peptides only, resulted in a protein completely devoid of activity [262]. In line with these findings, overexpression of *HNRNPA1* was shown to decrease HMGCR enzyme activity [263].

Reduced HMGCR activity has been associated with protection in AD (reviewed in [80,264], and see 1.3 Cholesterol in Alzheimer's disease). For example, statin use was shown to reduce risk of AD in retrospective human studies [22,23], however, in randomized control trials of statins in AD patients have failed to show any consistent beneficial effect [171,172]. Further, in a recent meta-analysis study, statins were confirmed as protective in AD [76]. The effects of reduced HMGCR activity on Aβ and TAU in cell culture systems have been mainly beneficial. Statins have been shown to stimulate neurite outgrowth [208], promote the non-amyloidogenic pathway, and reduce the production of Aβ [204–207].

Interestingly, the same cell line that was used to prove the correlation between rs3846662 and Δ13-*HMGCR* levels was also found to decrease Aβ pathology with statin treatment [33,204].

Aside from its implication in AD, rs3846662 is also highly correlated with peripheral blood cholesterol levels; the G allele associates with increased total and LDL cholesterol levels [226,227,265]. Interestingly, mid-life hypercholesterolemia is a suggested risk factor for AD [17–19,119,174,180,266], which is also supported by findings that cholesterol levels associate with Aβ pathology in the brain [179,184] and glucose hypometabolism in brain regions affected by AD [189].

In summary, there is strong evidence that rs3846662 can influence the alternative splicing of *HMGCR* to promote expression of Δ13-*HMGCR*. This transcript has been shown to be devoid of activity, strongly suggesting that rs3846662 can negatively influence the HMGCR pathway. In AD, there is strong evidence that inhibition of HMGCR is protective, possibly by regulating AD biomarkers Aβ and TAU on a cellular level. However, the causal relationship between rs3846662 and AD pathology has not been fully mapped.

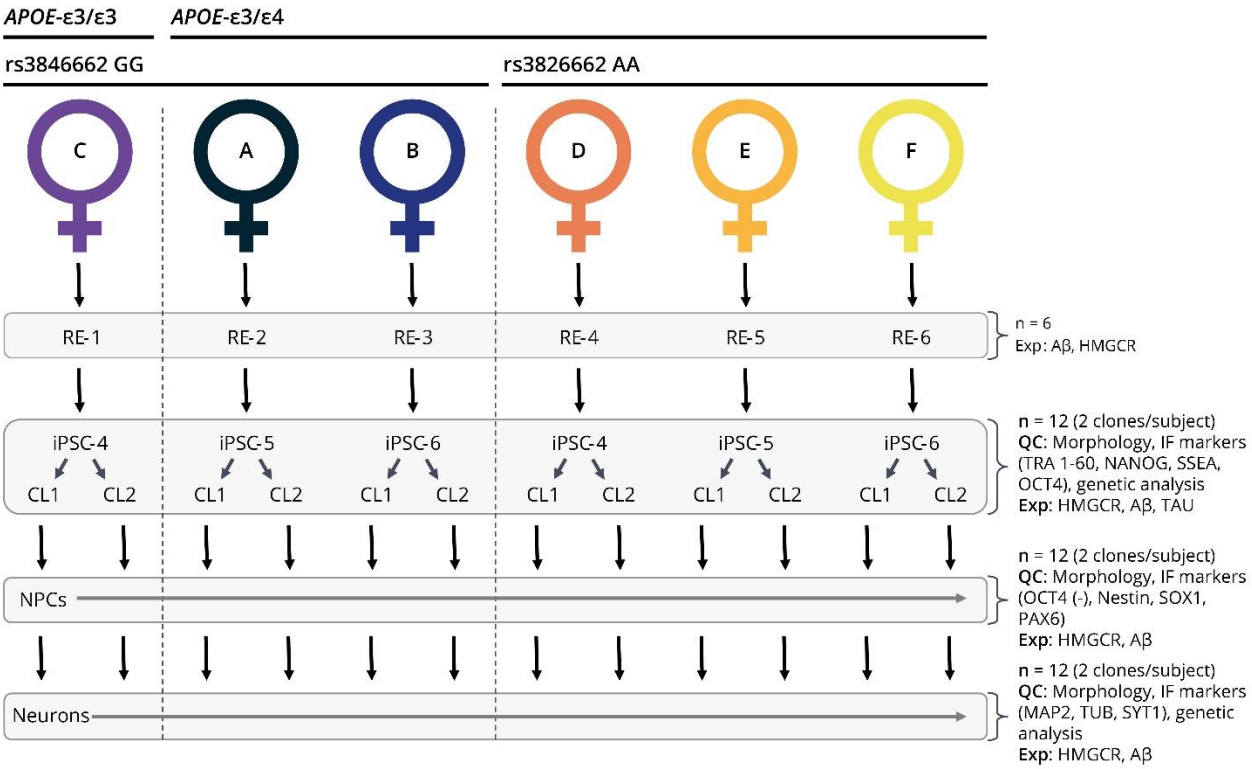
In the present study, we aimed to investigate the effect of rs3846662 on HMGCR and AD biomarkers across multiple cell types. We used stem cell technology to obtain induced pluripotent stem cells (iPSCs) from renal epithelial cells (RECs) which were then further differentiated into neural progenitor cells (NPCs) and neurons. The effect of rs3846662 was assessed at the different cell levels.

3.2 Material and methods

3.2.1 Experimental plan

In order to investigate the effects of rs3846662 on HMGCR and AD biomarkers on a cellular level and across multiple cell types, we decided to use stem cell technology (**Figure 3.1**). Individuals from the PREVENT-AD cohort were identified based on their rs3846662 genotype and asked to donate urine. From the urine, RECs were isolated and grown. These were used for creating iPSCs that

Figure 3.1 Experimental design



Six individuals were selected based on their rs3846662 genotype and RECs were isolated for everyone from collected urine. They were further reprogrammed to iPSCs and at least two clones/individual were expanded and differentiated into NPCs. Successful NPC cultures were differentiated into neurons. At each cell stage, QC measures and/or experiments were carried out as indicated.
Abbreviations: CL, clone; Exp, planned experiments; iPSC, induced pluripotent stem cell; NPC, neural progenitor cell, QC, quality control; RE, renal epithelial cell.

subsequently were differentiated into NPCs and neurons. At each cell type level, the levels of HMGCR and/or AD biomarkers was assessed.

3.2.2 Cohort description and individual selection (PREVENT-AD)

Individuals were recruited from the Pre-symptomatic Evaluation of Novel or Experimental Treatments for Alzheimer's Disease (PREVENT-AD) cohort [228]. This cohort consists of older, healthy individuals (55+) with a parental or multiple-sibling history of AD, followed over multiple years. All procedures were approved by the McGill University Faculty of Medicine Institutional

Review Board and complied with ethical principles of the Declaration of Helsinki. From the subset of individuals with genetic data, females either with the rs3846662 GG genotype (n = 3) or with the AA genotype (n = 3) were selected (**Table 3.1**).

3.2.3 Genetic data

Genotypes were determined with the Illumina Infinium Omni2.5M-8 array (Illumina, San Diege, CA, USA). The PLINK tool set (www.cog-genomics.org/plink2/) [232,233] was used to: 1) filter gender mismatches, 2) filter missingness at both the sample-level (< 5%) and SNP-level (< 5%), 3) assess sample heterozygosity and 4) filter SNPs in Hardy-Weinberg disequilibrium (p>0.001).

3.2.4 Isolation of RECs from urine

The selected individuals were asked to come in and their urine was collected on site. RECs were isolated according to Zhou et al., [267]. Briefly, individuals were asked to drink ample of water before the collection and to clean the urethral area with intimate wipes before collection. If possible, they were asked to not collect first micturition of the day and to collect urine mid-stream. The urine was collected in one or two sterile 80 ml specimen containers (NCS902-1O, New Century Scientific). From here on out, cells were handled using sterile techniques. The urine was transferred to 50 ml tubes (cat# 62.547.205, Sarstedt) and centrifuged at 400 g for 10 min. Cells were resuspended and washed in 10 ml wash buffer (**Supplemental sTable 3.1** in Appendix 3) and transferred to 15 ml tubes (cat# 62.554.205, Sarstedt). Samples were centrifuged at 200 g for 10 min, supernatant aspirated and cell pellet resuspended in 1 ml RE primary medium (**Supplemental sTable 3.1** in Appendix 3). Cells were then plated onto a gelatin coated (Gelatin, 0.1% (W/V) solution, cat# ES-006-B, Millipore) 12-well plate (cat# 3513, Corning); 1 well for each 80 ml collected urine, and 1 ml RE primary medium added/well. Cells were then incubated at 37 °C. For the following 48 h hours, 1 ml of RE primary medium/well was added daily and at 72 h post-plating the medium was collected and centrifuged at 400 g for 10 min. During centrifugation the plate was inspected for attached cells. If no attachment, centrifuged pellet was resuspended in 1 ml RE primary medium and added back to the plate, else RE

primary medium was directly added to the plate. For the following 72 h, 1 ml RE primary medium was added per well daily. On day 8 the medium was aspirated and new RE primary medium added. The plate was inspected for attached cells. Daily half medium changes were then performed until cells reached confluency. Cells were washed with Gibco™ PBS (cat# 10010-023, ThermoFisher Scientific), incubated with 0.05% Trypsin -EDTA (cat# 25300-054, Gibco™ ThermoFisher Scientific) for 3-5 min at 37 °C. Reaction was stopped by adding 10% FBS-DMEM/F12 (**Supplemental sTable 3.1** in Appendix 3), cells were centrifuged at 300 g for 5 min. Cells were resuspended in RE primary medium and plated onto gelatin coated plates.

3.2.5 Induction of pluripotency and differentiation to neurons

REC to iPSC: To induce pluripotency, RECs were transfected with the Yamanaka factors (*OCT3/4*, *SOX2*, *KLF4* and *MYC*) [268] according to the protocol by Bell et al., [269]. RECs were plated onto Corning™ Matrigel™ hESC-Qualified Matrix (cat# 354277, Fisher Scientific) coated dishes with RE primary medium. Cells were grown for two days and then collected, centrifuged and resuspended in 10% FBS (Gibco™, cat# 12483-020 ThermoFisher Scientific) in DMEM (Gibco™, cat# 11995-065, ThermoFisher Scientific). 200 000 – 300 000 RECs were transfected with the Episomal iPSC Reprogramming kit (cat# RF202, Alstem) using the Neon® Transfection System (Invitrogen™, cat# MPK10025, ThermoFisher Scientific). Cells were then plated onto a Matrigel® coated 6-well plate in RE primary medium and incubated at 37 °C overnight (day 1). After 24 h, 2 ug/ml puromycin (cat# 73342, STEMCELL Technologies) were added to the medium and cells incubated at 37 °C for another two days. On day four, RE primary medium was aspirated and TeSR™-E7™ medium (cat# 05910, STEMCELL Technologies) was added. Daily changes with TeSR™-E7™ medium were performed until cells reached appropriate morphology and size (500 – 1000 μm in diameter), usually ~14 days after change to TeSR™-E7™ medium. Individual colonies believed to be clonal, were then isolated using a combination of ReLeSR™ (cat# 05872, STEMCELL Technologies) and a 22-gauge needle, and transferred to TeSR™-E8™ (cat# 05940, STEMCELL Technologies), Essential 8™ (Gibco™, cat# A1517001, ThermoFisher Scientific) or mTeSR™-1 (cat# 85850, STEMCELL Technologies)

media. After expansion, colonies were further isolated before expansion and freezing. For each individual, two or three clones were propagated for experiments.

iPSC to NPC: High-quality iPSCs allowed to expand to a confluency of 15-20% for at least two days post passaging were used for differentiation into NPCs. The cells were incubated with neural induction medium 1 (**Supplemental sTable 3.2** in Appendix 3) for seven days, changing medium every other day, and then switched to neural induction medium 2 (**Supplemental sTable 3.2** in Appendix 3) for five days. Cells were then passaged using Gentle Cell Dissociation Reagent (cat# 07174, STEMCELL Technologies) to suspension culture in neural expansion medium (**Supplemental sTable 3.2** in Appendix 3). After a few days in suspension, NPC spheres formed that were collected by filtering through a 40 μm sterile cell strainer (Fisherbrand™, cat# 22-363-547, Fisher Scientific) and subsequently plated onto Matrigel® coated plates in STEMdiff™ Neural Progenitor Medium (cat# 05833, STEMCELL Technologies). Cells were grown to confluency, and at next passage cells were dissociated using Gentle Cell Dissociation Reagent (cat# 07174, STEMCELL Technologies) and plated onto new dishes. The cycle of suspension and dissociation were repeated until satisfactory purity.

NPC to neuron: High-quality NPCs grown on Poly-D-lysine/laminin double coated plates were grown to 70-80% confluency in STEMdiff™ Neural Progenitor Medium (cat# 05833, STEMCELL Technologies). Differentiation were started by switching to neuronal differentiation medium (**Supplemental sTable 3.2** in Appendix 3) with half medium changes every two – three days. Cells were collected after 30 days in differentiation medium.

3.2.6 Immunofluorescent staining & imaging

Cells were grown on appropriately coated circular glass coverslips (cat# 12CIR-1, ThermoFisher Scientific) placed in a suspension dish. The cells were fixed by washing the cells with PBS and then incubating with 4% PFA (made from Paraformaldehyde 8% aqueous solution (cat# 157-8, Electron Microscopy Sciences) by a 1:1 dilution with PBS) for 10 min at room temperature. Cells were then

washed with PBS, followed by permeabilization with 0.1% Triton X-100 (cat# 196145, GE Healthcare) for 10 min before blocking in 5% BSA (Gibco™, cat# 15260-037, ThermoFisher Scientific). Primary antibodies were added in indicated concentrations (**Supplemental sTable 3.3** in Appendix 3) over night at 4 °C. After overnight incubation, cells were washed with PBS for 5, 10 and 15 min before adding secondary antibody solution (**Supplemental sTable 3.3** in Appendix 3) and incubating for 1 h at room temperature (in the dark). Cells were washed 5, 10 and 15 min with PBS before mounting with Prolong Gold antifade reagent with DAPI (Invitrogen™, cat# P36931, ThermoFisher Scientific) onto microscope slides (Microscope Slides, Diamond White Glass, 25 x 75mm, Charged, 90° Ground Edges, cat# 1358, Globo Scientific). Imaging were done with a Zeiss Imager M2 with an ApoTome 2 fluorescence imaging system and analyzed with Fiji [270,271].

3.2.7 Genetic integrity analysis

Each iPSC cell line was assessed for genetic abnormalities. Briefly, DNA was extracted from iPSC cell pellets with the Genomic DNA Purification Kit (cat# 79020, STEMCELL Technologies). DNA quality was assessed by measuring the optical density using the Agilent Synergy H1 microplate reader. Finally, the DNA was analyzed with the hPSC Genetic Analysis Kit (cat# 07550, STEMCELL Technologies) and a subset of cell lines were also genotyped on an Illumina microarray (Illumina: CytoSNP-850K v1.2 BeadChip, WiCell Research Institute, Inc.).

3.2.8 Collection of cells for experiments

Each cell line was grown in duplicates or triplicates with respective growing conditions. Once reaching confluency, medium was collected in Eppendorf tubes and frozen at -80 °C. Cells were collected by washing three times with ice-cold PBS, then collected in PBS using a cell scraper (cat# 83.1832, Sarstedt). Samples were centrifuged at 14 000 rpm for 5 min. Cell pellet was washed once in PBS and re-centrifuged after which the supernatant was aspirated, and cell pellet frozen at -80°C until used for analyses.

3.2.9 qRT-PCR expression levels of *HMGCR*

RNA was extracted from REC and iPSC cell pellets using the Maxwell® 16 Cell LEV Total RNA Purification Kit (cat# AS1225, Promega) and optical density was measured with the Agilent Synergy H1 microplate reader. RNA was converted into cDNA using SuperScript™ VILO™ Master Mix (cat# 11755500, ThermoFisher Scientific). TaqMan™ Fast Advanced Master Mix (Applied Biosystems™, cat# 4444556, ThermoFisher Scientific) and QuantStudioTM 12K Flex Real Time PCR System and its accompanying software (v1.2.2) were used to quantify the cDNA. Human *HPRT1* was used as endogenous control (Applied Biosystems™, cat# 4326321E, ThermoFisher Scientific). Total, FL- and Δ13-*HMGCR* transcripts were measured with TaqMan Gene Expression Assays as described in [258].

3.2.10 ELISA for HMGCR protein levels, extracellular Aβ-42 and Aβ-40, and TAU

HMGCR, extracellular Aβ-42 and Aβ-40, and intracellular levels of TAU proteins were assessed in duplicates by ELISA. Briefly, cells and media were obtained according to the cell collection protocol described above. Collected cell pellets were homogenized and protein extracted by resuspending in 250 µl extraction buffer (PMSF [1mM], cocktail inhibitor [1x] (cat# 04693116001, Roche), in PBS (cat# 10010-023, Life Technologies)) and repeating three freeze/thaw cycles (-20 °C/room temperature), mixing gently after each thaw. Samples were then centrifuged for 15 min at 3000 rpm at 4 °C. Supernatant were transferred to a new tube and frozen at -80 °C until ELISA analyses. Total protein concentrations were assessed with Pierce™ BCA Protein Assay Kit (cat# 23225, ThermoFisher Scientific).

HMGCR protein levels were assessed with the Human 3-Hydroxy-3-methylglutaryl CoA reductase ELISA Kit (cat# MBS742031, MyBioSource). Upon thawing, cells were homogenized (see above) and then diluted 1:4 in extraction buffer before subjected to the protocol supplied by the company.

For assessment of AD biomarkers ELISA kits from Fujirebio were used. Extracellular Aβ-42 levels were assessed using the Aβ(1-42) HS Conj kit (cat# 81587) in combination with INNOTEST® β-

AMYLOID(1-42) ELISA kit (cat# 81576) and Aβ(1-42) CAL-RVC pack (cat# 81584). Extracellular Aβ-40 levels were assessed using the INNOTEST® β-AMYLOID(1-40) kit (cat# 81585). At day of analysis, media samples were thawed, centrifuged for 10 min at 3000 rpm at 4 °C. 75 (Aβ-40) or 100 µl (Aβ-42) of sample were used for analysis in protocols supplied by the company. Intracellular TAU was assessed by INNOTEST® hTAU Ag kit (cat# 81579). Cell pellet samples were prepared by homogenizing (see above) and diluting 1:10 in sample diluent (PBS, included in the kit) before subjected to protocol supplied by company.

3.2.11 LC-MS/MS for HMGCR activity

Cell growth conditions: iPSCs were plated in triplicates on Matrigel® in mTeSR-1™ medium. Media was changed after 24 h and then left to expand for 48 h. Cells were collected by scraping as described above.

Protein extraction: Cell pellets were lysed by adding 75 µl ice-cold lysis buffer and then subjecting samples to three freeze/thaw cycles (10 min at -80 °C/10 min at 37 °C), vortexing before each cycle. After lysis, samples were vortexed and centrifuged for 10 min at 14 000 rpm at 4 °C. The supernatant was collected, and protein concentrations were determined using Pierce™ BCA Protein Assay Kit (cat# 23225, ThermoFisher Scientific).

HMGCR activity assay: Enzymatic activity assay was performed according to Honda et al., [272] with modifications. To activate HMGCR, 25 µg of extracted protein was added to the enzyme buffer for a total volume of 120 µl, and incubated for 10 min at 37 °C. To assay HMGCR activity, 30 µl 3-hydroxy-3-methylglutaryl-CoA [2 mM] (HMG-CoA, cat# H6132, Sigma-Aldrich) was added to the mix and samples were incubated for 90 min at 37 °C. Reaction was stopped by adding 20 µl of HCl [6 M]. Samples incubated for 12-14 h at room temperature. To precipitate the proteins, samples were incubated over night at -20 °C with 300 µl ethanol containing 125 ng internal standard MVAL-D7 (cat# D-3743, CDN Isotopes). After centrifugation for 10 min at 14 000 rpm at 4 °C, supernatants were transferred to glass tubes (12*75 mm) and evaporated using a RapidVap Vertex Dry Evaporator

(Labconco). Precipitates were then diluted in 50 µl EtOAc/MeOH/H_2O (1:1:1 v/v/v) and transferred to HPLC vials.

LC-MS/MS analysis: Levels of mevalonolactone (MVAL) were measured as a proxy for HMGCR activity. The LC-MS/MS analysis was based on the method previously described [273]. An Agilent 1200 HPLC system coupled to a 3200 QTRAP® mass spectrometer with an API electrospray ion source was used. Analyst software (v1.4.2, AB Applied Biosystems) were used to set up protocol and analyze the data. For separation of MVAL, 10 µl of each sample was injected onto a Spherisorb ODS2 Column (80Å, 3 µm, 4.6 mm*100 mm, Waters™). A gradient program was used at a flow rate of 0.2 ml/min. It was adjusted to 93%/7% water/methanol containing 0.1% formic acid and 2 mM ammonium acetate for 2.5 min, followed by a linear gradient of 98% methanol containing 0.1% formic acid and 2 mM ammonium acetate for 5.5 min. The mobile phase was then readjusted to 93%/7% water/methanol for 17 min. The mass spectrometer, performed in electrospray positive ionization mode, with instrument settings as follows: CUR(10),CAD(0), NC(2), TEM(450), GS1(40), GS2(45), for MVAL and MVAL-D7 the DP(23), EP(10), CXP(5) and CE(15)for MVAL and CE(30) for MVAL-D7. Multiple reaction monitoring (MRM) mode was used for sample analysis, with mass transitions of *m/z* 131→69 and *m/z* 138.2→77 for MVAL and MVAL-D7 respectively.

3.2.12 Statistics

For each experiment, each clone was grown in duplicates or triplicates. Each replicate was further run in duplicates or triplicates in the specific assays. For statistical analyses, we obtained one data point for each clone by averaging the values of all replicates. All results were analyzed in R [248]. For each assay, if more than one cell type was analyzed, a two-factor ANOVA, examining the effect of cell type (REC, iPSC, NPC and neuron) and rs3846662 genotype (AA, GG) was used. Post hoc pairwise comparisons were evaluated with the "emmeans" package [254] using the Tukey method. Welch two sample t-tests were used if only one cell type was analyzed (HMGCR activity and TAU protein levels)

as well as for the descriptives. Figures were produced with the "ggplot2" package [250]. Values of groups are reported as mean ± standard error of the mean (SE).

3.3 Results

3.3.1 Cohort descriptives

Details for the six selected individuals are shown in **Table 3.1**. Findings from Leduc et al., [13] suggested a protective effect of rs3846662 AA genotype primarily in female *APOE*-ε4 carriers. Thus, we decided to draw a sample (n = 6) from females only, that predominantly were *APOE*-ε4 positive (five with ε3/ε4 genotype, one with ε3/ε3 genotype) and with either rs3846662 GG (n = 3) or AA (n = 3) genotypes. These individuals did not differ in terms of age ($t(2.29) = -1.78$, $p = 0.20$), BMI ($t(2.22) = 0.43$, $p = 0.71$) or weight ($t(2.41) = 0.61$, $p = 0.60$). One rs3846662 AA individual was taking statins at the time of enrollment in the main study.

3.3.2 Rs3846662 associates with HMGCR

***HMGCR* RNA expression levels**: The effect of rs3846662 on RNA levels of FL-*HMGCR* and Δ13-*HMGCR* (ratios over total *HMGCR* levels), as well as its interaction with cell type, was assessed in RECs and iPSCs (**Figure 3.2 A**). We found significant main effects of cell type ($F(1, 13) = 6.66$, $p =$

Table 3.1 Descriptives of PREVENT-AD subjects

	Sex	APOE-ε4 genotype	Statin use @ eligibility visit	Age [years] @ urine collection	Weight [kg]	BMI [kg/m2]
rs3846662 GG				**68.3#**	**64.5#**	**24.0#**
Subject A	F	ε3/ε4	none	73.7	67.6	24.1
Subject B	F	ε3/ε4	none	65.8	58.9	22.6
Subject C	F	ε3/ε3	none	65.3	67.1	25.4
rs3846662 AA				**63.3#**	**70.1#**	**25.5#**
Subject D	F	ε3/ε4	Teva-Rosuvastatin	63.7	86.9	31.9
Subject E	F	ε3/ε4	none	61.9	57.7	20.5
Subject F	F	ε3/ε4	none	64.3	65.7	24.1

Indicate group means, no statistical difference ($p > 0.05$) between GG and AA genotypes as assessed by Welch t-test

0.02) and rs3846662 ($F(1, 13) = 5.26$, $p = 0.04$) on FL-*HMGCR*, but no interaction effect. Levels of FL-*HMGCR* were increased in RECs compared to iPSCs (ratios 1.08 ± 0.04 *vs* 0.94 ± 0.03) and decreased in rs3846662 AA carriers compared to GG carriers (ratios 0.95 ± 0.03 *vs* 1.07 ± 0.04).

For Δ13-*HMGCR*, there was a significant main effect of rs3846662 ($F(1, 13) = 27.96$, $p = 0.0001$) such that levels of Δ13-HMGCR were increased in AA carriers compared to GG carriers (ratios 1.50 ± 0.10 *vs* 0.69 ± 0.12). No significant interaction or main effect of cell type was found on Δ13-*HMGCR* levels ($Fs(1, 13) \leq 1.48$, $ps \geq 0.25$).

HMGCR protein levels: HMGCR protein levels were assessed in RECs, iPSCs, NPCs and neurons with regards to cell type and rs3846662 (**Figure 3.2 B**). We found significant main effects for both factors ($F(3, 21) = 3.84$, $p = 0.02$ and $F(1, 21) = 7.35$, $p = 0.01$, respectively) but no interaction effect. Protein levels were increased in AA carriers compared to GG carriers (ratios 2.73 ± 0.48 *vs* 1.15 ± 0.19). Post hoc analysis of cell type revealed lower levels in neurons compared to RECs ($p = 0.031$, ratios 0.24 ± 0.02 *vs* 2.88 ± 0.41) and a trend for reduction compared to iPSCs ($p = 0.082$, ratio 2.06 ± 0.47). NPCs did not differ from any of the cell types (ratio 1.34 ± 0.27).

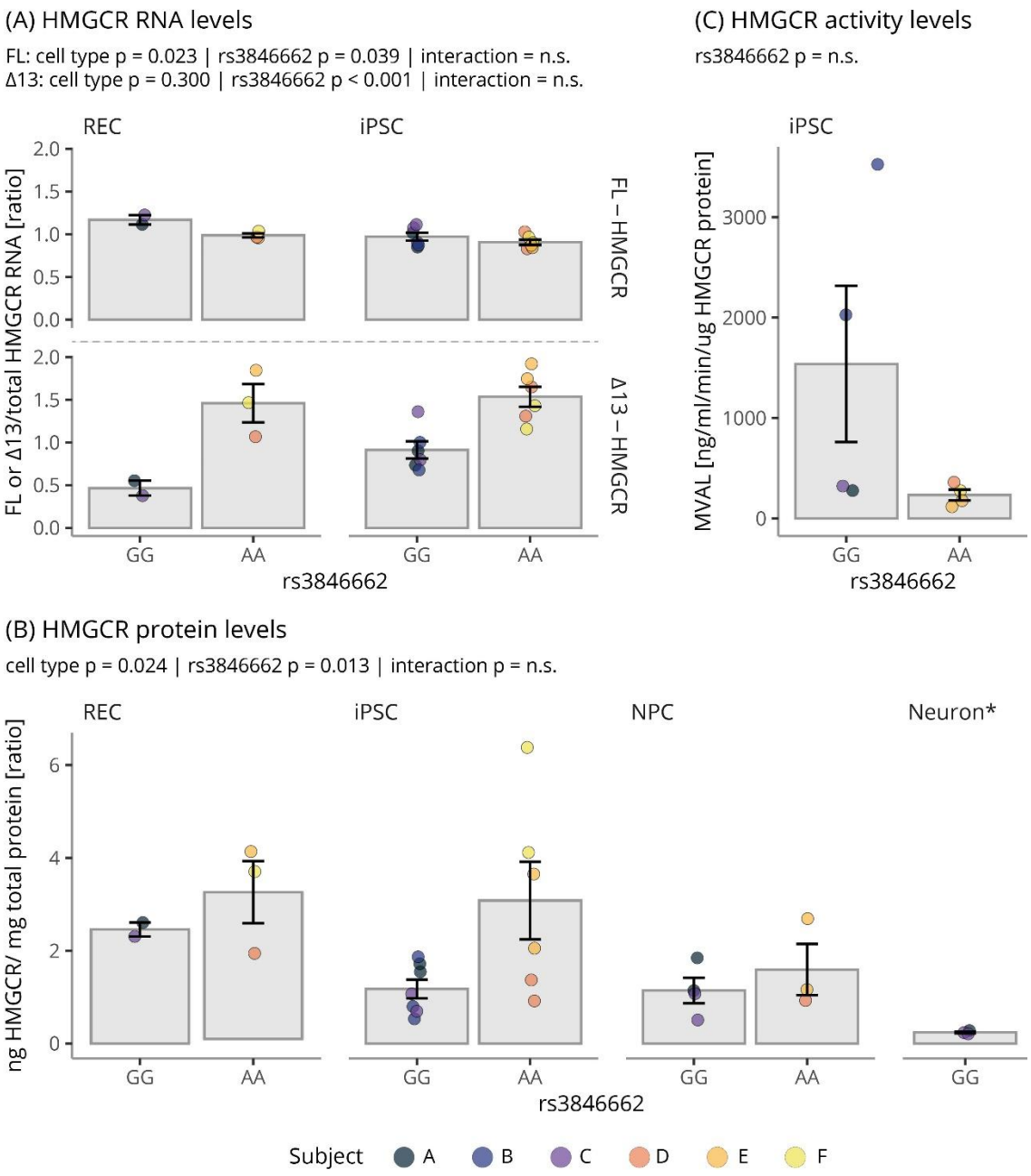
HMGCR activity levels: HMGCR activity levels were assessed using an adapted LC-MS/MS protocol in iPSCs (**Figure 3.2 C**). Although we did see a decrease in HMGCR activity in AA carriers compared to GG carriers (232.3 ± 53.7 *vs* 1537.8 ± 777.8 ng/ml/min/µg HMGCR protein), this was not significant ($t(3.8) = -0.86$, $p = 0.44$).

3.3.3 Rs3846662 has no detectable effect on AD biomarkers

Extracellular levels of Aβ-40 and Aβ-42, and intracellular levels of TAU were assessed by ELISA (**Figure 3.3**).

Aβ-40 levels: Extracellular Aβ-40 levels were assessed in RECs, iPSCs, NPCs, and in neurons (**Figure 3.3 A**). No interaction effect between cell type and rs3846662 could be detected, as no main effect of rs3846662 ($Fs \leq 0.003$, $ps = 1.00$). There was a significant effect of cell type ($F(3, 20) = 13.40$,

Figure 3.2 rs3846662 effect on HMGCR



Differences in HMGCR RNA levels (A), protein levels (B), and activity (C) dependent on rs3846662 genotype and cell type is depicted. In (A), ratios of FL- and Δ13-*HMGCR* RNA levels over total *HMGCR* levels in RECs and iPSCs were assessed by qRT-PCR. In (B), ratio of HMGCR protein over total protein levels in RECs, iPSCs, NPCs and neurons were assessed by ELISA and revealed an effect of cell type but post hoc analysis underpowered to detect which groups differed from each other. In (C), levels of MVAL as a proxy for HMGCR activity were assessed by LC-MS/MS. Error bars represent SE. For statistical details see 3.2.12 Statistics and 3.3 Results.

Figure 3.2 cont.
Abbreviations: Δ13-HMGCR, HMGCR transcript lacking exon 13; FL-HMGCR, full length HMGCR; iPSC, induced pluripotent stem cell; MVAL, mevalonolactone; NPC, neural progenitor cell; REC, renal epithelial cell; SE, standard error of the mean.
n.s. = non significant ($p > 0.05$)
* $p \leq 0.05$ Tukey corrected post-hoc analysis, compared to RECs

$p = 5*10^{-5}$) and post hoc analysis revealed higher levels in neurons compared to the other cell types (ps ≤ 0.0004, 741.77 ± 332.97 pg/ml) that further did not differ between each other (7.50 ± 1.21, 6.48 ± 1.09, and 97.44 ± 22.43 pg/ml for RECs, iPSCs, and NPCs respectively, ps ≥ 0.732).

Aβ-42 levels: Levels of secreted Aβ-42 were assessed in iPSCs, NPCs and neurons (**Figure 3.3 B**). The cell type * rs3846662 interaction as well as the main effect of rs3846662 were not significant (Fs ≤ 0.001, ps ≥ 0.97) whereas the cell type main effect was ($F(2, 9) = 5.22$, $p = 0.031$). Post hoc analysis revealed no difference between iPSCs and NPCs ($p = 0.99$, 12.20 ± 0.38 *vs* 19.45 ± 2.51 pg/ml) but higher levels in neurons compared to both other cell types (ps ≤ 0.049, 253.92 ± 137.85 pg/ml).

Aβ-42/Aβ-40 ratio: The ratio of secreted Aβ-42/Aβ-40 was calculated for each clone and was assessed in iPSCs, NPCs and neurons (**Figure 3.3 C**). We found a significant effect of cell type ($F(2, 9) = 11.8$, $p = 0.003$) with iPSCs having significantly higher ratios than NPCs and neurons (ps ≤ 0.015, ratios 2.26 ± 0.65 *vs* 0.23 ± 0.024 and 0.36 ± 0.064, respectively). We could not find any significant effects for either of the rs3846662 terms (Fs(1, 9) ≤ 0.92, ps ≥ 0.36).

TAU levels: Intracellular TAU levels were measured in iPSCs (**Figure 3.3 C**). TAU was expressed at a sufficient level (428.9 ± 5.8 *vs* 439.3 ± 11.6 pg/ml in AA and GG respectively) but we found no significant effect of rs3846662 ($t(8.7) = -0.81$, $p = 0.44$).

3.3.4 QC results

Expression of cell type specific markers: iPSCs, NPCs and neurons were stained for their respective markers (**Figure 3.4**). Overall, cell lines stained positive for iPSC markers octamer-binding transcription factor 4 (OCT4), Nanog, Stage-Specific Embryonic Antigen 4 (SSEA4), and TRA-1-60 (**Figure 3.4 A**). The NPC cultures showed more variability. Some cell lines (e.g. from

Figure 3.3 rs3846662 effect on AD biomarkers

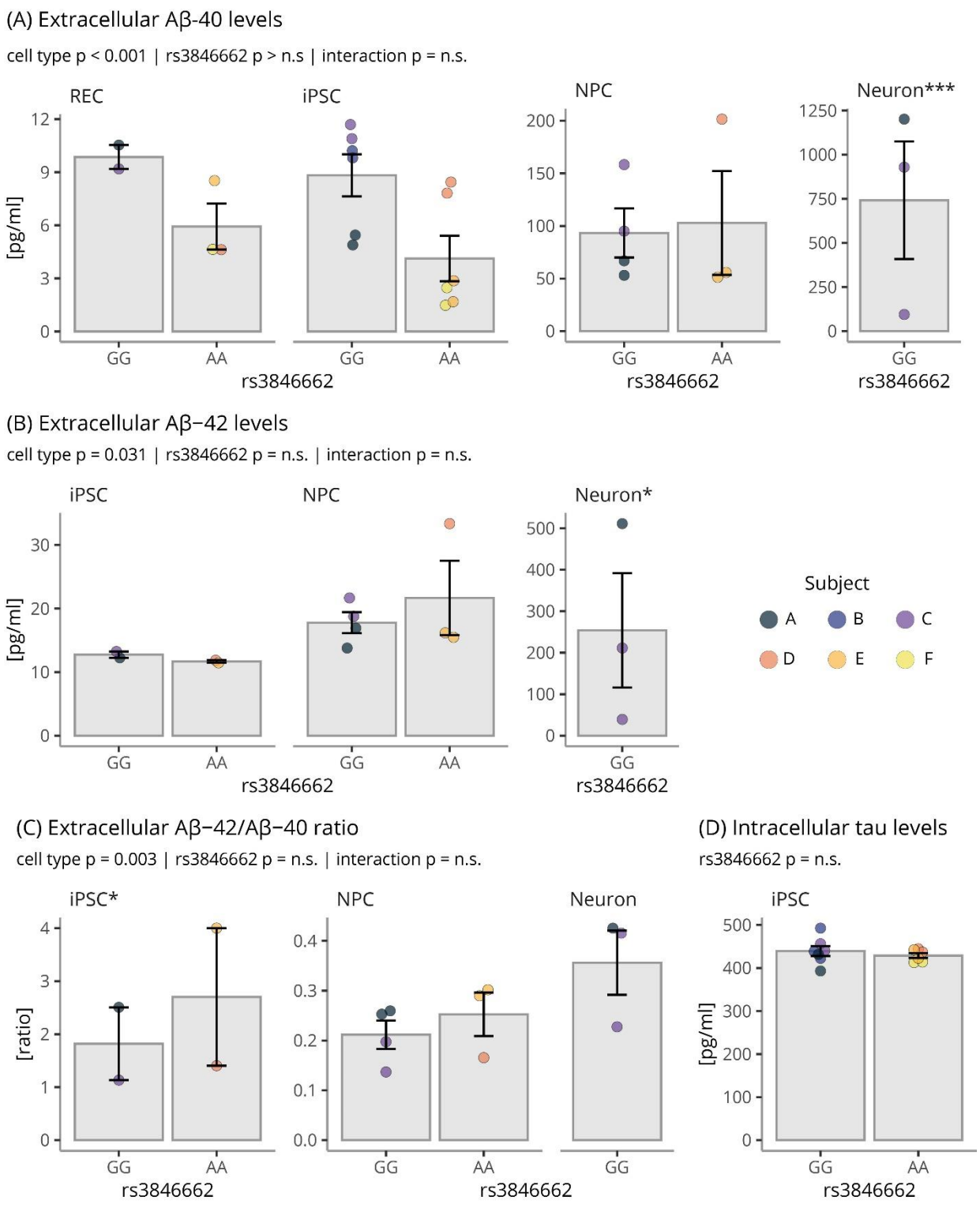
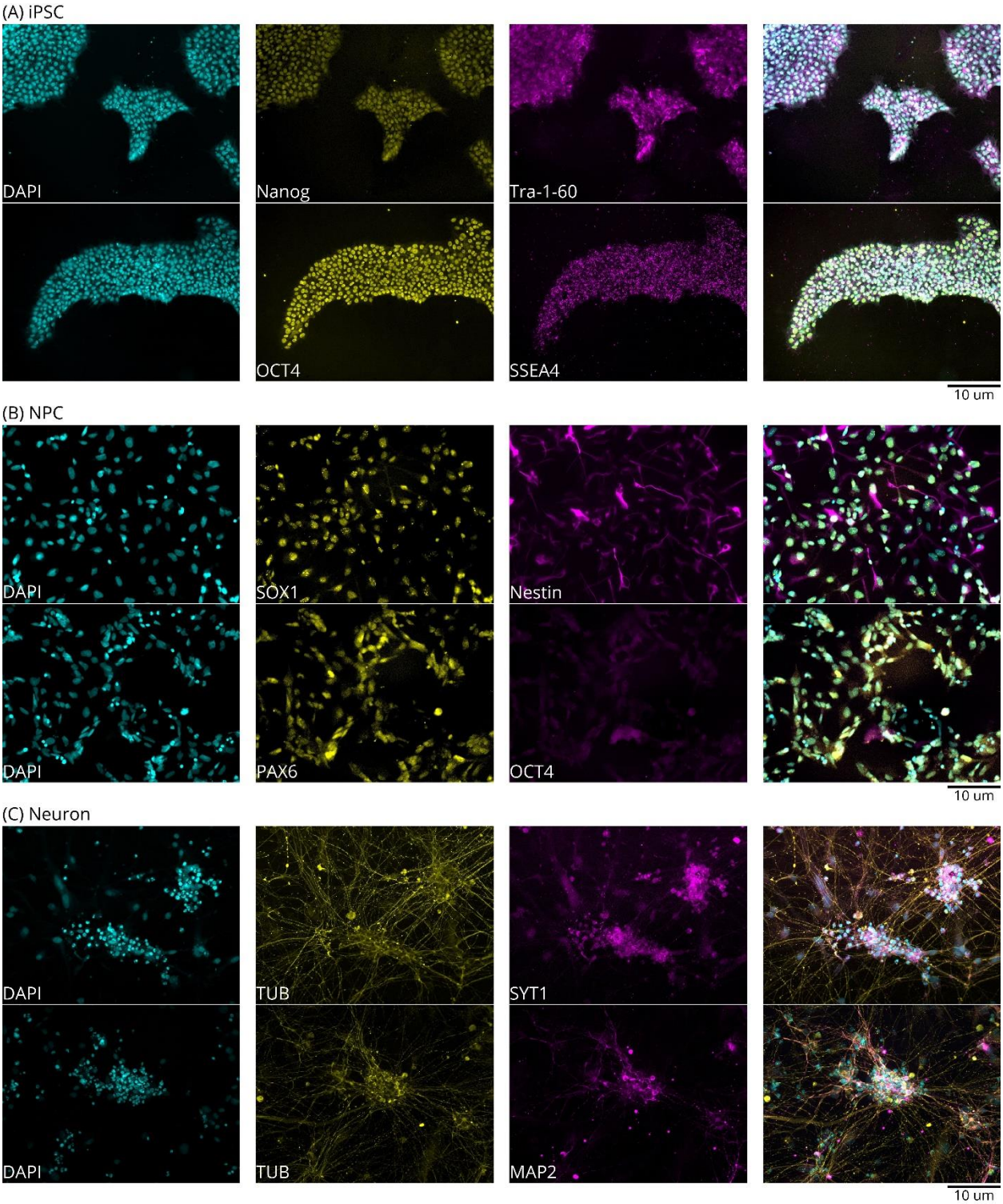


Figure 3.3 cont.
Levels of AD biomarkers were measured by ELISA; extracellular Aβ-40 (A), Aβ-42 (B), and the Aβ-42/Aβ-40 ratio, as well as intracellular TAU (D). Error bars represent SE. For statistical details see 3.2.12 Statistics and 3.3 Results.
Abbreviations: iPSC, induced pluripotent stem cell; NPC, neural progenitor cell; REC renal epithelial cell; SE, standard error of the mean.
*/*** $p \leq 0.05/p \leq 0.001$ Tukey corrected post-hoc analysis, compared to all other cell types

individual F) displayed uneven paired box 6 (PAX6), SRY-box transcription factor 1 (SOX1), and Nestin immunoreactivity, while some cell lines (e.g. from individual B) still expressed the iPSC marker OCT4. Not all cell lines were stained for NPC markers because initial evaluation based on morphology already suggested insufficient purity (**Figure 3.4 B**). Even less cell lines were stained for neurons for the same reason. However, we did manage to produce neurons from some cell lines and staining of these revealed expression of the neuronal markers β-III tubulin (TUB), synaptotagmin (SYT1) and microtubule associated protein 2 (MAP2, **Figure 3.4 C**).

Genetic analysis: Genetic analysis was performed on iPSCs from each clone with the hPSC Genetic Analysis Kit from STEMCELL Technologies (**Figure 3.5**). This kit investigates nine of the most common genetic abnormalities occurring as a consequence of inducing pluripotency by reprogramming. Out of 14 developed cell lines, we initially found all but one (B CL3) to carry a duplication of the 17q region. Furthermore, cell line A CL2 showed deletion of the Xp region, B CL2 a duplication of the 12p region, and D CL2 had duplications of both 12p and 20q regions. To confirm the accuracy of the 17q results, which seemed highly unlikely, four cell lines were sent to WiCell (WiCell Research Institute, Inc.) for characterization by SNP microarray. We were able to confirm the deletion of an X region in cell line A CL2, but also found duplications of the X chromosome in cell lines B CL2 and C CL1, not reported with the STEMCELL kit. Furthermore, the duplication of the 12p region in cell line B CL2 was not replicated. As for the 17q region, the results are contradictory. Although none of the four cell lines characterized by microarray, had any reportable gains of this region, the raw data indicated duplications of small regions (< 150 kb) for three of the cell lines (**Supplemental sTable 3.4** in Appendix 3).

Figure 3.4 Quality control expression of cell type specific markers



Expression of cell type specific markers was performed with immunofluorescent stainings; iPSCs expressing Nanog, TRA-1-60, OCT4 and SSEA4 (A), NPCs expressing SOX1, Nestin and PAX6 with

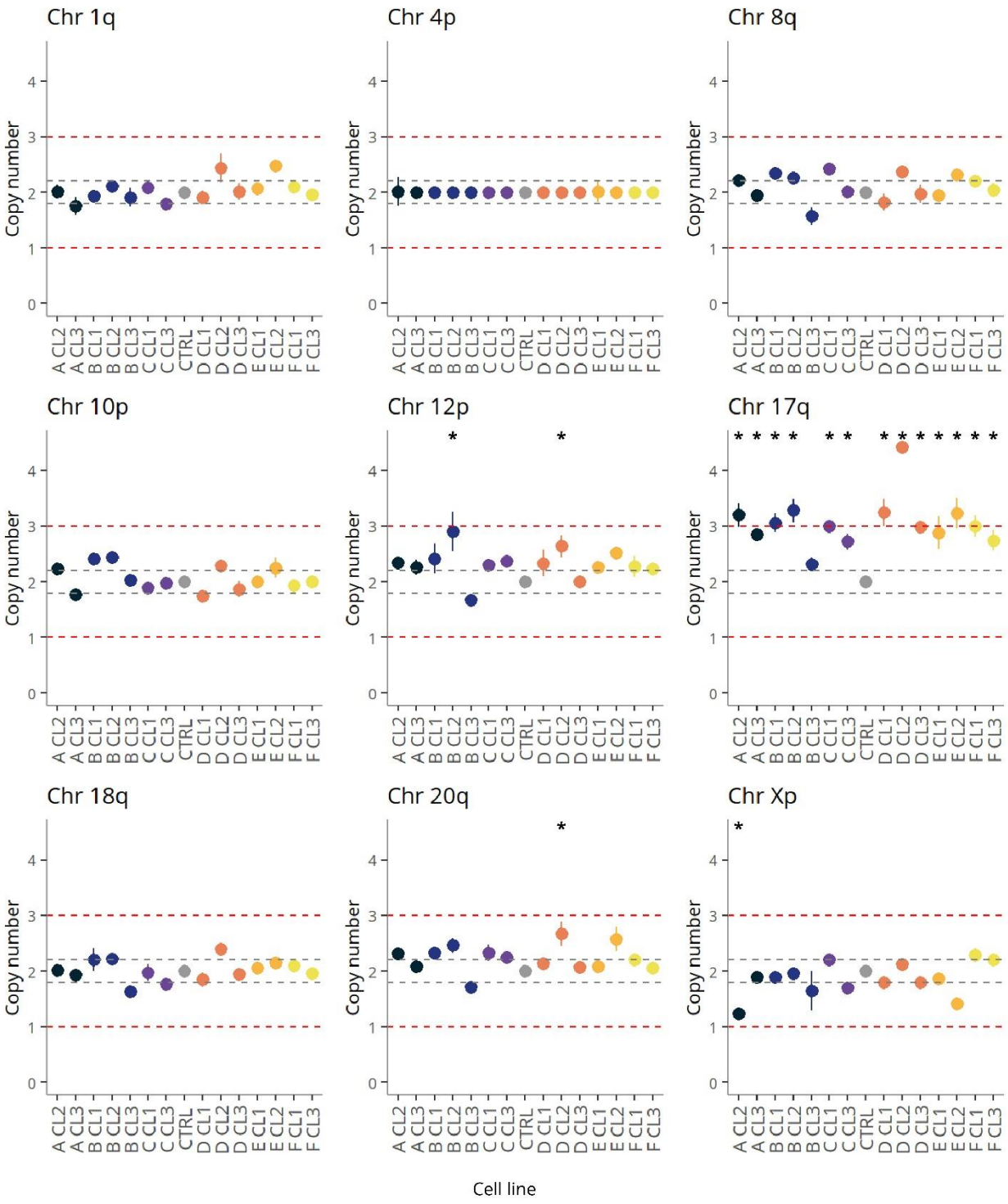
Figure 3.4 cont.
the absence of nuclear OCT4 expression (compare to OCT4 staining in (A)) (B), and neurons expressing TUB, SYT1 and MAP2 (C).
Abbreviations: iPSC, induced pluripotent stem cell; NPC, neural progenitor cell.

3.4 Conclusion

Comments on results: In this study we aimed to determine the cellular effects of rs3846662 on HMGCR and AD biomarkers by creating stem cells from individuals either carrying the AA or GG genotypes. In line with previous studies, we show that rs3846662 associates with measures of HMGCR; cells from rs3846662 AA carriers have higher Δ13-*HMGCR* RNA levels, accompanied by higher levels of HMGCR protein. Following our hypothesis that Δ13-*HMGCR* transcripts associates with lower activity, these findings go well in hand with what is previously known about HMGCR regulation; inhibition of HMGCR by statins leads to a dramatic upregulation HMGCR protein [165,167]. While not significant, the data indicate that although protein levels are higher, activity levels are lower.

While our findings on HMGCR are promising, we could not detect any effect of rs3846662 on extracellular levels of Aβ or intracellular levels of TAU. The lack of association between rs3846662 AA genotype and levels of AD biomarkers Aβ and TAU could be due to a number of reasons. The most obvious reason could be that we are critically underpowered to detect such a difference with a sample size of three individuals per genotype (at best). In line with this, differences in growth rate and potential to differentiate between the cell lines further affected the sample sizes. This in turn led to that we could not investigate any effect in certain cell types (primarily neurons, where we could only obtain data from GG carriers). Of course, one could hypothesize that rs3846662 directly affects growth rate due to its effect on HMGCR and cholesterol. Cholesterol is necessary for cells to grow and specifically for neurons to sprout dendrites and axons. Thus, although we could not prove that rs3846662 cells have less HMGCR activity in iPSCs, it is possible that this could have had an effect in other cell types, specifically neurons that are grown for 30 days without passaging. But again, we are underpowered to draw any conclusions regarding this hypothesis.

Figure 3.5 Genetic analysis of iPSC cell lines



Genetic integrity across 12 chromosomal loci assessed with the hPSC Genetic Analysis Kit (STEMCELL Technologies).

Figure 3.5 cont.
Abbreviations: Chr, chromosome; CL, clone.

Another aspect that needs to be taken into account is the rejuvenation that occurs when reprogramming somatic cells to iPSCs with the Yamanaka factors [268,274]. Aging is the most important risk factor for AD [119] and it is possible that by using trans-differentiation techniques (e.g. [275]) that allows the "age" of the cells to be retained, we would be better positioned to detect any changes in AD biomarkers.

Contradictory literature: The evidence of rs3846662 on AD is contradicting to begin with. There are definitely several papers supporting a protective association between rs3846662 A allele and AD, but there is also some supporting the opposite; Cao et al., found significant associations between the G allele and reduced hippocampal atrophy and increased temporal glucose metabolism [276]. Similarly, there has been suggestions for interactions with *APOE*-ε4 status and gender, but with mixed results; Leduc et al., [13] found an effect primarily in female *APOE*-ε4 carriers, while Chang et al., [261] found an effect in *APOE*-ε4 non-carriers. Wright et al., [34] who specifically tried to replicate the findings of Leduc et al., did find a stronger effect in *APOE*-ε4 carriers but could not replicate the sex effect. The source of these discrepancies could potentially be differences in ethnicity of the cohorts, some being performed exclusively in Europeans or Han Chinese, or in mixed cohorts. Rs3846662 differ greatly with regards to allele frequency between different ethnicities, in fact the A allele is the major allele in Europeans while it is the minor allele in most other ethnicities according to Ensembl (useast.ensembl.org/index.html). Of note, similar ethnic discrepancies have been reported in other GWAS risk factors such as CLU and PICALM [277–279]. In this study, three of the individuals were deemed to be of European ethnicity (data not shown), one was non-European, and two were not assessed. Nevertheless, most published studies support a protective role of rs3846662 A allele in AD.

Comments on quality control: Quality control of the cell lines did provide some answers to the differentiation problems. In the case of individual B, clones were repeatedly differentiated into NPCs

but with resulting cultures being impure. Visually inspecting the morphology and growth pattern at the end of the NPC induction protocol revealed iPSC like clusters of cells in the NPC culture. Upon staining it was indeed confirmed that these cells still expressed the iPSC marker OCT4 (**Supplemental sFigure 3.1** in Appendix 3). Further differentiation to neurons was possible, but due to the high level of impurity, only sparse neuronal cells was present and could only be detected upon staining. These cultures were thus not included in the neuronal analysis. Interestingly, two out of the three clones developed for this individual did display genetic abnormalities (**Figure 3.5**).

Other cell lines that proved problematic was those derived from individual F. In this case, CL1 was fully genotyped with SNP microarray and indicated no reportable gains or losses and IF staining revealed, although somewhat faint, expression of NPC markers in the absence of any positive OCT4 staining. Nevertheless, this cell line never resulted in any sufficient amount of neurons. Similarly, the cells were never grown for NPC experiments due to impurity of the NPC cell culture.

Overall, the problematic cell lines (mainly B and F) did seem to associate more with the individual rather than specific clones suggesting an issue with cell origin rather than issues occurring as a consequence of the reprogramming.

Another issue encountered during the quality control steps, were the issue of what appears to be a duplication of a chromosome 17q region in all but one cell line (B CL3, **Figure 3.5**). This seemed suspicious to us, and so we decided to send samples for genotyping to confirm. None of the replicated samples (n = 4) were reported to have any significant duplications of chromosome 17. We did try to run the analysis using our own control samples (from the same individuals but DNA extracted from blood) and found that the result varied depending on which control sample was used as reference (data not shown). Other chromosomes showed consistent results, the issue was only with chromosome 17.

The results from quality control steps taken in this study show that maintained genetic integrity and expression of cell type specific markers (as was the case for cell line F CL1) is not sufficient to ensure

successful differentiation of iPSCs to NPCs or neurons. It is also clear, that the assessment of genetic integrity after reprogramming differ widely depending on method used. The results from the STEMCELL Genetic Analysis kit and WiCell SNP microarray had very poor overlap. This strongly implores proper establishment of methods and use of multiple controls before interpreting any results pertaining to actual samples.

In summary: We have found further support for the role of rs3846662 in alternative splicing of *HMGCR* resulting in increased levels of the Δ13-*HMGCR* transcript potentially affecting the activity levels of the HMGCR protein. We could however not confirm any effect on AD biomarkers, but more studies with more appropriate samples sizes or isogenic cell lines are needed to determine if there is an effect of rs3846662 on AD biomarkers or not.

Chapter 4

A total cholesterol polygenic score in Alzheimer's disease

In Chapter 2 we identified rs72633963 to associate both with reduced blood cholesterol levels and protective AD phenotypes. Similarly, the literature suggests that rs3846662, examined in Chapter 3, associates with reduced risk of developing AD and with reduced blood cholesterol levels. Since midlife hypercholesterolemia is an established risk factor for AD, it begged the question whether these variants were influencing AD risk and phenotypes by decreasing peripheral blood cholesterol levels. Thus, in this chapter we wanted to test that hypothesis in a broader sense. We created a polygenic score with SNPs associating specifically with blood total cholesterol levels and evaluated it in the context of AD. This approach has indeed been used in a previous study [280] but showed no effect on AD. However, the polygenic score only captured a small portion of the variance in blood cholesterol levels (3.6%) and only clinical AD was evaluated. We thus hypothesized that optimizing the score by adjusting the selection process of SNPs to include, stratifying for factors such as sex and statin use, and considering AD as a biological construct, we could potentially unravel associations between distinct pathologies and at different stages of the disease (pre-clinical, clinical, and definite). In this study we show that the association between the score and cholesterol levels can be vastly improved by considering p-value threshold of SNPs to include in the score as well as statin use and sex in the cohort under investigation. We could however not find any associations with AD phenotypes or risk.

4.1 Introduction

The mechanism by which rs72633963 and rs3846662 exert their protective effect in AD remains to be determined. We hypothesized in Chapter 2 that the effect of rs72633963 was possibly mediated through its effect on peripheral TC and/or LDL-C levels. Interestingly, rs3846662 examined in Chapter 3, is also known to influence both TC and LDL-C levels [280]. Hypercholesterolemia or increased TC levels is a known risk factor for AD [1]; midlife hypercholesterolemia is associated with increased risk of developing AD [16–21,182,183] and high cholesterol (either in the form of TC or LDL-C) associates with increased amyloid load in the brain [184–186] as well as hypometabolism in brain regions affected by AD [189]. In line with these findings, rs72633963 and rs3846662 protective alleles associate with reduced TC levels [280], suggesting that they may be protective in AD by reducing TC levels.

Similar to AD, blood cholesterol levels are influenced by genetics [280–282]; heritability of AD is estimated to be 58%-79% [2] and 32%-63% for TC [283]. Considering the genetic background of both conditions, and the fact that they are linked in terms of risk, it is possible that some of the genetic variance seen in AD can be explained by variants influencing blood cholesterol levels.

An early study by Proitsi et al., [213] investigated the effect of a TC polygenic score (TC-PGS) in AD and did not reveal any significant effects. Two factors that could influence these results is the process of SNP selection for the score (reviewed in 1.2 Genetics in sporadic Alzheimer's disease) and the fact that AD is a clinicopathological construct (reviewed in 1.1.2 Definition). They used genome-wide significant SNPs to build the TC-PGS and while it did associate with TC levels, it only explained a mere 3.6 % of the variance which is in stark contrast to the heritability estimates. Further, only clinical AD was evaluated.

A clinical diagnosis of probable AD has a sensitivity of 81% and a specificity of 70% to predict definite AD [41], highlighting the fact that there are both a neuropathological and a clinical process. To overcome these discrepancies a new research framework was put forward by the NIA-AA [58] which

propose that, for research purposes only, AD should be defined as a biological construct determined by the presence of pathology as assessed with the AT(N) classification system [59]. Of note, this biological definition was proposed to also work with current clinical diagnoses of AD; e.g. AD neuropathological change with or without accompanying clinical AD diagnosis.

Considering the discrepancies of clinical and pathophysiologically proved AD and the fact that multiple studies have shown that incorporation of SNPs with p-values higher than the genome-wide significance threshold into a polygenic score can improve its performance, it is possible that the inclusion of weaker effect loci in the TC-PGS and using the new classification system for AD, could reveal important associations.

In this final study, we thus tested the hypothesis if the combined effect of SNPs associating with TC levels also influence AD risk or pathophysiological process. We derived an optimized TC-PGS by constructing multiple weighted TC-PGSs based on various p-value thresholds and investigated their interaction with sex and statin use; factors known to influence TC levels. We further aimed to characterize the optimized TC-PGS in relation to AD using three different cohorts spanning the spectrum of pre-clinical/pre-symptomatic, clinical, and pathophysiologically proved AD. To further minimize discrepancies between clinical and pathophysiologically proved AD, we used the AT(N) system to refine clinical diagnoses of AD, incorporating evidence of AD neuropathological change.

4.2 Material and methods

4.2.1 Meta-analysis summary data

Summary statistic data from the Global Lipids Genetics Consortium's meta-analysis of TC GWA studies [280] was available online and downloaded from [284]. Results from the joint analysis of metabochip and GWAS data was used. Before used for scoring, ambiguous SNPs were excluded and only SNPs present in all three target data sets were kept.

4.2.2 Study populations

PREVENT-AD: The Pre-symptomatic Evaluation of Novel or Experimental Treatments for Alzheimer's Disease (PREVENT-AD) cohort, based at the Centre for Studies on the Prevention of AD in Montreal, Canada (StoP-AD, douglas.research.mcgill.ca/stop-ad-centre) is a longitudinal study of older, healthy individuals (55+) with a parental or multiple-sibling history of AD [228]. Data for all variables were obtained from data release 5.0 (November 30, 2017, openpreventad.loris.ca/) except for *APOE* genotype, PET and genetic data. For these variables the latest available data at the center was used, to be included in future data releases. Each participant and study partner provided written informed consent. All procedures were approved by the McGill University Faculty of Medicine Institutional Review Board and complied with ethical principles of the Declaration of Helsinki.

ADNI: Data used in the preparation of this article were obtained from the Alzheimer's Disease Neuroimaging Initiative (ADNI) database (adni.loni.usc.edu). The ADNI was launched in 2003 as a public-private partnership, led by Principal Investigator Michael W. Weiner, MD in the U.S. The primary goal of ADNI has been to test whether serial MRI, PET, other biological markers, and clinical and neuropsychological assessment can be combined to measure the progression of MCI and early AD. For up-to-date information, see www.adni-info.org. For this study, a subset of ADNI consisting of individuals with genetic data and a family history of AD (first degree relative affected) were used. All data, but the CSF data, which were downloaded June 22, 2018, were downloaded on December 3, 2015.

ROSMAP: The Religious Orders Study and Rush Memory and Aging Project are two longitudinal clinical-pathologic cohort studies of aging and AD from the Rush Alzheimer's Disease Center in the U.S. (www.radc.rush.edu/) [285]. In this study a subset of individuals that had genetic and pathology data were used.

4.2.3 Genetic data

Quality control: QC procedures [286] for the genetic data was done similarly for all cohorts and were performed in PLINK v1.9 [232,233] as follows: heterozygous haploid genotypes excluded, sex check, relatives excluded (identity by descent > 0.1875), sample and genotyping call rate > 0.95, minor allele frequency > 0.05, and Hardy-Weinberg equilibrium < $1*10^{-6}$. SNPs were then matched with the GRCh37 genome (www.ncbi.nlm.nih.gov/assembly/GCF_000001405.13/#/st). A PCA with 1000 Genomes phase 3 data as a reference [236,240] was performed to determine ancestry and filter for individuals with European ancestry. Briefly, long range LD regions and ambiguous SNPs were first excluded from 1000 Genomes data and then merged with target cohort genetic data. Any ambiguous SNPs from target data were then excluded. Merged data were pruned with a sliding window of 2000 bp with a step size of 200 bp, excluding SNPs with an $r^2 > 0.2$ (--indep-pairwise 2000 200 0.2), and principle components (PCs) calculated (--pca) in PLINK [232,233]. Averages and standard deviations of PC1 and PC2 for Europeans in 1000 Genomes were determined and target cohort individuals were determined to be of European ancestry if their PC1 and 2 fell within ± 3 SD of the 1000 Genomes means. PCs were then calculated again within the European cohorts to include as covariates in subsequent analyses.

Imputation: PREVENT-AD, a subset of ADNI ("ADNI 1 GWAS" data set, see below), and ROSMAP data were imputed using the Sanger Imputation Service [237] (imputation.sanger.ac.uk/). Briefly, quality controlled genetic data was uploaded and pre-phased with SHAPEIT2 [238] and imputed with positional Burrows-Wheeler transform [239] using the 1000 Genomes cohort [236,240] as a reference panel. Only post-imputed SNPs with an info score greater than 0.7 were kept (similar to [287]) to balance quantity of excluded data (14% with 0.7 cut-off in [287]) with data quality.

Merging ADNI data: Two genomic data sets were used for ADNI; the "ADNI 1 GWAS" data set genotyped using the Illumina Human610-Quad BeadChip and the "ADNI WGS" data set genotyped using a whole genome sequencing platform. Before merging, both data sets were quality controlled

(including PCA) and "ADNI 1 GWAS" data was imputed. Some individuals were present in both data sets, in which case data from the "ADNI WGS" data set was used. The final genetic data set contain 6 164 853 SNPs and 1087 individuals.

4.2.4 TC polygenic score

Polygenic scoring was done with PLINK [232,233]. Using the PREVENT-AD cohort, SNPs were clumped with a sliding window of 250 kbp and filtering all SNPs with an LD r^2-value > 0.1 (--clump). Multiple weighted PGSs (using summary statistics β-values) were then calculated (--q-score-range) at various p-value cut-offs (1e-100, 1e-50, 1e-40, 1e-30, 1e-20, 1e-10, 1e-8, 1e-7, 1e-6, 1e-5, 1e-4, 1e-3, 1e-2, 0.05, 0.1, 0.5 and 1).

4.2.5 Phenotypes and statistics

All statistical analyses were performed in R [248]. For a full list of data sets, software and r packages used, and their respective links, see **Supplemental sTable 4.1** in Appendix 4. Values are reported as mean ± SE if not otherwise stated.

Descriptives: Differences in cohort characteristics, such as age, sex, and TC levels, were analyzed with either a Welch two sample t-test (comparing two cohorts) or an ANOVA (comparing the three cohorts) for continuous variables, and with a Pearson's chi-square test for categorical variables. Post-hoc analysis was performed if primary analysis was significant and comparisons were between all three cohorts. TukeyHSD was used for continuous variables and post-hoc chi-square test was used for categorical variables. R package "psych" [252] was used to compute summary statistics.

TC levels and p-value thresholding: In PREVENT-AD, TC levels were assessed in plasma drawn from non-fasting individuals mainly at the eligibility visit (i.e. before baseline measurements). In ADNI, TC levels were assessed in whole blood drawn at the screening visit from fasting individuals. ADNI TC measurements were transformed from mg/dl to mM to match the PREVENT-AD data by dividing values with 38.67. ROSMAP was not used for blood TC analyses. The relationship between each score and blood TC levels was evaluated with a linear regression with genetic PCs 1-10, sex,

statin use, *APOE*-ε4 status, age and age^2 as covariates. Standard deviations at each cut-off were determined by bootstrapping (n iterations = 5000) using the R package "boot" [288,289] and R^2-values were calculated using the "rcompanion" package [253]. Effects of statin use and sex on the relationship between the TC-PGS and TC levels were assessed by stratification; first by statin use and then by sex (in statin free individuals). The TC-PGS that explained most of the variance were selected for further analyses in all three cohorts.

Hypercholesterolemia: A hypercholesterolemia variable was created for PREVENT-AD and ADNI, by assuming that all individuals on statins and all non-treated individuals with TC levels > 6.2 mM, were hypercholesterolemic. Discrimination of hypercholesterolemic from healthy individuals was evaluated by receiver operator characteristics (ROC) curve analysis and quantified by AUC using the "pROC" package [290]. Data were stratified for sex and the difference between a model containing the covariates (PCs 1-10, age and age^2) and a model containing covariates plus the TC-PGS was evaluated with DeLong's test.

CSF measurements: In both PREVENT-AD and ADNI, CSF was obtained by lumbar puncture following an overnight fast. Levels of Aβ-42, p-TAU and total TAU were then measured by the Innotest® ELISA (Fujirebio) [242] and the Roche Elecsys CSF immunoassays (data file UPENNBIOMK9_04_19_17.csv) [243,244], for PREVENT-AD and ADNI respectively. Of note, the Elecsys Aβ-42 CSF immunoassay is currently under development for investigational use only and has an upper technical limit of 1700 pg/ml. Values above this limit are based on extrapolation of the calibration curve, and the performance of these values has not been formally established. These are still included in this study. In PREVENT-AD, Aβ-40 levels were further assessed by the MSD® MULTI-SPOT Assay System (V-PLEX Plus Aβ Peptide Panel 1 (6E10) Kit, MesoScale). Each dependent variable was examined for distribution pattern and transformed if not normally distributed and analyzed with a multiple linear regression. In PREVENTAD, models were corrected for genetic PCs 1-10, age, APOE-ε4 status, statin use and run with a sex*TC-PGS interaction term.

In ADNI, that has a bigger sample size, the same covariates were used but for statin that instead were included in the interaction term (statin*sex*TC-PGS).

PET measurements: PET scans were performed in PREVENT-AD using fluorine 18-labelled NAV4694 and 18-labelled AV-1451 (Flortaucipir) to estimate levels of brain Aβ and TAU, respectively. Standardized uptake value ratios (SUVR) were computed by dividing tracer uptake by cerebellar gray matter uptake (Aβ) or by inferior cerebellar gray matter uptake (TAU). For details on PET procedures, see [185]. The effect of TC-PGS on TAU PET data was investigated using a multiple linear regression similar to PREVENT-AD CSF analyses. Aβ data was not normally distributed and were therefore split into Aβ(-) and Aβ(+) similar to McSweeney and colleagues [291] (Aβ (+) defined as SUVR > 1.37). The Aβ(-) group had sufficient sample size to be analyzed similar to PREVENT-AD CSF analyses, whereas the sample size of the Aβ(+) group were to small to run the same regression. Thus, this latter regression was run with age, *APOE*-ε4 status, statin use and sex as covariates and only investigating the main effect of the TC-PGS.

Amyloid positivity status: According to the recently proposed biological definition of AD, we categorized individuals as on or off the AD spectrum by the presence of amyloid pathology in the brain [58,59]. In ADNI, we used the CSF p-TAU/Aβ-42 ratio as a proxy for brain amyloid pathology, as described by Hansson et al., [104]. Briefly, the CSF values were extracted from the last available visit for each individual, and a CSF p-TAU/Aβ-42 ratio ≥ 0.028 was considered as Aβ(+) and thus on the AD spectrum, whereas lower ratio was considered Aβ(-). In ROSMAP, semiquantitative estimates of neuritic plaque density as recommended by CERAD were used to define Aβ(+) individuals. This is a four point scale, and individuals with scores of three or four were considered Aβ(+). We evaluated whether the TC-PGS associated with risk of ending up on the AD spectrum by logistic regression. In ADNI the model included PCs 1-10, age, *APOE*-ε4 status and a statin*sex*TC-PGS interaction. Medication history was not available in ROSMAP, why the same model was run but without the statin variable.

Cognition: Cognition was analyzed in ADNI and ROSMAP. We limited our analyses to Aβ(+) individuals (see above for categorization) so as to only investigate individuals on the AD spectrum. In both ROSMAP and ADNI cognitive impairment (CI) was defined as having a clinical diagnosis of either MCI, AD or "Other dementia". To evaluate the association between the TC-PGS and overall risk of CI while also being Aβ(+), multiple logistic regressions were run including genetic PCs 1-10, age, *APOE*-ε4 status, and statin use (only for ADNI) as covariates and a sex*TC-PGS interaction.

Conversion rate: The effect of TC-PGS on conversion rate in ADNI and age of onset in ROSMAP, was evaluated with Kaplan-Meier survival analysis [255]. Including all individuals with a TC-PGS, individuals were categorized into tertiles (i.e. low, medium and high TC-PGS). In ADNI, individuals that were Aβ(+) with either no CI or with an MCI diagnosis at baseline were selected. Follow-up time ranged from three to 120 months. The "event" was defined as developing a clinical diagnosis of AD. In ROSMAP, Aβ(+) individuals (as defined above) were selected and the event was defined as a clinical diagnosis of either possible or probable AD. A larger sample size in ROSMAP allowed for stratification on sex. Analysis was done using the "survival" package [256,257], the "ggfortify" package [292,293] were used for plotting and "survminer" package [294] was used for creating survival tables.

4.3 Results

4.3.1 Cohort characteristics

Global Lipids Genetics Consortium: Meta-analysis summary data are based on 63 studies for a total of 114 230 individuals [280]. 15 studies were of non-European ancestry. The ratio of women in these studies ranged from 0 to 76.8% and the mean age ranged from 16 to 75 with a mean of 56.9 years. Most studies investigated individuals free of lipid lowering drugs (44/63) and the majority of studies had a fasting regime before cholesterol measurements (51/63). Raw data contained 2 446 981 SNPs, whereof 15.4% were ambiguous. These were excluded resulting in a data set with 2 069 037 SNPs. After matching with the target cohorts, the proportion of non-ambiguous SNPs present in each cohort was 86.4%, 89.7% and 91.3% for PREVENT-AD, ADNI and ROSMAP, respectively. After

filtering for SNPs not present in either data set, 1 653 356 SNPs remained, representing 67.6% of the original number of SNPs (see **Supplemental sFigure 4.1** in Appendix 4 for Manhattan plots of included and excluded SNPs).

Target cohorts: The three target cohorts, PREVENT-AD, ADNI and ROSMAP, differed in their percentages of females (70%, 48% and 71%, respectively, $p = 4.9*10^{-14}$), *APOE*-ε4 carriers (37%, 57% and 24%, respectively, $p = 1.5*10^{-23}$) and statin treated individuals (23% and 51% for PREVENT-AD and ADNI respectively, $p = 3.8*10^{-13}$, **Table 4.1**). Post-hoc analyses revealed that the proportion of females was higher in PREVENT-AD and ROSMAP compared to ADNI ($ps \leq 1.0*10^{-8}$) and that the proportion *APOE*-ε4 carriers were significantly different between all cohorts ($ps \leq 5.0*10^{-7}$). Age was recorded at the different assessments (blood, CSF, PET, Aβ(+) and cognition) and was different between the cohorts in all instances. Overall, PREVENT-AD (63.33 ± 0.42) is younger than ADNI (72.79 ± 0.35) that is younger than ROSMAP (88.42 ± 0.13, $ps < 2.0*10^{-16}$). TC levels was measured in PREVENT-AD and ADNI and were significantly higher in PREVENT-AD (5.42 ± 0.06 and 5.09 ± 0.06 mM, respectively, $p = 9.3*10^{-5}$).

4.3.2 Amount of variance explained in TC blood levels by TC-PGS varies with p-value cut-off, cohort, statin use and sex

To establish a TC-PGS that best associates with blood TC levels, various p-value cut-offs were investigated (**Figure 4.1**). First, the different scores were evaluated in a linear regression corrected for genetic PCs 1-10, sex, age, age^2, *APOE*-ε4 status and statin use, in PREVENT-AD (**Figure 4.1**, upper panel). At best, a TC-PGS based on a p-value cut-off of $1*10^{-6}$ explained 6.9% of the variance ($p = 2.93*10^{-8}$). Secondly, because statin use has a strong effect on TC levels, the scores were similarly evaluated but after stratification on statin use (consequently, statin use was excluded as a covariate from the model). Results revealed again a strong association of multiple TC-PGSs with TC levels, but only in the statin free individuals; in statin free individuals, again the p-value cut-off of $1*10^{-6}$ performed best with the TC-PGS explaining 13.5% of additional variance ($p = 2.83*10^{-9}$).

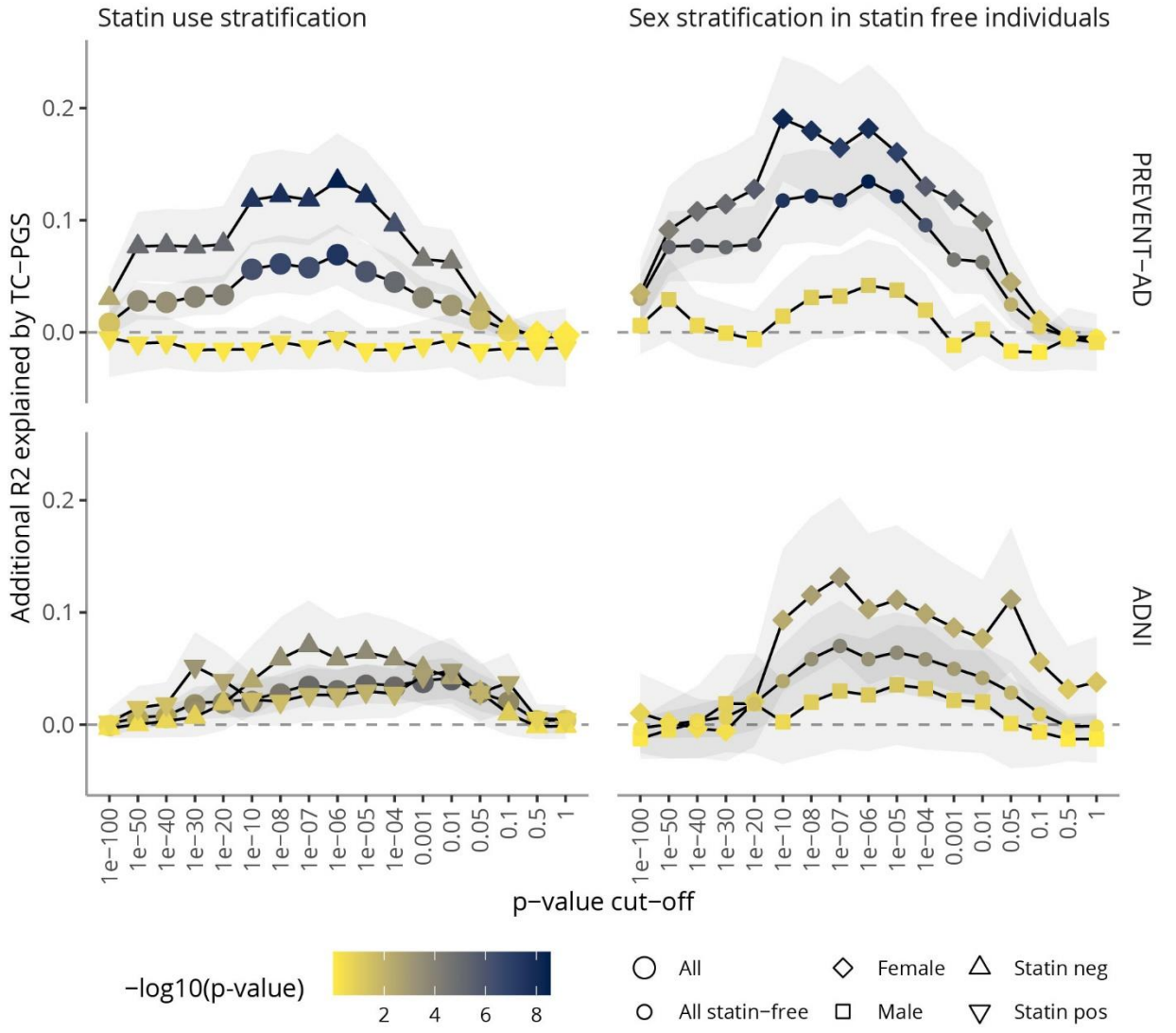
Table 4.1 Cohort characteristics

	PREVENT-AD		ADNI		ROSMAP		
Variable	**N**	**Mean (SE)**	**N**	**Mean (SE)**	**N**	**Mean (SE)**	**p-value**
Females [%]	306	69.9 (2.6)	401	47.9 (2.5)	547	71.1 (1.9)	<0.001 [a]
APOE-ε4 carriers [%]	302	37.1 (2.8)	401	56.9 (2.5)	546	24.2 (1.8)	<0.001 [a]
Age [years][#]	264	63.33 (0.42)	401	72.79 (0.35)	547	88.42 (0.13)	<0.001 [b]
Statin treated [%]	299	23.4 (2.5)	360	51.4 (2.6)	*NA*	*NA*	<0.001 [a]
TC measurments							
Age @ blood collection	287	63.15 (0.3)	355	72.63 (0.37)	*NA*	*NA*	<0.001 [c]
TC [mM]	287	5.42 (0.06)	355	5.09 (0.06)	*NA*	*NA*	<0.001 [c]
CSF measurements							
Age @ CSF collection	86	62.87 (0.59)	302	72.19 (0.41)	*NA*	*NA*	<0.001 [c]
Aβ-42 [pg/ml]	83	1160.31 (31.08)	301	1053.78 (37.17)	*NA*	*NA*	0.029 [c]
Aβ-40 [pg/ml]	75	6132.53 (222.57)	*NA*	*NA*	*NA*	*NA*	*NA*
p-TAU [pg/ml]	86	47.31 (1.89)	302	28.03 (0.76)	*NA*	*NA*	<0.001 [c]
TAU [pg/ml]	86	275.23 (13.87)	302	291.45 (6.95)	*NA*	*NA*	0.298 [c]
PET measurements							
Age @ Aβ PET	98	67.62 (0.49)	*NA*	*NA*	*NA*	*NA*	*NA*
Aβ PET [SUVR]	98	1.33 (0.04)	*NA*	*NA*	*NA*	*NA*	*NA*
Age @ TAU PET	100	70.79 (0.56)	*NA*	*NA*	*NA*	*NA*	*NA*
TAU PET [SUVR]	100	1.07 (0.01)	*NA*	*NA*	*NA*	*NA*	*NA*
Aβ(+)							
Age @ assessment	*NA*	*NA*	271	74.22 (0.45)	535	88.38 (0.13)	<0.001 [c]
Aβ(+) [%]	*NA*	*NA*	271	54.2 (3.0)	535	74.6 (1.9)	<0.001 [a]
Cognition							
Age @ assessment	*NA*	*NA*	139	76.62 (0.63)	399	88.5 (0.15)	<0.001 [c]
CI [%]	*NA*	*NA*	139	90.6 (2.5)	399	76.9 (2.1)	<0.001 [a]

Abbreviations: CI, cognitively impaired; CSF, cerebrospinal fluid; PET, positron emission tomography; SUVR, standardized uptake value ratio; TC, total cholesterol.
[#] Mean age was calculated for baseline in PREVENT-AD and ADNI and for age of death in ROSMAP.
[a] Calculated with χ^2 test
[b] Calculated with ANOVA
[c] Calculated with Welch two sample t-test

In contrast, none of the scores significantly associated with TC levels in statin treated individuals (ps ≥ 0.412). Finally, the statin free individuals were further stratified based on sex, revealing a highly significant effect in females; the best performing p-value cut-off of $1*10^{-10}$ explaining 19.0% of the

Figure 4.1 Variance explained by TC-PGS on TC levels over multiple p-value cut-offs



Multiple p-value thresholds were evaluated for association with TC levels in PREVENT-AD (upper panel) and ADNI (lower panel) stratified for statin use (left hand panel) and sex (right hand panel) using linear. Plotted is additional variance explained after adding the TC-PGS to the model. Shaded areas indicate standard deviation. For statistical details see 4.2.5 Phenotypes and statistics and 4.3 Results.
Abbreviations: TC, total cholesterol; TC-PGS, total cholesterol polygenic score.

variance ($p = 6.70*10^{-9}$), but only a trend in males with p-value cut-off $1*10^{-6}$ explaining 4.2% of the variance ($p = 0.066$).

The same procedure was repeated in ADNI (**Figure 4.1**, lower panel). When all individuals were analyzed together, at best, the TC-PGS explained 4.1 % of the variance ($p = 7.1*10^{-6}$) at a p-value cut-off of 0.01. Stratification on statin use revealed that the TC-PGS now at best explained 7.1 % of the variance in statin free individuals ($p = 1.4*10^{-4}$, p-value cut-off $1*10^{-7}$) and 5.2 % of the variance in statin treated individuals ($p = 7.2*10^{-4}$, p-value cut-off $1*10^{-30}$). Further stratification on sex in statin free individuals revealed that at best, the TC-PGS explained 13.1% of the variance ($p = 9.6*10^{-4}$, p-value cut-off $1*10^{-7}$) in females and 3.5 % of the variance ($p = 0.053$, p-value cut-off $1*10^{-5}$) in males. We decided to use the TC-PGS with a p-value cut-off of $1*10^{-6}$ based on its performance in the younger PREVENT-AD cohort and this score will from here on out be referred to solely as "TC-PGS".

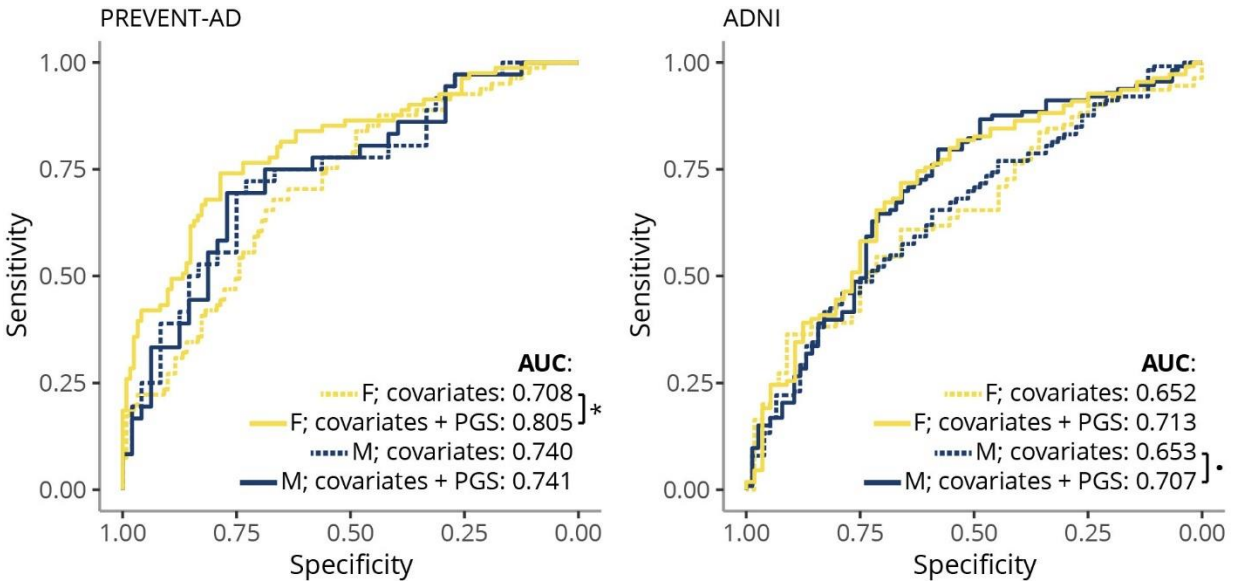
4.3.3 TC-PGS predicts hypercholesterolemia

We examined the TC-PGS' ability to predict hypercholesterolemia in PREVENT-AD and ADNI (**Figure 4.2**). ROC curve analysis was used to compare the performance of a logistic regression model including only covariates (genetic PCs 1-10, age and age^2) with a model including covariates and the TC-PGS, stratified for sex. In PREVENT-AD, adding the TC-PGS to the model significantly improved AUC in females (from 70.8 % to 80.5 %, $p = 0.0042$) but had no effect in males (74.0 % *vs* 74.1 %, $p = 0.91$). In ADNI, although adding the TC-PGS increased the AUC values for both females (65.2 % *vs* 71.3 %, $p = 0.14$) and males (65.3 % *vs* 70.7 %, $p = 0.087$), these increases did not reach significance.

4.3.4 TC-PGS does not associate with amyloid pathology

The effect of TC-PGS on Aβ pathology was assessed in PREVENT-AD and ADNI (**Figure 4.3**). In PREVENT-AD, linear regression correcting for covariates and with a sex*TC-PGS interaction term revealed no effect of TC-PGS, either as part of the interaction term or as a main effect, on CSF Aβ-42 ($-0.172 \leq t(16, 66) \leq 0.269$, $ps \geq 0.788$) or its ratio with Aβ-40 ($-0.271 \leq t(16, 58) \leq 0.024$, $ps \geq 0.787$). Similarly, in ADNI no significant effect of TC-PGS could be detected on CSF Aβ-42, neither as a main

Figure 4.2 Effect of TC-PGS on prediction of hypercholesterolemia



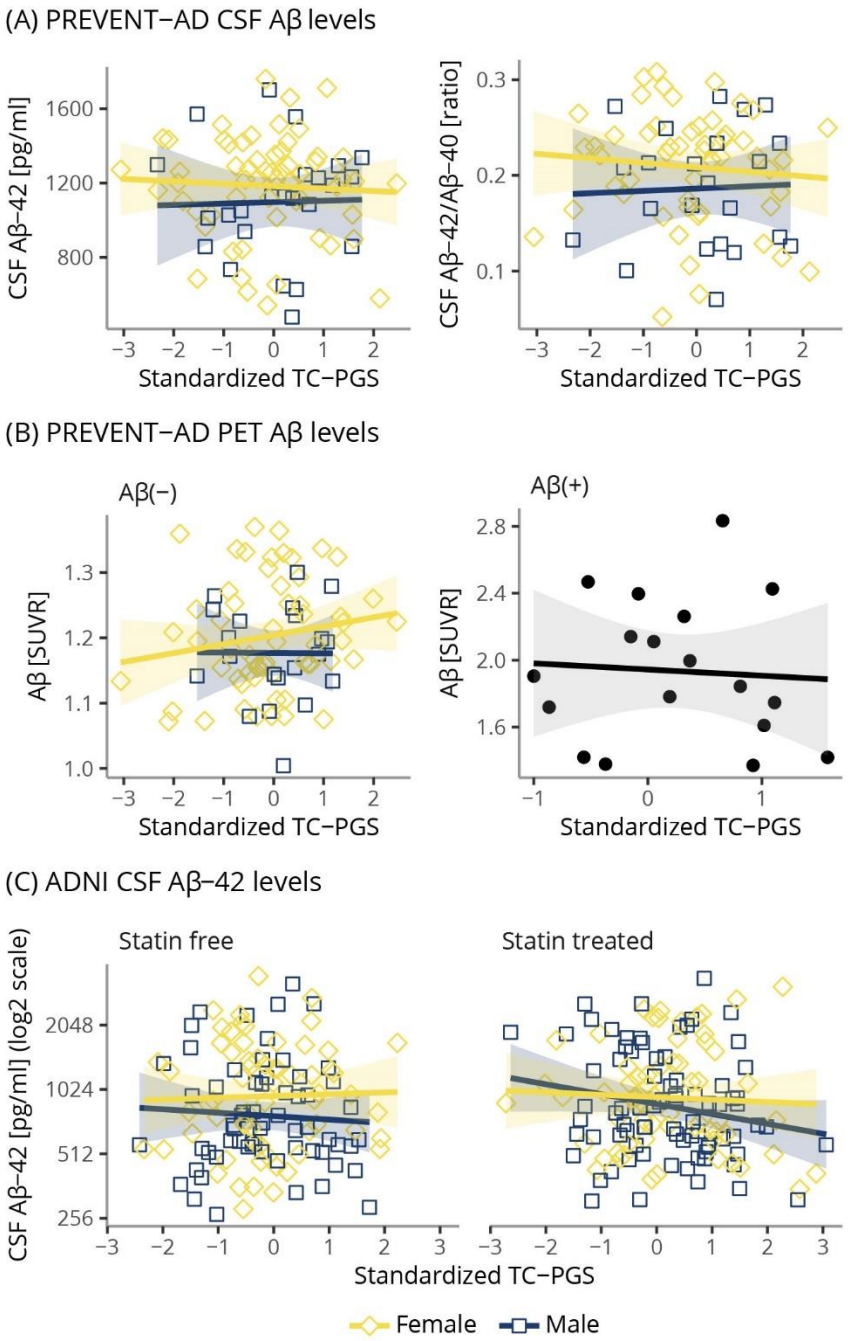
ROC curves showing the effect of covariates and TC-PGS on predicting hypercholesterolemia in PREVENT-AD (left hand panel) and ADNI (right hand panel) stratified for sex. Individuals were deemed hypercholesterolemic if they were on statins or if they had TC levels > 6.2 mM (i.e. 240 mg/dl). For statistical details see 4.2.5 Phenotypes and statistics and 4.3 Results.
Abbreviations: AUC, area under the curve; F, females; M, males; PC, genetic principle component; TC-PGS, total cholesterol polygenic score.
* $p \leq 0.05$
• $p \leq 0.1$

effect nor as part of any of the interaction terms (sex*TC-PGS, statin*TC-PGS, statin*sex*TC-PGS ($-0.869 \leq ts(19, 250) \leq 0.663$, $ps \geq 0.371$). Further, the effect of TC-PGS was evaluated on Aβ pathology as assessed by PET in PREVENT-AD. To normalize the distribution of the data, data was split into Aβ(-) and Aβ(+) individuals (see 4.2.5 Phenotypes and statistics; PET measurements). In addition, due to the very small sample size of the Aβ(+) group, the genetic PCs were excluded from the model and was run without any interaction terms. We found no effect of the TC-PGS in either Aβ(-) ($-0.672 \leq t(16, 63) \leq 1.095$, $ps \geq 0.278$) or Aβ(+) individuals ($t(5, 11) = -0.213$, $p = 0.835$).

4.3.5 TC-PGS does not associate with TAU pathology

The TC-PGS was evaluated for associations with biomarkers of p-TAU pathology in PREVENT-AD (CSF and PET) and ADNI (CSF, **Figure 4.4**). In PREVENT-AD, linear regression corrected for

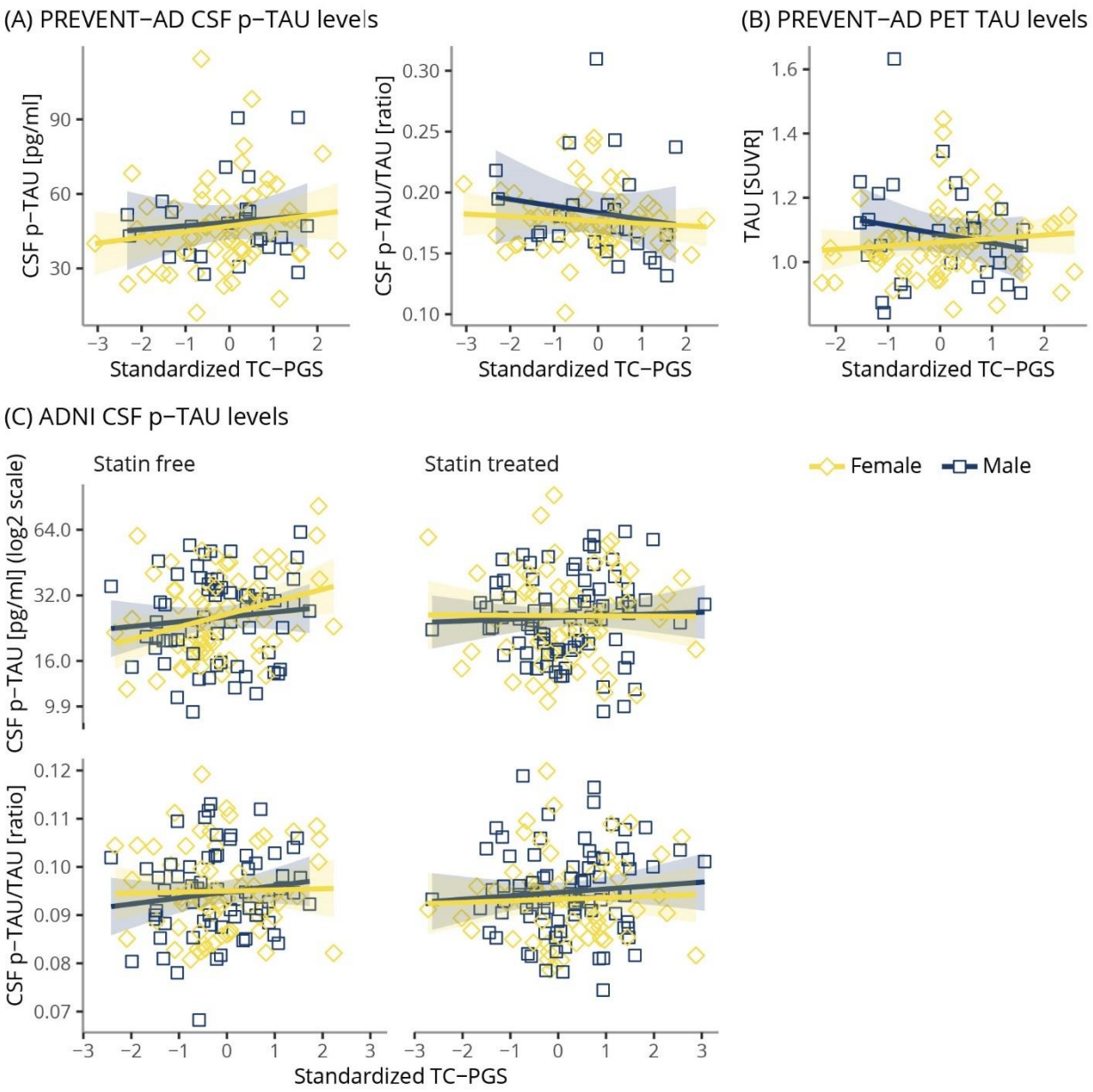
Figure 4.3 Associations of TC-PGS with biomarkers of amyloid pathology



Aβ pathology biomarkers plotted against TC-PGS and assessed by multiple linear regressions; PREVENT-AD CSF levels of Aβ-42 and its ratio with Aβ-40 (A), PREVENT-AD Aβ brain levels, as assessed by PET after stratification by Aβ(+) status (SUVR > 1.37) (B), and ADNI CSF levels of Aβ-42 stratified for statin use (C). Shaded areas indicate 95% confidence intervals. For statistical details see 4.2.5 Phenotypes and statistics and 4.3 Results. Abbreviations: CSF, cerebrospinal fluid; PET, positron emission tomography; SUVR, standardized uptake value ratio; TC-PGS, total cholesterol polygenic score.

covariates and run with a sex*TC-PGS interaction revealed no associations between TC-PGS and biomarkers of p-TAU pathology as assessed by CSF p-TAU ($-0.069 \leq t(16, 69) \leq 0.641$, $ps \geq 0.523$), CSF p-TAU/TAU ratio ($-0.919 \leq t(16, 69) \leq 0.819$, $ps \geq 0.361$) and TAU PET ($-1.279, t(16, 83) \leq$

Figure 4.4 Associations of TC-PGS with biomarkers of p-TAU pathology



p-TAU pathology biomarkers plotted against TC-PGS and assessed by multiple linear regressions; PREVENT-AD CSF levels of p-TAU and its ratio with total TAU levels (A), PREVENT-AD TAU brain levels, as assessed by PET (B), and ADNI CSF levels of p-TAU and its ratio with total TAU (C). Shaded areas indicate 95% confidence intervals. For statistical details see 4.2.5 Phenotypes and statistics and 4.3 Results.

Abbreviations: CSF, cerebrospinal fluid; PET, positron emission tomography; SUVR, standardized uptake value ratio; TC-PGS, total cholesterol polygenic score.

1.641, ps ≥ 0.105). Similarly, in ADNI, we found no significant associations between the TC-PGS and p-TAU pathology as assessed by CSF p-TAU (-1.212 ≤ ts(19, 251) ≤ 1.247, ps ≥ 0.214) and its ratio with TAU (-0.599 ≤ ts(19, 251) ≤ 0.123, ps ≥ 0.550).

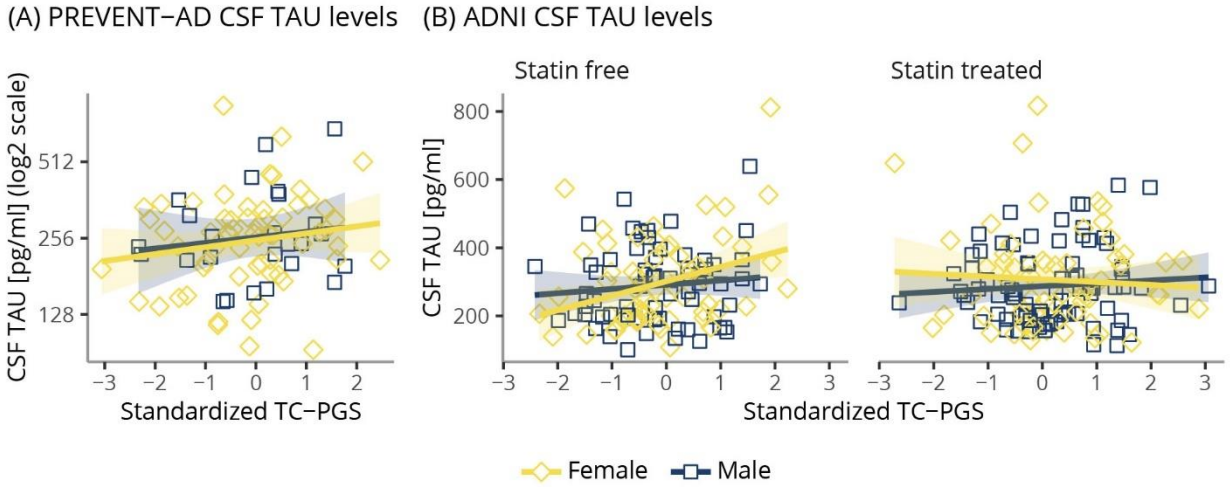
4.3.6 TC-PGS does not associate with markers of neurodegeneration

The TC-PGS was evaluated for associations with biomarkers of neurodegeneration in PREVENT-AD and ADNI by measuring levels of CSF TAU (**Figure 4.5**). We found no evidence for an association of the TC-PGS with CSF TAU in neither PREVENT-AD (-0.111 ≤ t(16, 69) ≤ 0.758, ps ≥ 0.451) nor ADNI (-1.393 ≤ ts(19, 251) ≤ 1.533, ps ≥ 0.127).

4.3.7 TC-PGS does not associate with increased risk of becoming Aβ(+)

The association between TC-PGS and risk of ending up on the AD spectrum were evaluated in ADNI and ROSMAP (**Table 4.2**). Individuals in ADNI and ROSMAP were categorized based on the presence of Aβ pathology in the brain as either Aβ(-) or Aβ(+) (see 4.2.5 Phenotypes and statistics;

Figure 4.5 Associations of TC-PGS with biomarkers of neurodegeneration



Biomarkers of neurodegeneration as assessed by CSF levels of TAU plotted against TC-PGS. Data analyzed by multiple linear regressions; PREVENT-AD CSF levels of TAU (A) and ADNI CSF TAU levels (B). Shaded areas indicate 95% confidence intervals. For statistical details see 4.2.5 Phenotypes and statistics and 4.3 Results.
Abbreviations: CSF, cerebrospinal fluid; TC-PGS, total cholesterol polygenic score.

Amyloid positivity status). In neither ADNI (-0.821 ≤ zs ≤ 1.108, ps ≥ 0.268) nor ROSMAP (-0.607 ≤ zs ≤ -0.293, ps ≥ 0.544) did we find any significant effect of TC-PGS on the risk of becoming Aβ(+).

4.3.8 TC-PGS does not associate with cognition in Aβ(+) individuals

Finally, we evaluated whether the TC-PGS associated with risk of becoming cognitively impaired in ADNI and ROSMAP (**Table 4.2**). For this analysis we used the subset of individuals that were Aβ(+) and defined cognitive impairment as having any diagnosis of cognitive impairment (e.g. including MCI and other dementias) at the last recorded visit. In neither ADNI (0.179 ≤ zs ≤ 0.283, ps ≥ 0.777) nor ROSMAP (-0.952 ≤ zs ≤ 1.088, ps ≥ 0.276) could we detect any significant association between the TC-PGS and risk of becoming cognitively impaired.

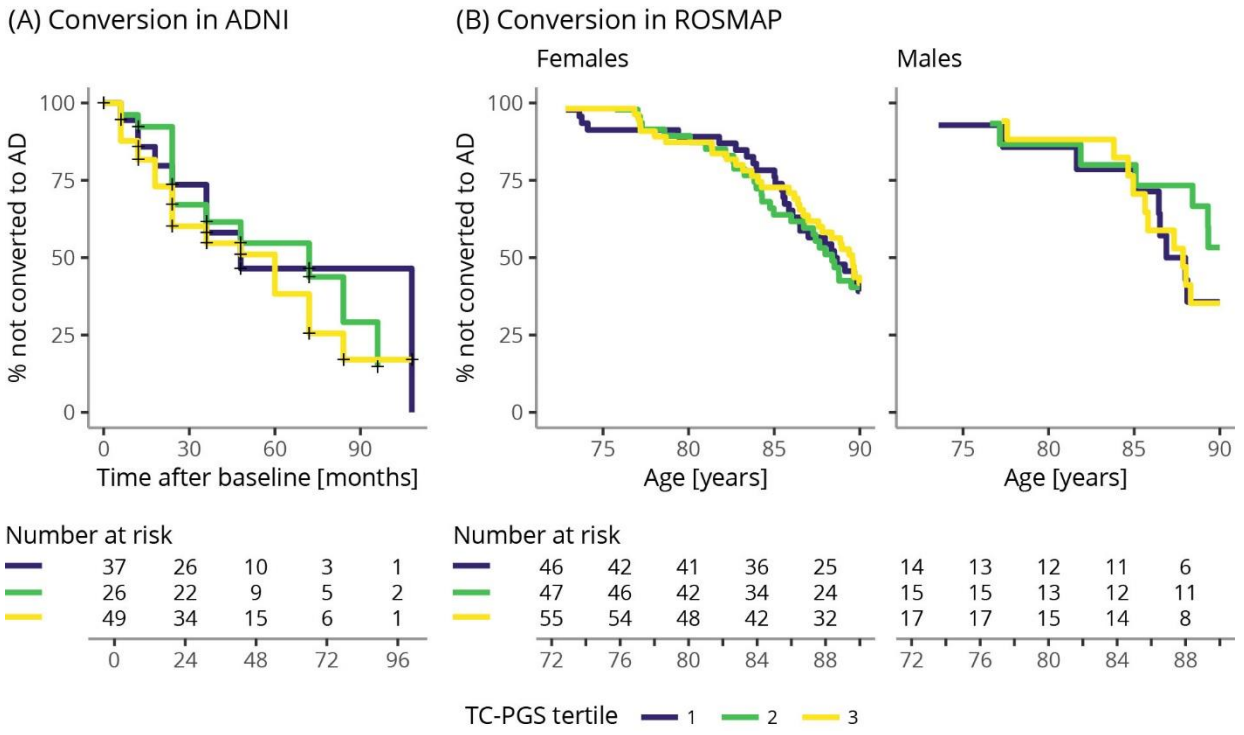
We also evaluated whether the TC-PGS had any effect on the conversion rate (ADNI, **Figure 4.6 A**) or age of onset (ROSMAP, **Figure 4.6 B**). A subset of ADNI was selected; Aβ(+) individuals either non-CI or with an MCI diagnosis at baseline. The event was defined as receiving a clinical diagnosis of AD. Tertiles of the TC-PGS were then evaluated in a Kaplan-Meier analysis. We found no difference between the TC-PGS tertiles on conversion rate in ADNI ($\chi^2(2) = 1.1$, $p = 0.6$). In ROSMAP we examined age at onset of a clinical diagnosis of possible or probable AD. Again, we found no difference between the TC-PGS tertiles ($\chi^2(2) = 0.2$, $p = 0.9$).

Table 4.2 Associations of TC-PGS with disease status

	ADNI				ROSMAP			
	Aβ(+)		CI		Aβ(+)		CI	
TC-PGS terms	ORs	p-value	ORs	p-value	ORs	p-value	ORs	p-value
TC-PGS	0.82	0.577	1.08	0.858	0.95	0.769	0.76	0.341
Statin*TC-PGS	1.68	0.268	*NA*	*NA*	*NA*	*NA*	*NA*	*NA*
Sex*TC-PGS	1.08	0.867	1.18	0.777	0.87	0.544	1.40	0.276
Statin*Sex*TC-PGS	0.60	0.412	*NA*	*NA*	*NA*	*NA*	*NA*	*NA*

ORs for interaction terms refer to statin treated, females with an increase in 1 SD of TC-PGS. NA indicates that interaction was not investigated due to insufficient sample size. For statistical details see 4.2.5 Phenotypes and statistics and 4.3 Results.
Abbreviations: CI, cognitively impaired; OR, odds ratio; TC-PGS, total cholesterol polygenic score.

Figure 4.6 No effect of TC-PGS on conversion rate and age of onset



Kaplan-Meier survival curves displaying conversion rates to a clinical diagnosis of AD depending on TC-PGS tertiles in Aβ(+) individuals; conversion rate from healthy or MCI to AD as assessed by months after baseline visit in ADNI (A) and age of AD onset in ROSMAP, stratified for sex (B). For statistical details see 4.2.5 Phenotypes and statistics and 4.3 Results.
Abbreviations: TC-PGS, total cholesterol polygenic score.

4.4 Conclusion

4.4.1 TC-PGS and TC levels

In this study we have created a TC-PGS that associates with blood TC levels in two different AD related cohorts. We show that the variability explained by the score depends on cohort, selection of SNPs to include in the score, statin use and sex. We used clumping and p-value thresholding as a method for pruning SNPs to include in the scores, and thus evaluated a number of p-value threshold in both PREVENT-AD in ADNI. We saw that the score which explained most of the variance, varied between the two cohorts and depending on stratification on statin use and sex. In the non-stratified analyses, a p-value threshold of $1*10^{-6}$ performed best in PREVENT-AD, while a threshold of 0.01

performed best in ADNI. In addition, the scores in general performed better in PREVENT-AD than in ADNI (**Figure 4.1**). Further, stratification on both statin use and sex had remarkable effect on the scores' performance in PREVENT-AD, and less so in ADNI. For example, in PREVENT-AD, the TC-PGSs had significant associations in statin free females, with no significant associations in statin treated males.

Similarly, examining the predictive ability of the TC-PGS on hypercholesterolemia, revealed a significant improvement in PREVENT-AD females, increasing AUC from 0.708 to 0.805 but not in males (**Figure 4.2**). In ADNI we detected similar trends for improved AUCs in females and males, but this did not reach significance (ps > 0.08).

The discrepancies between cohorts could possibly be due to the differences between PREVENT-AD and ADNI (**Table 4.1**) that differ on the proportions of females, *APOE*-ε4 carriers and statin users, as well as age. With sex and statin stratifications we see that results do become more similar, further supporting the importance of taking these factors into account. Another factor that could affect the associations, is that cholesterol measurements were taken after fasting in ADNI, whereas in PREVENT-AD non-fasted samples were used, although studies have shown that TC levels are little influenced by fasting conditions [295,296].

The *APOE* gene locus is one of the most important for TC levels. The top SNP in the meta-analysis results used here [280] is indeed rs7412 - one of the SNPs determining the *APOE*-ε2 allele. rs7412 C allele gives rise to either the ε3 or ε4 isoform and associates with increased TC levels ($\beta = 0.374$, $p = 1.560*10^{-283}$). The second SNP in the *APOE* haplotype, rs429358, determines the ε4 allele, and is not present in the summary data. Nevertheless, rs429358 have been shown to associate with TC levels in other big GWA studies [297,298] such that the C allele, that results in the ε4 isoform, associates with increased TC levels. Although we did not examine how stratification for *APOE*-ε4 status would affect the relationship between the TC-PGS and TC levels it is possible that the effect of the TC-PGS would be less in *APOE*-ε4 carriers due to its strong relationship with TC levels. Furthermore, if this is true,

that would explain the reduced effect in ADNI since this cohort has a higher percentage of *APOE*-ε4 carriers than PREVENT-AD (**Table 4.1**).

There is also the fact that the two cohorts differ in terms of age; ADNI being on average 10 years older than PREVENT-AD. TC levels change with age; between younger adulthood over midlife and into older adulthood, TC levels increase with age [299], however, levels appear to be decreasing with age in older adults (>70 years old) [300,301]. This altered metabolism of cholesterol with age possibly involves different sets of genes and could thus explain why the TC-PGS behave differently in the two differently aged cohorts. This hypothesis would however have to be further investigated using either longitudinal studies or cross-sectional studies covering a bigger range of ages. Considering that increased midlife levels of TC [16,19,20,189] associate with increased risk of AD, it is interesting that our TC-PGS performs better in PREVENT-AD, which is closer to midlife than ADNI, thus suggesting it is better capturing midlife than late-life cholesterol levels.

The interaction between age and sex could also be of interest. For example, menopause in women are associated with increased TC levels and risk of cardiovascular disease [302,303] and hormone replacement therapy has shown to decrease TC levels [304]. PREVENT-AD are younger and have a higher percentage of females compared to ADNI and one could thus hypothesize that discrepancies in TC metabolism could also be influenced by discrepancies in the proportions of individuals that underwent menopause and treatment thereof.

Finally, comparing to the study by Proitsi [213], our results show that the variance explained in blood TC levels by a TC-PGS can be vastly improved (3.6% in [213] *vs* 18.2% for p-value cut-off $1*10^{-6}$ in statin free females in PREVENT-AD) by considering statin use and sex.

4.4.2 TC-PGS and AD

Contrary to our significant findings between a TC-PGS and TC blood levels, the TC-PGS showed no associations with any biomarkers of AD (**Figures 4.3** and **4.4**) or neurodegeneration (**Figure 4.5**),

nor with cognition (**Figure 4.6**, **Table 4.2**). Similarly, the TC-PGS did no associate with the risk of becoming Aβ(+) (**Table 4.2**).

The relationship between vascular factors and AD biomarkers was recently assessed in PREVENT-AD and showed that vascular factors, including TC levels, associates with increased Aβ pathology, but only in individuals free of vascular medication [185]. There are some differences between this study and that one, in that here individuals were grouped based solely on statin use, whereas Köbe et al., included other medications relevant to cardiovascular disease. Further, we decided to focus on the interaction with sex rather than with statin use, and thus a statin interaction was not included in our PREVENT-AD analyses (due to small sample size). However, in ADNI we had a sufficient sample size to include statin use as an interaction, and also in this cohort we could not find evidence for any influence on the TC-PGS on AD biomarkers. Thus, it is possible that in order to see an effect between TC-PGS and AD biomarkers, all vascular medication, rather than just statin use, need to be taken into consideration.

It is also possible that there is an additive effect of vascular risk factors such that the TC-PGS alone is not sufficient to have an effect on AD. Indeed, there are ample evidence that such an additive effect exists. Kivipelto et al., showed in multiple studies that there is an additive effect of TC levels, blood pressure and *APOE*-ε4 [16–18], leading to the development of the cardiovascular risk factors, aging, and dementia (CAIDE) score [305]. This score takes into account age, sex, education, systolic blood pressure, BMI, cholesterol, physical activity and *APOE*-ε4 status and has been validated as a predictor for AD [306]. Similarly, vascular burden scores, taking into account factors such as hyperlipidemia, diabetes and hypertension, associates with impaired executive function and lower the threshold of amyloid burden needed to result in cognitive impairment [307].

As mentioned above, *APOE* is important both for TC levels and AD. In this study design, we decided to keep the *APOE* gene locus in the TC-PGS but to correct for *APOE*-ε4 status in each regression model. Thus, the associations between the TC-PGS and TC levels are in addition to any effect of

APOE-ε4 status. Similarly, the lack of association between TC-PGS and AD is after correcting for *APOE*-ε4 status. It is thus possible that the increased risk of AD seen in *APOE*-ε4 carriers could still be mediated in part by increasing cholesterol. Thus, our results show that genetic variants, other than the genetic variants resulting in the *APOE*-ε4 isoform, strongly correlate with TC levels but fail to associate with AD pathology.

In summary, we have created a TC-PGS that associates with TC levels and significantly improves prediction of hypercholesterolemia, specifically in statin free females with European ancestry. We could however not prove any significant associations with AD; neither on the neuropathological underpinnings nor on cognition. It is possible that explaining ~18% of the variance in blood TC levels is still not enough to find significant associations with AD. For example, while it has previously been shown that TC levels associates with Aβ pathology in PREVENT-AD [185], the TC-PGS did not in the same cohort, which would suggest that we would possibly need a bigger sample size. Furthermore, considering the fact that there is an additive effect of vascular risk factors on AD, it is still possible that the TC-PGS could have an effect on AD in individuals at higher cardiovascular risk (e.g. *APOE*-ε4 carriers). Further research is warranted to establish the role of a TC-PGS in AD.

Chapter 5

Discussion

In this thesis we have tried to identify some of the missing genetic heritability in AD using hypothesis driven experimental designs to enable detection of low effect loci. Based on the knowledge of the implications of altered cholesterol metabolism in AD, the studies presented in this thesis build upon the overarching hypothesis that SNPs in loci related to cholesterol metabolism influence AD risk by changing levels of cholesterol, either centrally on a cellular level or peripherally on a systematic level (i.e. blood cholesterol levels).

5.1 Main findings on the involvement of cholesterol related genetics in AD

Our first hypothesis regarded HMGCR, the rate-limiting step in cholesterol synthesis and the target of statins [165,167], drugs that have been shown to associate with reduced risk of AD [22,23]. We hypothesized that genetic variants in the *HMGCR* gene locus would associate with AD by either increasing or decreasing the expression of *HMGCR* resulting in changes in cholesterol levels; systemic increases of blood cholesterol levels in mid-life has been proven to associate with increased AD risk (discussed in Chapter 4) and reducing cellular cholesterol has been proven to reduce Aβ production (discussed in Chapter 2 and 3). In line with our hypothesis, in Chapter 2, we found a SNP, rs72633963 (A allele), in the *HMGCR* gene locus that associated with reduced neuritic plaque density in the brain and improved immediate memory using two Quebec-based cohorts. These changes were seen specifically in *APOE*-ε4 carriers and were accompanied by associations with reduced TC and LDL-C levels. Findings on HMGCR metabolism was non-conclusive and possibly hampered by discrepancies

between cohorts and the lack of complete characterization of HMGCR measures within cohorts (discussed in detail below).

In Chapter 3 we aimed to evaluate the effect of an already identified *HMGCR* SNP, rs3846662. Using stem cell technology, we were able to investigate effects in multiple cell types (RECs, iPSCs, NPCs, and neurons) from the same set of individuals. In accordance with the literature, we found associations with cholesterol metabolism; the protective AA genotype associated with lower levels of FL-*HMGCR* and higher levels of the alternatively spliced transcript Δ13-*HMGCR*. This was accompanied by a compensatory increase in HMGCR protein levels and, although not significant, a decrease in activity levels. We could not find any effect on extracellular Aβ-40 and Aβ-42 or intracellular TAU levels. However, we believe we are critically underpowered to detect any potential effects on these markers and are therefore conservative in drawing any definitive conclusions regarding these particular findings.

Due to the weak and sometimes contradictory associations of rs72633963 and rs3846662 on AD risk and AD biomarkers, we hypothesized that by combining the effect of multiple SNPs using polygenic scores, we would be better positioned to detect an effect on AD. Hence, in Chapter 4, the combined effect of SNPs influencing peripheral blood TC levels was assessed by constructing a TC-PGS and then evaluating it with regards to AD risk, pathology and biomarkers thereof. We showed that the TC-PGS associated with TC levels in two different cohorts, and that these associations were affected by statin use and sex such that the TC-PGS performed best in statin free, females. The TC-PGS improved the prediction accuracy of hypercholesterolemia by ~10% and explained 18% of the variance in TC levels. This is comparable to the effect of *APOE*-ε4 on TC levels that explains about 12% of the variance [308]. We could however not detect any significant associations with AD; neither on risk nor on the levels of biomarkers.

Taken together, while our findings from Chapter 2 and 3, suggest a role for individual SNPs with effects on cholesterol metabolism in AD, our TC-PGS suggest little overlap between genetics

underlying peripheral TC levels and AD. These findings and discrepancies are discussed in detail below.

5.2 Possible mechanisms of cholesterol related genetics in AD

5.2.1 Causality of SNPs

The findings in this thesis are solely *associations* and does not infer *causality*. As such, both rs3846662 and rs72633963 could be markers of other genetic events that indeed causes the effects that we see on HMGCR metabolism and TC levels. Effects of rs3846662 have been reported in the literature and would suggest that rs3846662 is indeed a causal SNP. For example, transfecting cells with minigene constructs with either the AA or GG genotype, showed that transfections with the AA genotype resulted in higher rates of splicing as assessed by levels of Δ13-*HMGCR*/total *HMGCR* RNA ratio [32]. Thus, our findings go well with what has been shown before and we believe that the effects that we see in Chapter 3, are likely due to the rs3846662 genotype and follow the hypothesis of rs3846662 A allele promoting alternative splicing of *HMGCR* ultimately leading to reduced activity of the HMGCR protein (**Figure 5.1**, lower panel).

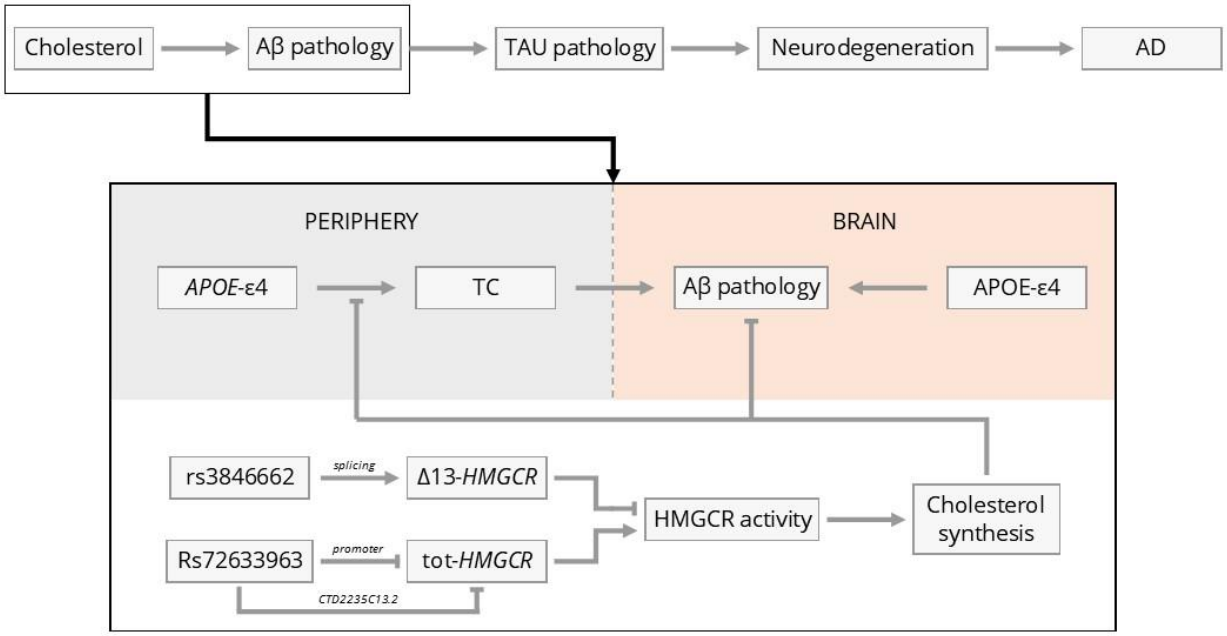
On the other hand, rs72633963 is a hitherto unexplored variant and it is thus possible that the effects are actually mediated by other genetic elements. This could explain the discrepancies we see between ADNI and the Quebec based cohorts in Chapter 2; a causal SNP would have a higher probability of associating with traits across populations than a SNP that "tag" the effect of a causal element because LD between the SNP and the causal element might differ between populations and cohorts (the effects of population structure is discussed in detail below, see 5.3.1 Ancestry/population stratification).

Rs72633963 is located to a region encompassing two features; the *HMGCR* promoter region and exon 2 of the *CTD2235C13.2* antisense transcript (see **Figure 2.1** in Chapter 2). If indeed causal, one could hypothesize that rs72633963 exerts its effect either by influencing transcription of *HMGCR* by altering motifs for transcription factors (such as SREBF2) or by having a direct effect on the *CTD2235C13.2* transcript. *CTD2235C13.2* is denoted as a novel transcript encoding a long non-coding RNA in

Ensembl (ensemble.org) and although this specific transcript is not explored in the literature, it has been shown that these specific types of RNAs can alter transcription of its nearby gene [309]. As a variant that is located in an exon, rs72633963 could thus influence the function of the *CTD2235C13.2* transcript which would subsequently affect expression of *HMGCR*. Taken together, this suggests that rs72633963 could have direct or indirect effects on the transcription and expression levels of *HMGCR* and in this way affecting activity (**Figure 5.1**, lower panel).

We thus hypothesize a model in which rs72633963 and rs3846662 through distinct mechanisms on *HMGCR* expression (total *HMGCR* vs Δ13-*HMGCR*, respectively) ultimately affect HMGCR activity to lead to decreased cholesterol synthesis (**Figure 5.1**, lower panel).

Figure 5.1 Proposed mechanism of cholesterol related genetics in AD



We propose that TC levels, either peripherally or centrally primarily affect Aβ levels, ultimately increasing the risk of developing AD. The effect of HMGCR SNPs are hypothesized to occur through direct or indirect effects of *HMGCR* expression, subsequently affecting activity resulting in reduced cholesterol synthesis. This reduction can have effects in the periphery or directly in the brain.
Abbreviations: Δ13-HMGCR, *HMGCR* transcript lacking exon 13; tot-HMGCR, total HMGCR.

5.2.2 Peripheral effect

SNPs mediating effect through peripheral TC levels: Previously described *HMGCR* variants (rs3846662, rs3761740 and rs5909) as well as rs72633963 described in this thesis, associate with peripheral blood TC levels [32,226,227]. Elevated TC levels, especially in midlife, are known to associate with increased AD risk [16–20,189] and multiple biomarkers of AD [184–186,189]. It is thus tempting to hypothesize that these SNPs influence AD risk by increasing or decreasing TC levels due to their effect on *HMGCR* expression levels and/or activity (**Figure 5.1**). Indeed, in PREVENT-AD it has recently been shown that TC levels associate with increased Aβ pathology (as assessed by CSF Aβ-42 and PET) [185]. Using the same cohort, we here show rs72633963 A allele to associate negatively with TC levels which was accompanied by a trend in reductions in Aβ pathology (as assessed by CSF Aβ-42, p = 0.108).

Because of the weak and sometimes contradictory effects of rs72633963 on AD identified in Chapter 2, and due to the implication of rs72633963 and rs3846662 effect on TC levels, we hypothesized that by combining multiple SNPs associating with TC levels into one score we would be able to see a more robust effect on AD and to test in a broader sense if SNPs affect AD risk by influencing TC levels. This led to the construction of the TC-PGS evaluated in Chapter 4. However, contrary to our hypothesis based on the findings from Chapter 2, the TC-PGS did no associate with neither AD risk nor AD biomarkers but can possibly be explained by a number of confounding factors discussed below.

***APOE*-ε4 interaction**: Previous studies of rs3846662 have indicated protective effects specifically in *APOE*-ε4 carriers [13,34] and we found similar interaction effects of rs72633963 on both Aβ pathology and cognition. As *APOE*-ε4 allele dose correlates positively with TC levels [310–312] and the *APOE* locus is one of the top hits for TC GWAS [226], it has been suggested that part of the *APOE* effect on AD is dependent on its effect on TC levels [126]. One could thus hypothesize that APOE contributes to an increase in TC levels that infer a higher AD risk. This increase can be counteracted by *HMGCR* SNPs

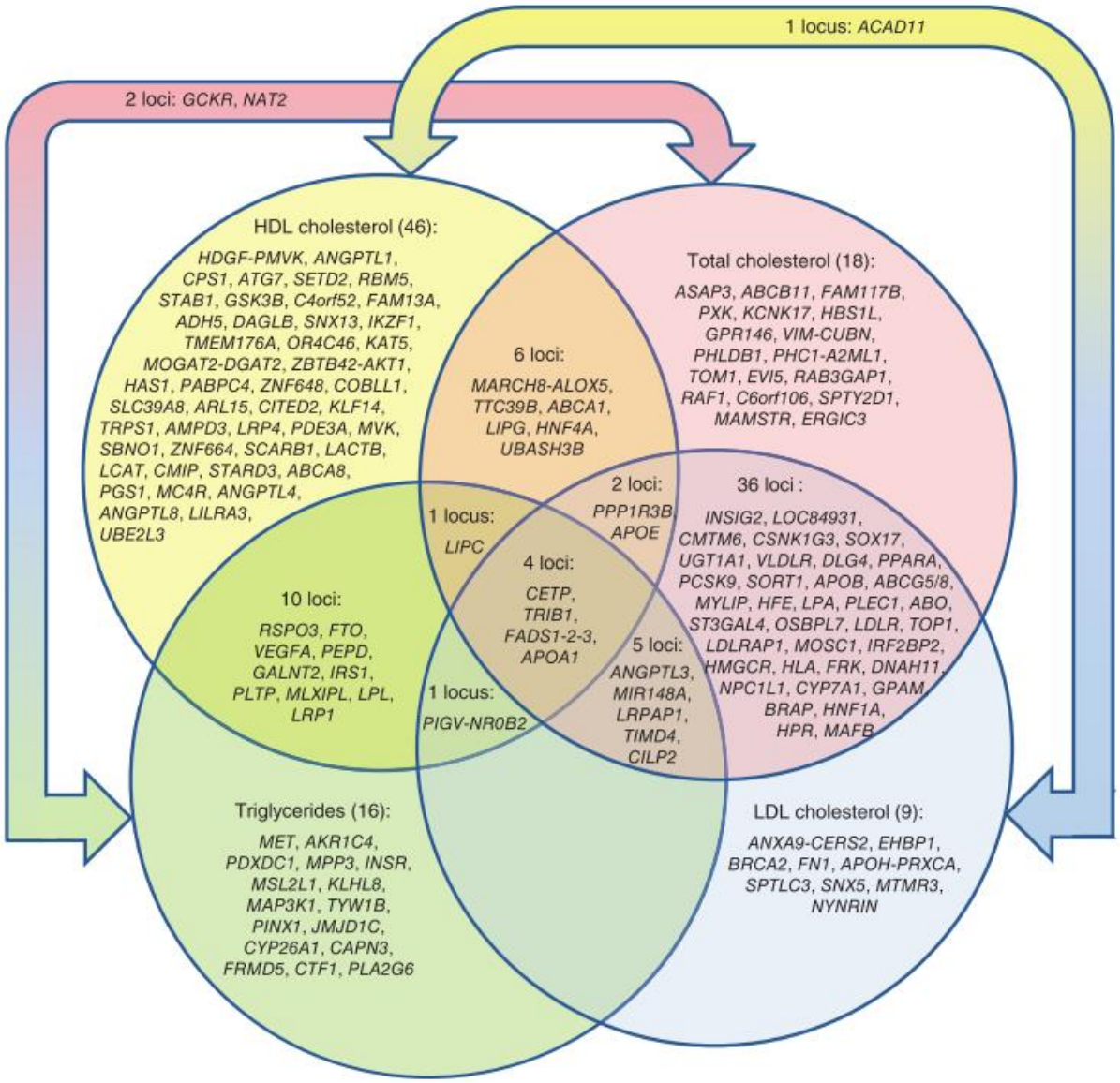
that influence cholesterol synthesis by regulating HMGCR (**Figure 5.1**, middle panel). The fact that we do not see an effect in *APOE*-ε4 non-carriers could be explained by that they already have low TC levels by virtue of not being an ε4 carrier and any further reduction is non-consequential.

Interactions with *APOE*-ε4 could also potentially explain, at least in part, the discrepancy between Chapter 2 and 4; in Chapter 4, we chose to investigate interactions with statin use and sex instead of *APOE*-ε4. This was done because initial exploratory analysis revealed statin use and sex as the two major factors influencing TC levels and due to sample sizes, analyses was limited to investigating the statin use * sex * TC-PGS interaction. Interestingly, research suggest that the AD risk associated with TC levels are dependent on *APOE*-ε4 status; cholesterol levels associate with increased risk of AD or impaired cognition in ε4 non-carriers and not in ε4 carriers [266,310,313,314]. Thus future studies investigating individuals with a low vs high TC-PGS in ε4 carriers and non-carriers, could potentially reveal a similar pattern as in Chapter 2, such that low TC-PGS in ε4 carriers would be protective, and high TC-PGS would associate with increased risk in ε4 non-carriers.

SNPs mediating effect through other lipid parameters: Both rs3846662 and rs72633963 associates not only with TC levels but with LDL-C levels [226]. While TC captures cholesterol present in all lipoprotein fractions (e.g. HDL and LDL), the cholesterol associated with the different fractions have different implications when it comes to cardiovascular disease [167] and the same might be true for AD. For example, high LDL-C but low HDL-C associates with increased risk of cardiovascular disease [167]. In AD, similar to TC levels, increased LDL-C levels have been shown to associate with increased AD risk whereas the relationship between HDL-C and AD is less clear [183,315,316]. In fact, in a recent meta-analysis, LDL-C but not TC correlated with AD [315]. Interestingly, we found that rs72633963 A allele, similar to rs3846662 A allele (described in [226]), associates with reduced levels of LDL-C but does not associate with HDL-C. This suggests that the effect could be mediated through TC *or,* more precisely LDL-C, but not HDL-C levels.

If indeed the effect is mediated through LDL-C rather than TC levels, this would have implications for the TC-PGS in Chapter 4. The genetics underlying the different lipid traits are somewhat overlapping, but each lipid trait also has its own specific loci associated (**Figure 5.2**) [226]. Thus, in Chapter 4 we investigated genetics underlying TC levels, but it could also be that genetics underlying

Figure 5.2 Lipid traits and their associated genetic loci



Venn diagram of the genetic overlap between the different lipid traits. Figure reproduced with permission from [226].

each of the specific traits, or genetics that are common for traits (e.g. SNPs involved in both TC and LDL-C), would correlate with AD. In fact, Proitsi et al., [213], constructed polygenic scores of genome-wide significant SNPs for each of the cholesterol traits (TC, LDL-C, HDL-C and TG), and for each trait created one trait-specific score (only including the loci exclusively associating with the trait) and one full score (including all loci associated with the trait). These analyses showed that while the full scores better predicted their respective traits, only the HDL-C trait-specific score associated (positively) with AD. The limitations of this study is discussed in Chapter 4 and thus while the LDL-C score did not associate with AD, it is possible that by allowing more SNPs in the polygenic score, considering pathophysiologically proved AD, and interactions with factors such as sex, age, statin use, and *APOE*-ε4, associations would become apparent.

5.2.3 Central effect

HMGCR metabolism: There is also the possibility that the effects of rs72633963 and rs3846662 on AD is mediated by central mechanisms. We found rs72633963 to associate with frontal cortex levels of total *HMGCR* and to interact with *APOE*-ε4 status; the A allele associated with increased *HMGCR* levels in ε4 non-carriers and had no effect in ε4 carriers. These findings suggest that rs72633963 does have an effect centrally on HMGCR metabolism, but the absence of effect in ε4 carriers does thus not support mediation of increased Aβ pathology through altered RNA levels of *HMGCR* in this group. However, a central effect cannot be ruled out, since our findings on rs72633963 only concerned *HMGCR* RNA levels. It is well known that HMGCR is extensively regulated by feedback loops and full blockage of HMGCR activity (e.g. by statins) can increase RNA expression eightfold [165]. Thus, a more extensive characterization of HMGCR metabolism with regards to rs72633963 in brain tissue and cell types, investigating outcomes such as HMGCR activity, would be more informative of a central effect.

In Chapter 3, one of our objectives was to determine the effect of rs3846662 AA genotype across multiple cell types, particularly in the neuronal related cell types. Any such cell type specific effect

could shed light on central effects. Unfortunately, due to differences in growth rate and potential to differentiate between cell lines, we were unable to differentiate cells from AA carriers to neurons (discussed in Chapter 3) and were underpowered to detect any potential cell type * rs3846662 interaction effects. Nevertheless, HMGCR protein levels were assessed in all cell types and while the interaction did not allow for post-hoc testing, the means suggest a diminishing effect of rs3846662 across cell types: iPSC > REC > NPC. One could thus speculate if this implies a diminished effect in central cell types such as NPCs, a finding that would be supported by literature showing that rs3846662 A allele associates with increased Δ13-*HMGCR* in liver but not in brain tissue [33,241]. However, these are premature conclusions considering the drawbacks of this study and would have to be further investigated using proper sized groups and different cell types, as well as investigating more important indices such as HMGCR activity or cholesterol levels.

Cholesterol *vs* non-sterol isoprenoids: HMGCR catalyzes the conversion of HMG-CoA to mevalonate, which in turn is a precursor for cholesterol, but also other non-sterol isoprenoids and it has been argued that the protective effects of statins can be mediated through reductions in cholesterol and/or non-sterol isoprenoids. Indeed, in AD brains, levels of the non-sterol isoprenoids farnesyl and geranylgeranyl pyrophosphates are increased compared to control brains, whereas there are no difference in cholesterol levels [317,318]. Levels of the enzymes responsible for their synthesis was further shown to associate with increased levels of TAU pathology (p-TAU levels and NFT density) [318] supporting a role for the non-sterol isoprenoids in the brain.

In vitro studies have shown that full blockage of HMGCR activity in rat neuronal cultures increases TAU phosphorylation, destabilizes microtubules and ultimately decreases cell viability [209,210], and that these phenotypes can be rescued by supplementation of non-sterol isoprenoids but not cholesterol [209]. On the other hand, cholesterol specific reductions have been shown to decrease Aβ production [30,31]. These findings would support a role for non-sterol isoprenoids in TAU pathology, but opposite

to what has been found in the human brain. Cholesterol levels on the other hand, seem to have an effect on Aβ pathology.

The relationship between these two pathways are complex, and importantly, a partial reduction of HMGCR activity does not affect the production of the two groups of end products equally. In the case of low mevalonate production, the production of non-sterol isoprenoids are prioritized over cholesterol production [165]. Consequently, since rs72633963 and rs3846662 likely has a partial influence on HMGCR activity, the effect on the non-sterol isoprenoids might arguably be less than their effect on cholesterol. Nevertheless, the effect on non-sterol isoprenoids were not a subject for this thesis, and any cholesterol and non-sterol isoprenoid specific effects would have to be evaluated.

5.3 Limitations

5.3.1 Ancestry/population stratification

Cohorts used throughout this thesis are mainly of European descent and results should thus only be interpreted in the context of European ancestry. This decision was made based on the data that was available, with the cohorts used in this study being mainly of European descent. The most diverse cohort that we used was ADNI where still 741 out of 795 individuals were deemed to be of European descent. Due to the small sample size of the non-European population, investigations of effects in this population would likely not yield robust results and were therefore excluded.

Within the field of genomic studies, the effect of ancestry, or population stratification, on gene-phenotype associations is widely acknowledged; heterogenous population structures can lead to both false positives and reduced power to detect true positives [319]. The importance of population structure is evident also in AD research, where multiple loci primarily have been identified in a Caucasian population and then been shown to associate differently with disease in other populations such as African Americans and Asian. The effects of *APOE* [70], *ABCA7* [320], *CLU* [278] and *PICALM* [279] amongst other loci, have been found to differ in such a way depending on ancestry.

The associations between TC levels and AD have been demonstrated for multiple ancestries, (e.g. Asian [321] and African American [314]) and the GWA study meta-analysis on TC levels used in Chapter 4 included multiple ancestry groups [226]. Thus, one could argue for an inclusion of all individuals regardless of ancestry. As mentioned above, effect of loci on AD seems to differ depending on ancestry, and it would thus follow that even though the TC loci are derived from a multiethnic cohort their relationship to AD might differ. While we did not find an effect of the TC-PGS in Europeans, it could still be of value in individuals of other ancestry. The lack of diversity would have to be addressed in follow up studies.

In this thesis, we have used four different cohorts; PREVENT-AD, ext-QFP, ADNI and ROSMAP which were all evaluated for ancestry by mapping their genetic data to the 1000 Genomes data set [236]. By doing so, we found that even though mapping to, and selecting for, European ancestry, the substructure can still be highly diverse (e.g. compare ext-QFP and ADNI in **Supplemental sFigure 2.1** in Appendix 2). Although the cohorts were not mapped together, from their individual plots with the 1000 Genomes, one can see that ADNI consists mainly of individuals with Central European/British ancestry whereas PREVENT-AD does not map to either of the 1000 Genomes subpopulations (Northern and Western European, Toscani in Italy, Finnish in Finland, British in England and Scotland and Iberian population in Spain, data not shown). This is in line with literature showing that even within populations such as Europeans, genetic substructures are still very much apparent [322] and can influence the associations between genetic loci (e.g. *APOE* [323] and *ACE* [324]) and AD risk. Viewed in this light, it is possible that the associations of rs72633963 A allele from Chapter 2 is specific to individuals with specifically French ancestry, but the lack of an actual French cohort, leaves this to be investigated.

In Chapter 4, we addressed population stratification by filtering all cohorts for European ancestry and further substructures were handled by correcting for the top genetic PCs (as proposed by [259]). In the correlations between the TC-PGS and cholesterol levels, we did observe that correcting for PCs in

PREVENT-AD and ADNI helped bring the results closer to each other (data not shown), indicating that there indeed is an effect of substructures that were addressed to some extent by including the PCs in the model.

5.3.2 Clinical and pathological heterogeneity

In our initial study of *HMGCR* SNPs we found different effects of rs72633963 on AD risk between the ext-QFP and ADNI cohorts; in ADNI the A allele associated with increased risk in *APOE*-ε4 carriers, whereas we saw an indication for protection in ext-QFP (Chapter 2). We argued that one of the reasons for the discrepancy could be due to differences in the certainty of the AD diagnoses. ADNI, being on ongoing longitudinal study, diagnoses individuals clinically, and thus with possible or probable AD, whereas the ext-QFP cohort contains a mix of clinical and pathologically defined cases. Indeed, it has been shown that using pathophysiologically proven AD as outcome in a GWAS, renders different effect sizes for some SNPs compared to their effect size associated with clinically defined AD [325].

Using the AT(N) classification system, we used CSF biomarkers to refine the control (Aβ(-), cognitively healthy) and AD (Aβ(+), clinical AD diagnosis) groups in ADNI. Although we still saw an indication for increased AD risk with the rs72633963 A allele, the association was not significant. These findings thus lend support to the idea that the heterogeneity associated with clinical diagnoses can influence SNP-AD associations and could be of importance for the identification of low effect loci.

In addition to refining clinical diagnoses, presence of pathologies other than Aβ, TAU and neurodegeneration may influence the results. For example, Lesser et al., [188] found that correlations between TC levels and neuritic plaque density became stronger when excluding AD cases with non-AD pathologies (mainly vascular) thus comparing control to "pure AD". Similarly, they found a linear positive relationship between TC levels and certainty of AD diagnosis (CERAD possible, probable and definite AD) [179]. Considering the marked co-occurrence of other pathologies in AD [326], especially

vascular alterations such as white matter hyperintensities, vascular amyloid or micro-bleeds, these studies suggest that we might have to look further than the core pathologies and that groups need to be further refined with regards to these pathologies. For example, classifying individuals according to AT(N) with a vascular pathology component (e.g. AT(N)-V) could help tease apart relationships between rs72633963 or TC-PGS and AD.

5.3.3 Interactions with other factors

TC genetics: Cholesterol metabolism, both peripheral and central, is highly dynamic and interacts with numerous other factors that could all contribute to confounding results. Factors such as age, sex, menopause, physical activity, smoking, medication amongst others, all affect cholesterol metabolism as evidenced by their effect on peripheral blood lipids (discussed in Chapter 4). In the studies presented here, we have considered ancestry, age, *APOE*-ε4 status, statin use, and sex. However, the importance of other factors is apparent in the literature, and it has further been shown that they may interact with genetics to influence lipid profiles. By considering interactions with smoking status [327] and physical activity [328] multiple new loci were identified for associating with lipid traits, and for some loci, effects were only detectable when considering the interaction. While we have explored an *APOE*-ε4 * SNP interaction (Chapter 2) and a statin use * sex * TC-PGS interaction (Chapter 4), it is clear that the genetics underlying cholesterol metabolism is far more complicated than that and future studies would have to address the specific question of these interactions.

Vascular burden: As discussed in Chapter 4, we also hypothesized that there might be an additive vascular burden effect, such that the effect of the TC-PGS would only be apparent in individuals affected by other cardiovascular risk factors. Findings form rs72633963 and rs3846662 would lend some support to that notion with effects occurring only in *APOE*-ε4 carriers and not in ε4 non-carriers [13,34].

5.4 Future directions

Results from GWA studies suggest that genome-wide significant gene loci associating with AD are involved in lipid metabolism [7]. *APOE*-ε4 as the most impactful factor for both TC levels and AD risk strongly suggest a genetic link between the two. In this thesis, we have investigated variants apart from *APOE* and tested two specific hypotheses; SNPs that we believe affect cholesterol synthesis through *HMGCR* specifically (Chapter 2 and 3) and SNPs that affect peripheral TC levels (Chapter 4).

Brain cell type heterogeneity: Although speculative, our findings on the rs72633963 and rs3846662 suggest that they might act by altering peripheral levels of either TC or LDL-C. The lack of association between rs3846662 and HMGCR metabolism in the brain reported in the literature [33,241], were argued to potentially be obscured by the heterogeneity of the tissue; the brain homogenates contain diverse cell types with different roles such as neurons, astrocytes, microglia, oligodendrocytes and cells associated with the vascular system. This hypothesis could be tested by using the stem cells developed in this thesis to create different peripheral (e.g. liver cells) and central (e.g. neurons, astrocytes, microglia, oligodendrocytes) cells and investigate HMGCR activity and cholesterol levels. To detect effects on AD processes however, sample sizes would need to be increased.

Cell heterogeneity of the brain tissue could further be addressed by single-cell RNA sequencing in which nuclei are isolated from brain tissue and separated based on their expression profiles [329]. Exploring data from one such experiment using mouse brain tissue, expression of *HMGCR* is apparent in multiple cell types; e.g. neurons, oligodendrocytes and astrocytes [330]. It would also seem that expression is highest in astrocytes, supporting the fact that cholesterol synthesis mainly occurs in glia in the adult brain. Further, using this approach in combination with genetic data on rs3846662, possible cell type specific effects could be identified.

Causality of SNPs: Another aspect lacking in this thesis is the causality of identified SNPs and would be a future endeavor. To this end, isogenic cell lines developed from either of the iPSC lines developed

here could shed light on the causality of rs3846662 or rs72633963 on at least HMGCR metabolism and cholesterol.

Lipid polygenic scores: Even though the TC-PGS investigated in this thesis did not result in significant correlations with AD, others have investigated a lipid related polygenic score in AD and found significant correlations. Darst et al. [212] constructed scores based on AD genome-wide significant SNPs after assigning them to a GO pathway. Three scores was constructed; one containing genes pertaining to immune response, one with genes involved in Aβ clearance and finally one with genes involved in cholesterol metabolism, and correlated with cognition and biomarkers of Aβ, TAU and neurodegeneration. The cholesterol score was the only score that significantly associated with CSF Aβ-42/Aβ-40 ratio independently of *APOE*. Of note, only two loci were included (*ABCA7* and *CLU*) in the score and the SNPs in the score were weighted for their association with AD. Thus, that score and the TC-PGS constructed here is very different, since our TC-PGS are weighted for associations with TC levels rather than AD. In addition, while *ABCA7* and *CLU* are cholesterol related genes, their role in the TC-PGS constructed here is unclear.

As discussed above, rs72633963 and rs3846662 also associate with LDL-C levels, raising the possibility that genetics underlying LDL-C could be of more importance than the ones underlying TC. In addition, Proitsi et al., [213] found an association between an HDL-C trait-specific score. Thus, creating trait specific, optimized polygenic scores for these traits could shed light on the genetic overlap between peripheral blood cholesterol metabolism and AD.

Taken together these findings suggest that the process of selecting SNPs for including in the score, and the weighting of the score, influence the outcome on AD while also highlighting the nature of polygenic scores. Our TC-PGS, by design, captures variance in peripheral TC levels specifically. As such, any associations between the TC-PGS would be through its effect on TC levels. To assess central effects of TC related genetics, one would have to use a different approach for selecting and weighting SNPs. One such approach could be investigating SNPs in cholesterol related genes (e.g. genes

associated with lipid metabolism GO terms identified in AD GWA studies [7]) weighted either by their effect on brain gene expression or by their effect on AD.

5.5 Closing remarks

With this thesis we aimed to unravel some of the missing heritability in AD by investigating cholesterol related variants in AD. We have provided evidence for the involvement of HMGCR SNPs rs72633963 and rs3846662 in cholesterol metabolism and provide a hypothesis for how they also influence AD (**Figure 5.1**). As both rs72633963 and rs3846662 associate with peripheral TC levels, and increased midlife TC levels is a risk factor for AD, we further hypothesized that variants linked with TC levels would associate with AD. To this end we created a TC-PGS and while the TC-PGS strongly associated with TC levels and was successful in improving the prediction of hypercholesterolemia, we could not detect any significant effect on AD risk or pathology.

Although contradictory, the results presented here contribute to the literature of the involvement of cholesterol related genetics in AD; while we could not prove an association with AD risk, rs72633963 might still contribute to AD by influencing Aβ pathology in populations of French ancestry, and genetics underlying peripheral TC levels do not seem to have an effect on AD in Europeans.

References

1. Alzheimer's Association. 2019 Alzheimer's disease facts and figures. *Alzheimer's & Dementia*. 2019;15(3):321-387. doi:10.1016/j.jalz.2019.01.010

2. Gatz M, Reynolds CA, Fratiglioni L, et al. Role of Genes and Environments for Explaining Alzheimer Disease. *Arch Gen Psychiatry*. 2006;63(2):168-174. doi:10.1001/archpsyc.63.2.168

3. Ridge PG, Hoyt KB, Boehme K, et al. Assessment of the genetic variance of late-onset Alzheimer's disease. *Neurobiology of Aging*. 2016;41:200.e13-200.e20. doi:10.1016/j.neurobiolaging.2016.02.024

4. Lee SH, Yang J, Goddard ME, Visscher PM, Wray NR. Estimation of pleiotropy between complex diseases using single-nucleotide polymorphism-derived genomic relationships and restricted maximum likelihood. *Bioinformatics*. 2012;28(19):2540-2542. doi:10.1093/bioinformatics/bts474

5. Mormino EC, Sperling RA, Holmes AJ, et al. Polygenic risk of Alzheimer disease is associated with early- and late-life processes. *Neurology*. 2016;87(5):481. doi:10.1212/WNL.0000000000002922

6. Escott-Price V, Myers AJ, Huentelman M, Hardy J. Polygenic risk score analysis of pathologically confirmed Alzheimer disease. *Ann Neurol*. 2017;82(2):311-314. doi:10.1002/ana.24999

7. Kunkle BW, Grenier-Boley B, Sims R, et al. Genetic meta-analysis of diagnosed Alzheimer's disease identifies new risk loci and implicates Aβ, tau, immunity and lipid processing. *Nature Genetics*. 2019;51(3):414-430. doi:10.1038/s41588-019-0358-2

8. Jansen IE, Savage JE, Watanabe K, et al. Genome-wide meta-analysis identifies new loci and functional pathways influencing Alzheimer's disease risk. *Nat Genet*. 2019;51(3):404-413. doi:10.1038/s41588-018-0311-9

9. Ebbert MTW, Ridge PG, Kauwe JSK. Bridging the gap between statistical and biological epistasis in Alzheimer's disease. *Biomed Res Int*. 2015;2015:870123. doi:10.1155/2015/870123

10. Jun G, Naj AC, Beecham GW, et al. Meta-analysis confirms CR1, CLU, and PICALM as alzheimer disease risk loci and reveals interactions with APOE genotypes. *Arch Neurol.* 2010;67(12):1473-1484. doi:10.1001/archneurol.2010.201

11. Santos-Rebouças CB, Gonçalves AP, dos Santos JM, et al. rs3851179 Polymorphism at 5′ to the PICALM Gene is Associated with Alzheimer and Parkinson Diseases in Brazilian Population. *NeuroMolecular Medicine.* 2017;19(2-3):293-299. doi:10.1007/s12017-017-8444-z

12. Gharesouran J, Rezazadeh M, Khorrami A, Ghojazadeh M, Talebi M. Genetic evidence for the involvement of variants at APOE, BIN1, CR1, and PICALM loci in risk of late-onset Alzheimer's disease and evaluation for interactions with APOE genotypes. *J Mol Neurosci.* 2014;54(4):780-786. doi:10.1007/s12031-014-0377-5

13. Leduc V, De Beaumont L, Théroux L, et al. HMGCR is a genetic modifier for risk, age of onset and MCI conversion to Alzheimer's disease in a three cohorts study. *Molecular Psychiatry.* 2015;20(7):867-873. doi:10.1038/mp.2014.81

14. Liao Y-C, Lee W-J, Hwang J-P, et al. ABCA7 gene and the risk of Alzheimer's disease in Han Chinese in Taiwan. *Neurobiol Aging.* 2014;35(10):2423.e7-2423.e13. doi:10.1016/j.neurobiolaging.2014.05.009

15. Keenan BT, Shulman JM, Chibnik LB, et al. A coding variant in CR1 interacts with APOE-ε4 to influence cognitive decline. *Hum Mol Genet.* 2012;21(10):2377-2388. doi:10.1093/hmg/dds054

16. Kivipelto M, Ngandu T, Fratiglioni L, et al. Obesity and Vascular Risk Factors at Midlife and the Risk of Dementia and Alzheimer Disease. *Arch Neurol.* 2005;62(10):1556-1560. doi:10.1001/archneur.62.10.1556

17. Kivipelto M, Helkala E-L, Laakso MP, et al. Apolipoprotein E ϵ4 Allele, Elevated Midlife Total Cholesterol Level, and High Midlife Systolic Blood Pressure Are Independent Risk Factors for Late-Life Alzheimer Disease. *Ann Intern Med.* 2002;137(3):149. doi:10.7326/0003-4819-137-3-200208060-00006

18. Kivipelto M, Helkala E-L, Laakso MP, et al. Midlife vascular risk factors and Alzheimer's disease in later life: longitudinal, population based study. *BMJ.* 2001;322(7300):1447–1451. doi:10.1136/bmj.322.7300.1447

19. Solomon A, Kivipelto M, Wolozin B, Zhou J, Whitmer RA. Midlife Serum Cholesterol and Increased Risk of Alzheimer's and Vascular Dementia Three Decades Later. *Dement Geriatr Cogn.* 2009;28(1):75-80. doi:10.1159/000231980

20. Toro P, Degen Ch, Pierer M, Gustafson D, Schröder J, Schönknecht P. Cholesterol in mild cognitive impairment and Alzheimer's disease in a birth cohort over 14 years. *European Archives of Psychiatry and Clinical Neuroscience*. 2014;264(6):485-492. doi:10.1007/s00406-013-0468-2

21. Notkola I-L, Sulkava R, Pekkanen J, et al. Serum Total Cholesterol, Apolipoprotein E ε4 Allele, and Alzheimer's Disease. *NED*. 1998;17(1):14-20. doi:10.1159/000026149

22. Jick H, Zornberg GL, Jick SS, Seshadri S, Drachman DA. Statins and the risk of dementia. *Lancet*. 2000;356(9242):1627-1631.

23. Wolozin B, Kellman W, Ruosseau P, Celesia G, Siegel G. Decreased prevalence of Alzheimer disease associated with 3-hydroxy-3-methyglutaryl coenzyme A reductase inhibitors. *Arch Neurol*. 2000;57(10):1439-1443. doi:10.1001/archneur.57.10.1439

24. Sparks D, Scheff S, Hunsaker J, Liu H, Landers T, Gross D. Induction of Alzheimer-like beta-amyloid immunoreactivity in the brains of rabbits with dietary cholesterol. *Exp Neurol*. 1994;126(1):88-94. doi:10.1006/exnr.1994.1044

25. Sparks D. Intraneuronal beta-amyloid immunoreactivity in the CNS. *Neurobiol Aging*. 1996;17(2):291-299.

26. Streit W, Sparks D. Activation of microglia in the brains of humans with heart disease and hypercholesterolemic rabbits. *J Mol Med*. 1997;75(2):130-138.

27. Refolo L, Pappolla M, Malester B, J L. Hypercholesterolemia accelerates the Alzheimer's amyloid pathology in a transgenic mouse model. *Neurobiology of Disease*. 2000;7(4):321-331. doi:10.1006/nbdi.2000.0304

28. Shie F-S, Jin L-W, Cook DG, Leverenz JB, C L Renée. Diet-induced hypercholesterolemia enhances brain A accumulation in transgenic mice. *Neuroreport*. 2002;13(4):455-459.

29. Umeda T, Tomiyama T, Kitajima E, et al. Hypercholesterolemia accelerates intraneuronal accumulation of A oligomers resulting in memory impairment in Alzheimer's disease model mice. *Life Sci*. 2012;91(23):1169-1176. doi:10.1016/j.lfs.2011.12.022

30. Fassbender K, Simons M, Bergmann C, et al. Simvastatin strongly reduces levels of Alzheimer's disease β-amyloid peptides Aβ42 and Aβ40 in vitro and in vivo. *Proc Natl Acad Sci USA*. 2001;98(10):5856. doi:10.1073/pnas.081620098

31. Simons M, Keller P, De Strooper B, Beyreuther K, Dotti CG, Simons K. Cholesterol depletion inhibits the generation of β-amyloid in hippocampal neurons. *Proc Natl Acad Sci USA*. 1998;95(11):6460. doi:10.1073/pnas.95.11.6460

32. Burkhardt R, Kenny EE, Lowe JK, et al. Common SNPs in HMGCR in micronesians and whites associated with LDL-cholesterol levels affect alternative splicing of exon13. *Arterioscler Thromb Vasc Biol*. 2008;28(11):2078-2084. doi:10.1161/ATVBAHA.108.172288

33. Simmons CR, Zou F, Younkin SG, Estus S. Evaluation of the global association between cholesterol-associated polymorphisms and Alzheimer's disease suggests a role for rs3846662 and HMGCR splicing in disease risk. *Mol Neurodegener*. 2011;6(1):62. doi:10.1186/1750-1326-6-62

34. Wright SM, Jensen SL, Cockriel KL, et al. Association study of rs3846662 with Alzheimer's disease in a population-based cohort: the Cache County Study. *Neurobiology of Aging*. 2019;84:242.e1-242.e6. doi:10.1016/j.neurobiolaging.2019.03.004

35. Keller L, Murphy C, Wang H-XX, et al. A functional polymorphism in the HMGCR promoter affects transcriptional activity but not the risk for Alzheimer disease in Swedish populations. *Brain Res*. 2010;1344:185-191. doi:10.1016/j.brainres.2010.04.073

36. Porcellini E, Calabrese E, Guerini F, et al. The hydroxy-methyl-glutaryl CoA reductase promoter polymorphism is associated with Alzheimer's risk and cognitive deterioration. *Neurosci Lett*. 2007;416(1):66-70. doi:10.1016/j.neulet.2007.01.046

37. Johnson JM, Castle J, Garrett-Engele P, et al. Genome-wide survey of human alternative pre-mRNA splicing with exon junction microarrays. *Science*. 2003;302(5653):2141-2144. doi:10.1126/science.1090100

38. Medina MW, Krauss RM. The role of HMGCR alternative splicing in statin efficacy. *Trends Cardiovasc Med*. 2009;19(5):173-177. doi:10.1016/j.tcm.2009.10.003

39. Mangialasche F, Solomon A, Winblad B, Mecocci P, Kivipelto M. Alzheimer's disease: clinical trials and drug development. *The Lancet Neurology*. 2010;9(7):702-716. doi:10.1016/S1474-4422(10)70119-8

40. Liu P-P, Xie Y, Meng X-Y, Kang J-S. History and progress of hypotheses and clinical trials for Alzheimer's disease. *Signal Transduction and Targeted Therapy*. 2019;4(1):1-22. doi:10.1038/s41392-019-0063-8

41. McKhann GM, Knopman DS, Chertkow H, et al. The diagnosis of dementia due to Alzheimer's disease: Recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease. *Alzheimers Dement.* 2011;7(3):263-269. doi:10.1016/j.jalz.2011.03.005

42. Scheltens P, Blennow K, Breteler MMB, et al. Alzheimer's disease. *The Lancet.* 2016;388(10043):505-517. doi:10.1016/S0140-6736(15)01124-1

43. WHO website. Dementia. Accessed August 19, 2020. https://www.who.int/news-room/fact-sheets/detail/dementia

44. Hebert LE, Weuve J, Scherr PA, Evans DA. Alzheimer disease in the United States (2010-2050) estimated using the 2010 census. *Neurology.* 2013;80(19):1778-1783. doi:10.1212/WNL.0b013e31828726f5

45. Robinson M, Lee BY, Hane FT. Recent Progress in Alzheimer's Disease Research, Part 2: Genetics and Epidemiology. *J Alzheimers Dis.* 2017;57(2):317-330. doi:10.3233/JAD-161149

46. Folstein MF, Folstein SE, McHugh PR. "Mini-mental state". A practical method for grading the cognitive state of patients for the clinician. *J Psychiatr Res.* 1975;12(3):189-198. doi:10.1016/0022-3956(75)90026-6

47. Nasreddine ZS, Phillips NA, Bédirian V, et al. The Montreal Cognitive Assessment, MoCA: a brief screening tool for mild cognitive impairment. *J Am Geriatr Soc.* 2005;53(4):695-699. doi:10.1111/j.1532-5415.2005.53221.x

48. Hughes CP, Berg L, Danziger WL, Coben LA, Martin RL. A new clinical scale for the staging of dementia. *Br J Psychiatry.* 1982;140:566-572. doi:10.1192/bjp.140.6.566

49. Hyman BT, Trojanowski JQ. Consensus recommendations for the postmortem diagnosis of Alzheimer disease from the National Institute on Aging and the Reagan Institute Working Group on diagnostic criteria for the neuropathological assessment of Alzheimer disease. *J Neuropathol Exp Neurol.* 1997;56(10):1095-1097. doi:10.1097/00005072-199710000-00002

50. Alzheimer A. Uber eine eigenartige Erkrankung der Hirnrinde. *Zentralbl Nervenh Psych.* 1907;18:177–179.

51. Fischer O. Miliare Nekrosen mit drusigen Wucherungen der Neurofibrillen, eine regelmässige Veränderung der Hirnrinde bei seniler Demenz. *Monatsschr Psychiatr Neurol.* 1907;22:361–372.

52. Goedert M. Oskar Fischer and the study of dementia. *Brain.* 2008;132(4):1102-1111. doi:10.1093/brain/awn256

53. Toodayan N. Professor Alois Alzheimer (1864–1915): Lest we forget. *Journal of Clinical Neuroscience*. 2016;31:47-55. doi:10.1016/j.jocn.2015.12.032

54. Blessed G, Tomlinson BE, Roth M. The Association Between Quantitative Measures of Dementia and of Senile Change in the Cerebral Grey Matter of Elderly Subjects. *The British Journal of Psychiatry*. 1968;114(512):797-811. doi:10.1192/bjp.114.512.797

55. McKhann G, Drachman D, Folstein M, Katzman R, Price D, Stadlan EM. Clinical diagnosis of Alzheimer's disease: *Neurology*. 1984;34:939-944.

56. Sperling RA, Aisen PS, Beckett LA, et al. Toward defining the preclinical stages of Alzheimer's disease: Recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease. *Alzheimer's & Dementia*. 2011;7(3):280–292. doi:10.1016/j.jalz.2011.03.003

57. Albert MS, DeKosky ST, Dickson D, et al. The diagnosis of mild cognitive impairment due to Alzheimer's disease: Recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease. *Alzheimer's & Dementia*. 2011;7(3):270-279. doi:10.1016/j.jalz.2011.03.008

58. Jack CR, Bennett DA, Blennow K, et al. NIA-AA Research Framework: Toward a biological definition of Alzheimer's disease. *Alzheimer's & Dementia: The Journal of the Alzheimer's Association*. 2018;14(4):535-562. doi:10.1016/j.jalz.2018.02.018

59. Jack CR, Bennett DA, Blennow K, et al. A/T/N: An unbiased descriptive classification scheme for Alzheimer disease biomarkers. *Neurology*. 2016;87(5):539-547. doi:10.1212/WNL.0000000000002923

60. Iturria-Medina Y, Sotero RC, Toussaint PJ, Mateos-Pérez JM, Evans AC, Alzheimer's Disease Neuroimaging Initiative. Early role of vascular dysregulation on late-onset Alzheimer's disease based on multifactorial data-driven analysis. *Nat Commun*. 2016;7:11934. doi:10.1038/ncomms11934

61. Karch CM, Goate AM. Alzheimer's Disease Risk Genes and Mechanisms of Disease Pathogenesis. *Biological Psychiatry*. 2015;77(1):43-51. doi:10.1016/j.biopsych.2014.05.006

62. Liao X, Cai F, Sun Z, et al. Identification of Alzheimer's disease-associated rare coding variants in the ECE2 gene. *JCI Insight*. 2020;5(4). doi:10.1172/jci.insight.135119

63. Korvatska O, Leverenz JB, Jayadev S, et al. R47H Variant of TREM2 Associated With Alzheimer Disease in a Large Late-Onset Family: Clinical, Genetic, and Neuropathological Study. *JAMA Neurol*. 2015;72(8):920-927. doi:10.1001/jamaneurol.2015.0979

64. Kawas C, Gray S, Brookmeyer R, Fozard J, Zonderman A. Age-specific incidence rates of Alzheimer's disease: The Baltimore Longitudinal Study of Aging. *Neurology*. 2000;54(11):2072-2077. doi:10.1212/WNL.54.11.2072

65. Toepper M. Dissociating Normal Aging from Alzheimer's Disease: A View from Cognitive Neuroscience. *J Alzheimers Dis*. 2017;57(2):331-352. doi:10.3233/JAD-161099

66. Lautenschlager NT, Cupples LA, Rao VS, et al. Risk of dementia among relatives of Alzheimer's disease patients in the MIRAGE study: What is in store for the oldest old? *Neurology*. 1996;46(3):641-650. doi:10.1212/wnl.46.3.641

67. Poirier J, Davignon J, Bouthillier D, et al. Apolipoprotein E polymorphism and Alzheimer's disease. *The Lancet*. 1993;342(8873):697-699. doi:10.1016/0140-6736(93)91705-Q

68. Corder EH, Saunders AM, Strittmatter WJ, et al. Gene dose of apolipoprotein E type 4 allele and the risk of Alzheimer's disease in late onset families. *Science*. 1993;261(5123):921-923. doi:10.1126/science.8346443

69. Saunders AM, Strittmatter WJ, Schmechel D, et al. Association of apolipoprotein E allele epsilon 4 with late-onset familial and sporadic Alzheimer's disease. *Neurology*. 1993;43(8):1467-1472. doi:10.1212/wnl.43.8.1467

70. Farrer LA, Cupples LA, Haines JL, et al. Effects of Age, Sex, and Ethnicity on the Association Between Apolipoprotein E Genotype and Alzheimer Disease: A Meta-analysis. *JAMA*. 1997;278(16):1349-1356. doi:10.1001/jama.1997.03550160069041

71. Lambert J-C, Ibrahim-Verbaas CA, Harold D, et al. Meta-analysis of 74,046 individuals identifies 11 new susceptibility loci for Alzheimer's disease. *Nature Genetics*. 2013;45(12):1452-1458. doi:10.1038/ng.2802

72. Marioni RE, Harris SE, Zhang Q, et al. GWAS on family history of Alzheimer's disease. *Translational Psychiatry*. 2018;8(1):99. doi:10.1038/s41398-018-0150-6

73. Harold D, Abraham R, Hollingworth P, et al. Genome-wide association study identifies variants at CLU and PICALM associated with Alzheimer's disease. *Nat Genet*. 2009;41(10):1088-1093. doi:10.1038/ng.440

74. Sando SB, Melquist S, Cannon A, et al. Risk-reducing effect of education in Alzheimer's disease. *International Journal of Geriatric Psychiatry*. 2008;23(11):1156-1162. doi:10.1002/gps.2043

75. McDowell I, Xi G, Lindsay J, Tierney M. Mapping the connections between education and dementia. *Journal of Clinical and Experimental Neuropsychology*. 2007;29(2):127-141. doi:10.1080/13803390600582420

76. Xu W, Tan L, Wang H-FF, et al. Meta-analysis of modifiable risk factors for Alzheimer's disease. *J Neurol Neurosurg Psychiatr*. 2015;86(12):1299-1306. doi:10.1136/jnnp-2015-310548

77. Wang H-X, Karp A, Winblad B, Fratiglioni L. Late-life engagement in social and leisure activities is associated with a decreased risk of dementia: a longitudinal study from the Kungsholmen project. *Am J Epidemiol*. 2002;155(12):1081-1087. doi:10.1093/aje/155.12.1081

78. Whitmer R, Gustafson D, E B-C, Haan M, Gunderson E, Yaffe K. Central obesity and increased risk of dementia more than three decades later. *Neurology*. 2008;71(14):1057-1064. doi:10.1212/01.wnl.0000306313.89165.ef

79. Guerreiro R, Brás J, Hardy J. SnapShot: Genetics of Alzheimer's Disease. *Cell*. 2013;155(4):968-968.e1. doi:10.1016/j.cell.2013.10.037

80. Shobab LA, Hsiung G-YR, Feldman HH. Cholesterol in Alzheimer's disease. *The Lancet Neurology*. 2005;4(12):841-852. doi:10.1016/S1474-4422(05)70248-9

81. Wood W, Li L, Müller WE, Eckert GP. Cholesterol as a causative factor in Alzheimer's disease: a debatable hypothesis. *J Neurochem*. 2014;129(4):559-572. doi:10.1111/jnc.12637

82. Canada PHA of. A Dementia Strategy for Canada: Together We Aspire. aem. Published August 19, 2019. Accessed August 19, 2020. https://www.canada.ca/en/public-health/services/publications/diseases-conditions/dementia-strategy.html#s3.2

83. Duyckaerts C, Delatour B, Potier M-C. Classification and basic pathology of Alzheimer disease. *Acta Neuropathol*. 2009;118(1):5-36. doi:10.1007/s00401-009-0532-1

84. Wilquet V, Strooper BD. Amyloid-beta precursor protein processing in neurodegeneration. *Current Opinion in Neurobiology*. 2004;14(5):582-588. doi:10.1016/j.conb.2004.08.001

85. O'Brien RJ, Wong PC. Amyloid Precursor Protein Processing and Alzheimer's Disease. *Annu Rev Neurosci*. 2011;34:185-204. doi:10.1146/annurev-neuro-061010-113613

86. Kang J, Lemaire HG, Unterbeck A, et al. The precursor of Alzheimer's disease amyloid A4 protein resembles a cell-surface receptor. *Nature*. 1987;325(6106):733-736. doi:10.1038/325733a0

87. Masters CL, Simms G, Weinman NA, Multhaup G, McDonald BL, Beyreuther K. Amyloid plaque core protein in Alzheimer disease and Down syndrome. *Proceedings of the National Academy of Sciences*. 1985;82(12):4245-4249. doi:10.1073/pnas.82.12.4245

88. Mirra SS, Heyman A, McKeel D, et al. The Consortium to Establish a Registry for Alzheimer's Disease (CERAD). Part II. Standardization of the neuropathologic assessment of Alzheimer's disease. *Neurology*. 1991;41(4):479-486. doi:10.1212/wnl.41.4.479

89. Thal DR, Rüb U, Orantes M, Braak H. Phases of A beta-deposition in the human brain and its relevance for the development of AD. *Neurology*. 2002;58(12):1791-1800. doi:10.1212/wnl.58.12.1791

90. Alzheimer A. über eigenartige Krankheitsfälle des späteren Alters. *Z f d g Neur u Psych*. 1911;4(1):356. doi:10.1007/BF02866241

91. Ittner A, Ittner LM. Dendritic Tau in Alzheimer's Disease. *Neuron*. 2018;99(1):13-27. doi:10.1016/j.neuron.2018.06.003

92. Goedert M, Crowther RA, Garner CC. Molecular characterization of microtubule-associated proteins tau and map2. *Trends in Neurosciences*. 1991;14(5):193-199. doi:10.1016/0166-2236(91)90105-4

93. Ballatore C, Lee VM-Y, Trojanowski JQ. Tau-mediated neurodegeneration in Alzheimer's disease and related disorders. *Nature Reviews Neuroscience*. 2007;8(9):663-672. doi:10.1038/nrn2194

94. Braak H, Braak E. Neuropathological stageing of Alzheimer-related changes. *Acta Neuropathologica*. 1991;82(4):239-259. doi:10.1007/BF00308809

95. Jouanne M, Rault S, Voisin-Chiret A-S. Tau protein aggregation in Alzheimer's disease: An attractive target for the development of novel therapeutic agents. *Eur J Med Chem*. 2017;139:153-167. doi:10.1016/j.ejmech.2017.07.070

96. Gómez-Isla T, Price JL, Jr DWM, Morris JC, Growdon JH, Hyman BT. Profound Loss of Layer II Entorhinal Cortex Neurons Occurs in Very Mild Alzheimer's Disease. *J Neurosci*. 1996;16(14):4491-4500. doi:10.1523/JNEUROSCI.16-14-04491.1996

97. Hyman BT, Phelps CH, Beach TG, et al. National Institute on Aging–Alzheimer's Association guidelines for the neuropathologic assessment of Alzheimer's disease. *Alzheimer's & Dementia*. 2012;8(1):1-13. doi:10.1016/j.jalz.2011.10.007

98. Blennow K, Hampel H. CSF markers for incipient Alzheimer's disease. *Lancet Neurol*. 2003;2(10):605-613. doi:10.1016/s1474-4422(03)00530-1

99. Seppälä TT, Nerg O, Koivisto AM, et al. CSF biomarkers for Alzheimer disease correlate with cortical brain biopsy findings. *Neurology*. 2012;78(20):1568-1575. doi:10.1212/WNL.0b013e3182563bd0

100. Fagan AM, Mintun MA, Mach RH, et al. Inverse relation between in vivo amyloid imaging load and cerebrospinal fluid Abeta42 in humans. *Ann Neurol*. 2006;59(3):512-519. doi:10.1002/ana.20730

101. Jack CR, Barrio JR, Kepe V. Cerebral amyloid PET imaging in Alzheimer's disease. *Acta Neuropathol*. 2013;126(5):643-657. doi:10.1007/s00401-013-1185-7

102. Lowe VJ, Lundt ES, Albertson SM, et al. Neuroimaging correlates with neuropathologic schemes in neurodegenerative disease. *Alzheimers Dement*. 2019;15(7):927-939. doi:10.1016/j.jalz.2019.03.016

103. Clark CM, Pontecorvo MJ, Beach TG, et al. Cerebral PET with florbetapir compared with neuropathology at autopsy for detection of neuritic amyloid-β plaques: a prospective cohort study. *The Lancet Neurology*. 2012;11(8):669-678. doi:10.1016/S1474-4422(12)70142-4

104. Hansson O, Seibyl J, Stomrud E, et al. CSF biomarkers of Alzheimer's disease concord with amyloid-β PET and predict clinical progression: A study of fully automated immunoassays in BioFINDER and ADNI cohorts. *Alzheimer's & Dementia*. 2018;14(11):1470-1481. doi:10.1016/j.jalz.2018.01.010

105. Blennow K, Hampel H, Weiner M, Zetterberg H. Cerebrospinal fluid and plasma biomarkers in Alzheimer disease. *Nat Rev Neurol*. 2010;6(3):131-144. doi:10.1038/nrneurol.2010.4

106. Schöll M, Lockhart SN, Schonhaut DR, et al. PET Imaging of Tau Deposition in the Aging Human Brain. *Neuron*. 2016;89(5):971-982. doi:10.1016/j.neuron.2016.01.028

107. Fleisher AS, Pontecorvo MJ, Devous MD, et al. Positron Emission Tomography Imaging With [18F]flortaucipir and Postmortem Assessment of Alzheimer Disease Neuropathologic Changes. *JAMA Neurol*. 2020;77(7):829-839. doi:10.1001/jamaneurol.2020.0528

108. Landau SM, Harvey D, Madison CM, et al. Associations between cognitive, functional, and FDG-PET measures of decline in AD and MCI. *Neurobiol Aging*. 2011;32(7):1207-1218. doi:10.1016/j.neurobiolaging.2009.07.002

109. Fox NC, Crum WR, Scahill RI, Stevens JM, Janssen JC, Rossor MN. Imaging of onset and progression of Alzheimer's disease with voxel-compression mapping of serial magnetic resonance images. *The Lancet*. 2001;358(9277):201-205. doi:10.1016/S0140-6736(01)05408-3

110. Olsson B, Lautner R, Andreasson U, et al. CSF and blood biomarkers for the diagnosis of Alzheimer's disease: a systematic review and meta-analysis. *Lancet Neurol*. 2016;15(7):673-684. doi:10.1016/S1474-4422(16)00070-3

111. Hansson O, Zetterberg H, Buchhave P, Londos E, Blennow K, Minthon L. Association between CSF biomarkers and incipient Alzheimer's disease in patients with mild cognitive impairment: a follow-up study. *Lancet Neurol*. 2006;5(3):228-234. doi:10.1016/S1474-4422(06)70355-6

112. Ossenkoppele R, Jansen WJ, Rabinovici GD, et al. Prevalence of Amyloid PET Positivity in Dementia Syndromes: A Meta-analysis. *JAMA*. 2015;313(19):1939-1950. doi:10.1001/jama.2015.4669

113. Jack CR, Knopman DS, Jagust WJ, et al. Hypothetical model of dynamic biomarkers of the Alzheimer's pathological cascade. *The Lancet Neurology*. 2010;9(1):119-128. doi:10.1016/S1474-4422(09)70299-6

114. Jack CR, Knopman DS, Jagust WJ, et al. Tracking pathophysiological processes in Alzheimer's disease: an updated hypothetical model of dynamic biomarkers. *The Lancet Neurology*. 2013;12(2):207-216. doi:10.1016/S1474-4422(12)70291-0

115. Bateman RJ, Xiong C, Benzinger TLS, et al. Clinical and Biomarker Changes in Dominantly Inherited Alzheimer's Disease. *New England Journal of Medicine*. 2012;367(9):795-804. doi:10.1056/NEJMoa1202753

116. Saunders A, Strittmatter W, Schmechel D, et al. Association of apolipoprotein E allele epsilon 4 with late-onset familial and sporadic Alzheimer's disease. *Neurology*. 1993;43(8):1467-1472.

117. Strittmatter W, Saunders A. Apolipoprotein E: high-avidity binding to beta-amyloid and increased frequency of type 4 allele in late-onset familial Alzheimer disease. *Proc Natl Acad Sci USA*. Published online 1993. doi:10.1073/pnas.90.5.1977

118. Shen L, Jia J. An Overview of Genome-Wide Association Studies in Alzheimer's Disease. *Neurosci Bull*. 2016;32(2):183-190. doi:10.1007/s12264-016-0011-3

119. Alz. Ass. 2018 Alzheimer's disease facts and figures. *Alzheimer's & Dementia*. 2018;14(3):367-429. doi:10.1016/j.jalz.2018.02.001

120. Mahley RW. Apolipoprotein E: cholesterol transport protein with expanding role in cell biology. *Science*. 1988;240(4852):622-630. doi:10.1126/science.3283935

121. Corder EH, Saunders AM, Risch NJ, et al. Protective effect of apolipoprotein E type 2 allele for late onset Alzheimer disease. *Nat Genet*. 1994;7(2):180-184. doi:10.1038/ng0694-180

122. Deelen J, Evans DS, Arking DE, et al. A meta-analysis of genome-wide association studies identifies multiple longevity genes. *Nat Commun*. 2019;10(1):3669. doi:10.1038/s41467-019-11558-2

123. Kuo C-L, Pilling LC, Atkins JL, Kuchel GA, Melzer D. ApoE e2 and aging-related outcomes in 379,000 UK Biobank participants. *Aging (Albany NY)*. 2020;12(12):12222-12233. doi:10.18632/aging.103405

124. Altmann A, Tian L, Henderson VW, Greicius MD, Alzheimer's Disease Neuroimaging Initiative Investigators. Sex modifies the APOE-related risk of developing Alzheimer disease. *Ann Neurol*. 2014;75(4):563-573. doi:10.1002/ana.24135

125. Tang M-X, Stern Y, Marder K, et al. The APOE-∈4 Allele and the Risk of Alzheimer Disease Among African Americans, Whites, and Hispanics. *JAMA*. 1998;279(10):751-755. doi:10.1001/jama.279.10.751

126. Jarvik G, Wijsman E, Kukull W, Schellenberg G, Yu C, Larson E. Interactions of apolipoprotein E genotype, total cholesterol level, age, and sex in prediction of Alzheimer's disease: a case-control study. *Neurology*. 1995;45(6):1092-1096.

127. Reiman EM, Chen K, Alexander GE, et al. Functional brain abnormalities in young adults at genetic risk for late-onset Alzheimer's dementia. *Proc Natl Acad Sci USA*. 2004;101(1):284-289. doi:10.1073/pnas.2635903100

128. Reiman EM, Caselli RJ, Yun LS, et al. Preclinical evidence of Alzheimer's disease in persons homozygous for the epsilon 4 allele for apolipoprotein E. *N Engl J Med*. 1996;334(12):752-758. doi:10.1056/NEJM199603213341202

129. Sunderland T, Mirza N, Putnam KT, et al. Cerebrospinal fluid β-amyloid1–42 and tau in control subjects at risk for Alzheimer's disease: The effect of APOE ε4 allele. *Biological Psychiatry*. 2004;56(9):670-676. doi:10.1016/j.biopsych.2004.07.021

130. Reiman EM, Chen K, Liu X, et al. Fibrillar amyloid-beta burden in cognitively normal people at 3 levels of genetic risk for Alzheimer's disease. *Proc Natl Acad Sci USA*. 2009;106(16):6820-6825. doi:10.1073/pnas.0900345106

131. Morris JC, Roe CM, Xiong C, et al. APOE predicts amyloid-beta but not tau Alzheimer pathology in cognitively normal aging. *Ann Neurol*. 2010;67(1):122-131. doi:10.1002/ana.21843

132. Beffert U, Poirier J. Apolipoprotein E, plaques, tangles and cholinergic dysfunction in Alzheimer's disease. *AnnNYAcadSci*. 1996;777:166-174.

133. Tiraboschi P, Hansen LA, Masliah E, Alford M, Thal LJ, Corey-Bloom J. Impact of APOE genotype on neuropathologic and neurochemical markers of Alzheimer disease. *Neurology*. 2004;62(11):1977-1983. doi:10.1212/01.wnl.0000128091.92139.0f

134. Leduc V, Stéphanie J-B, Poirier J. APOE and cholesterol homeostasis in Alzheimer's disease. *Trends Mol Med*. 2010;16(10):469-477. doi:10.1016/j.molmed.2010.07.008

135. Poirier J. Apolipoprotein E, cholesterol transport and synthesis in sporadic Alzheimer's disease. *Neurobiol Aging*. 2005;26(3):355-361. doi:10.1016/j.neurobiolaging.2004.09.003

136. Zhao N, Liu C-C, Qiao W, Bu G. Apolipoprotein E, Receptors, and Modulation of Alzheimer's Disease. *Biological Psychiatry*. 2018;83(4):347-357. doi:10.1016/j.biopsych.2017.03.003

137. Namba Y, Tomonaga M, Kawasaki H, Otomo E, Ikeda K. Apolipoprotein E immunoreactivity in cerebral amyloid deposits and neurofibrillary tangles in Alzheimer's disease and kuru plaque amyloid in Creutzfeldt-Jakob disease. *Brain Res*. 1991;541(1):163-166. doi:10.1016/0006-8993(91)91092-f

138. Bales KR, Verina T, Cummins DJ, et al. Apolipoprotein E is essential for amyloid deposition in the APP(V717F) transgenic mouse model of Alzheimer's disease. *ProcNatlAcadSciUSA*. 1999;96(26):15233-15238.

139. Holtzman DM, Bales KR, Tenkova T, et al. Apolipoprotein E isoform-dependent amyloid deposition and neuritic degeneration in a mouse model of Alzheimer's disease. *Proc Natl Acad Sci USA*. 2000;97(6):2892-2897. doi:10.1073/pnas.050004797

140. Dodart J-C, Marr RA, Koistinaho M, et al. Gene delivery of human apolipoprotein E alters brain Abeta burden in a mouse model of Alzheimer's disease. *Proc Natl Acad Sci USA*. 2005;102(4):1211-1216. doi:10.1073/pnas.0409072102

141. Bales KR, Liu F, Wu S, et al. Human APOE isoform-dependent effects on brain beta-amyloid levels in PDAPP transgenic mice. *J Neurosci*. 2009;29(21):6771-6779. doi:10.1523/JNEUROSCI.0887-09.2009

142. Castellano JM, Kim J, Stewart FR, et al. Human apoE isoforms differentially regulate brain amyloid-β peptide clearance. *Sci Transl Med*. 2011;3(89):89ra57. doi:10.1126/scitranslmed.3002156

143. Brecht WJ, Harris FM, Chang S, et al. Neuron-specific apolipoprotein e4 proteolysis is associated with increased tau phosphorylation in brains of transgenic mice. *J Neurosci*. 2004;24(10):2527-2534. doi:10.1523/JNEUROSCI.4315-03.2004

144. Hollingworth P, Harold D, Sims R, et al. Common variants at ABCA7, MS4A6A/MS4A4E, EPHA1, CD33 and CD2AP are associated with Alzheimer's disease. *Nat Genet*. 2011;43(5):429-435. doi:10.1038/ng.803

145. Naj AC, Jun G, Beecham GW, et al. Common variants at *MS4A4/MS4A6E*, *CD2AP*, *CD33* and *EPHA1* are associated with late-onset Alzheimer's disease. *Nature Genetics*. 2011;43(5):436-441. doi:10.1038/ng.801

146. Choi SW, Mak TSH, O'Reilly P. A guide to performing Polygenic Risk Score analyses. *bioRxiv*. Published online January 1, 2018. doi:10.1101/416545

147. Desikan RS, Fan CC, Wang Y, et al. Genetic assessment of age-associated Alzheimer disease risk: Development and validation of a polygenic hazard score. Brayne C, ed. *PLOS Medicine*. 2017;14(3):e1002258. doi:10.1371/journal.pmed.1002258

148. Tosto G, Bird TD, Tsuang D, et al. Polygenic risk scores in familial Alzheimer disease. *Neurology*. 2017;88(12):1180-1186. doi:10.1212/WNL.0000000000003734

149. Tasaki S, Gaiteri C, Mostafavi S, De Jager PL, Bennett DA. The Molecular and Neuropathological Consequences of Genetic Risk for Alzheimer's Dementia. *Front Neurosci*. 2018;12. doi:10.3389/fnins.2018.00699

150. Adams HHH, de Bruijn RFAG, Hofman A, et al. Genetic risk of neurodegenerative diseases is associated with mild cognitive impairment and conversion to dementia. *Alzheimers Dement*. 2015;11(11):1277-1285. doi:10.1016/j.jalz.2014.12.008

151. Marden JR, Mayeda ER, Walter S, et al. Using an Alzheimer Disease Polygenic Risk Score to Predict Memory Decline in Black and White Americans Over 14 Years of Follow-up. *Alzheimer Dis Assoc Disord*. 2016;30(3):195-202. doi:10.1097/WAD.0000000000000137

152. Lupton MK, Strike L, Hansell NK, et al. The effect of increased genetic risk for Alzheimer's disease on hippocampal and amygdala volume. *Neurobiol Aging*. 2016;40:68-77. doi:10.1016/j.neurobiolaging.2015.12.023

153. Ridge PG, Mukherjee S, Crane PK, Kauwe JSK, Consortium ADG. Alzheimer's Disease: Analyzing the Missing Heritability. *PLOS ONE*. 2013;8(11):e79771. doi:10.1371/journal.pone.0079771

154. Jonsson T, Atwal JK, Steinberg S, et al. A mutation in APP protects against Alzheimer's disease and age-related cognitive decline. *Nature*. 2012;488(7409):96-99. doi:10.1038/nature11283

155. Jonsson T, Stefansson H, Steinberg S, et al. Variant of TREM2 associated with the risk of Alzheimer's disease. *N Engl J Med*. 2013;368(2):107-116. doi:10.1056/NEJMoa1211103

156. Wingo TS, Cutler DJ, Wingo AP, et al. Association of Early-Onset Alzheimer Disease With Elevated Low-Density Lipoprotein Cholesterol Levels and Rare Genetic Coding Variants of APOB. *JAMA Neurol*. 2019;76(7):809-817. doi:10.1001/jamaneurol.2019.0648

157. Barral S, Bird T, Goate A, et al. Genotype patterns at PICALM, CR1, BIN1, CLU, and APOE genes are associated with episodic memory. *Neurology*. 2012;78(19):1464-1471. doi:10.1212/WNL.0b013e3182553c48

158. Combarros O, van Duijn CM, Hammond N, et al. Replication by the Epistasis Project of the interaction between the genes for IL-6 and IL-10 in the risk of Alzheimer's disease. *Journal of Neuroinflammation*. 2009;6(1):22. doi:10.1186/1742-2094-6-22

159. Lehmann DJ, Schuur M, Warden DR, et al. Transferrin and HFE genes interact in Alzheimer's disease risk: the Epistasis Project. *Neurobiol Aging*. 2012;33(1):202.e1-13. doi:10.1016/j.neurobiolaging.2010.07.018

160. Kölsch H, Lehmann DJ, Ibrahim-Verbaas CA, et al. Interaction of insulin and PPAR-α genes in Alzheimer's disease: the Epistasis Project. *Journal of Neural Transmission*. 2012;119(4):473-479. doi:10.1007/s00702-011-0732-4

161. Medway C, Combarros O, Cortina-Borja M, et al. The sex-specific associations of the aromatase gene with Alzheimer's disease and its interaction with IL10 in the Epistasis Project. *Eur J Hum Genet.* 2014;22(2):216-220. doi:10.1038/ejhg.2013.116

162. Combarros O, Warden DR, Hammond N, et al. The dopamine β-hydroxylase -1021C/T polymorphism is associated with the risk of Alzheimer's disease in the Epistasis Project. *BMC Med Genet.* 2010;11:162. doi:10.1186/1471-2350-11-162

163. Benedet AL, Labbe A, Lemay P, et al. Epistasis analysis links immune cascades and cerebral amyloidosis. *J Neuroinflammation.* 2015;12:227. doi:10.1186/s12974-015-0436-z

164. Zhang J, Liu Q. Cholesterol metabolism and homeostasis in the brain. *Protein Cell.* 2015;6(4):254-264. doi:10.1007/s13238-014-0131-3

165. Goldstein J, Brown M. Regulation of the mevalonate pathway. *Nature.* 1990;343(6257):425-430. doi:10.1038/343425a0

166. Picard C, Julien C, Frappier J, et al. Alterations in cholesterol metabolism-related genes in sporadic Alzheimer's disease. *Neurobiol Aging.* 2018;66:180.e1-180.e9. doi:10.1016/j.neurobiolaging.2018.01.018

167. Goldstein JL, Brown MS. A century of cholesterol and coronaries: from plaques to genes to statins. *Cell.* 2015;161(1):161-172. doi:10.1016/j.cell.2015.01.036

168. Tan ZS, Seshadri S, Beiser A, et al. Plasma Total Cholesterol Level as a Risk Factor for Alzheimer Disease: The Framingham Study. *Arch Intern Med.* 2003;163(9):1053-1057. doi:10.1001/archinte.163.9.1053

169. Mielke MM, Zandi PP, Shao H, et al. The 32-year relationship between cholesterol and dementia from midlife to late life. *Neurology.* 2010;75(21):1888-1895. doi:10.1212/WNL.0b013e3181feb2bf

170. Rönnemaa E, Zethelius B, Lannfelt L, Kilander L. Vascular Risk Factors and Dementia: 40-Year Follow-Up of a Population-Based Cohort. *DEM.* 2011;31(6):460-466. doi:10.1159/000330020

171. Shepherd J, Blauw GJ, Murphy MB, et al. Pravastatin in elderly individuals at risk of vascular disease (PROSPER): a randomised controlled trial. *The Lancet.* 2002;360(9346):1623-1630. doi:10.1016/S0140-6736(02)11600-X

172. Heart Protection Study Collaborative Group. MRC/BHF Heart Protection Study of cholesterol lowering with simvastatin in 20 536 high-risk individuals: a randomised placebocontrolled trial. *The Lancet*. 2002;360(9326):7-22. doi:10.1016/S0140-6736(02)09327-3

173. Sano M, Bell K, Galasko D, et al. A randomized, double-blind, placebo-controlled trial of simvastatin to treat Alzheimer disease. *Neurology*. 2011;77(6):556-563. doi:10.1212/WNL.0b013e318228bf11

174. Lesser G, Kandiah K, Libow LS, et al. Elevated Serum Total and LDL Cholesterol in Very Old Patients with Alzheimer's Disease. *DEM*. 2001;12(2):138-145. doi:10.1159/000051248

175. Romas SN, Tang M-X, Berglund L, Mayeux R. APOE genotype, plasma lipids, lipoproteins, and AD in community elderly. *Neurology*. 1999;53(3):517-517. doi:10.1212/WNL.53.3.517

176. Reitz C, Tang M-X, Luchsinger J, Mayeux R. Relation of plasma lipids to Alzheimer disease and vascular dementia. *Arch Neurol*. 2004;61(5):705-714. doi:10.1001/archneur.61.5.705

177. Li G, Shofer JB, Kukull WA, et al. Serum cholesterol and risk of Alzheimer disease: a community-based cohort study. *Neurology*. 2005;65(7):1045-1050. doi:10.1212/01.wnl.0000178989.87072.11

178. Schilling S, Tzourio C, Soumaré A, et al. Differential associations of plasma lipids with incident dementia and dementia subtypes in the 3C Study: A longitudinal, population-based prospective cohort study. *PLOS Medicine*. 2017;14(3):e1002265. doi:10.1371/journal.pmed.1002265

179. Lesser GT, Haroutunian V, Purohit DP, et al. Serum Lipids Are Related to Alzheimer's Pathology in Nursing Home Residents. *DEM*. 2009;27(1):42-49. doi:10.1159/000189268

180. Kuo YM, Emmerling MR, Bisgaier CL, et al. Elevated low-density lipoprotein in Alzheimer's disease correlates with brain abeta 1-42 levels. *Biochem Biophys Res Commun*. 1998;252(3):711-715. doi:10.1006/bbrc.1998.9652

181. Lepara O, Valjevac A, Alajbegović A, Zaćiragić A, Nakas-Ićindić E. Decreased serum lipids in patients with probable Alzheimer's disease. *Bosn J Basic Med Sci*. 2009;9(3):215-220. doi:10.17305/bjbms.2009.2809

182. Anstey KJ, Lipnicki DM, Low L-F. Cholesterol as a risk factor for dementia and cognitive decline: a systematic review of prospective studies with meta-analysis. *Am J Geriatr Psychiatry*. 2008;16(5):343-354. doi:10.1097/JGP.0b013e31816b72d4

183. Anstey KJ, Ashby-Mitchell K, Peters R. Updating the Evidence on the Association between Serum Cholesterol and Risk of Late-Life Dementia: Review and Meta-Analysis. *J Alzheimers Dis.* 2017;56(1):215-228. doi:10.3233/JAD-160826

184. Pappolla MA, Bryant-Thomas TK, Herbert D, et al. Mild hypercholesterolemia is an early risk factor for the development of Alzheimer amyloid pathology. *Neurology.* 2003;61(2):199. doi:10.1212/01.WNL.0000070182.02537.84

185. Köbe T, Gonneaud J, Binette AP, et al. Association of Vascular Risk Factors With β-Amyloid Peptide and Tau Burdens in Cognitively Unimpaired Individuals and Its Interaction With Vascular Medication Use. *JAMA Netw Open.* 2020;3(2):e1920780-e1920780. doi:10.1001/jamanetworkopen.2019.20780

186. Reed B, Villeneuve S, Mack W, DeCarli C, Chui HC, Jagust W. Associations Between Serum Cholesterol Levels and Cerebral Amyloidosis. *JAMA Neurol.* 2014;71(2):195-200. doi:10.1001/jamaneurol.2013.5390

187. Okamura N, Arai H, Maruyama M, et al. Serum Cholesterol and Cerebrospinal Fluid Amyloid β Protein in Alzheimer'S Disease. *Journal of the American Geriatrics Society.* 2001;49(12):1738-1739. doi:10.1046/j.1532-5415.2001.49293.x

188. Lesser GT, Beeri MS, Schmeidler J, Purohit DP, Haroutunian V. Cholesterol and LDL Relate to Neuritic Plaques and to APOE4 Presence but Not to Neurofibrillary Tangles. *Current Alzheimer Research.* 2011;8(3):303-312.

189. Reiman EM, Chen K, Langbaum JBS, et al. Higher serum total cholesterol levels in late middle age are associated with glucose hypometabolism in brain regions affected by Alzheimer's disease and normal aging. *NeuroImage.* 2010;49(1):169-176. doi:10.1016/j.neuroimage.2009.07.025

190. Wu C-W, Liao P-C, Lin C, et al. Brain region-dependent increases in ?-amyloid and apolipoprotein E levels in hypercholesterolemic rabbits. *J Neural Transm.* 2003;110(6):641-649. doi:10.1007/s00702-002-0809-1

191. Kálmán J, Kudchodkar BJ, Krishnamoorthy R, Dory L, Lacko AG, Agarwal N. High cholesterol diet down regulates the activity of activator protein-1 but not nuclear factor-kappa B in rabbit brain. *Life Sci.* 2001;68(13):1495-1503. doi:10.1016/s0024-3205(01)00955-9

192. Sparks D. The early and ongoing experience with the cholesterol-fed rabbit as a model of Alzheimer's disease: the old, the new and the pilot. *J Alzheimers Dis.* 2008;15(4):641-656.

193. Sparks D, Liu H, Gross D, Scheff S. Increased density of cortical apolipoprotein E immunoreactive neurons in rabbit brain after dietary administration of cholesterol. *Neurosci Lett.* 1995;187(2):142-144.

194. Deci S, Lemieux S, Carrie S-B, Sparks D, Schreurs B. Cholesterol increases ventricular volume in a rabbit model of Alzheimer's disease. *J Alzheimer's Dis Jad.* 2012;29(2):283-292. doi:10.3233/jad-2011-111415

195. Schreurs BG, Smith-Bell CA, Lemieux SK. Dietary cholesterol increases ventricular volume and narrows cerebrovascular diameter in a rabbit model of Alzheimer's disease. *Neuroscience.* 2013;254:61-69. doi:10.1016/j.neuroscience.2013.09.015

196. Schreurs B, Wang D, Carrie S-B, Burhans L, Bell R, Jimena G-J. Dietary Cholesterol Concentration and Duration Degrade Long-Term Memory of Classical Conditioning of the Rabbit's Nictitating Membrane Response. *Int J Alzheimer's Dis.* 2012;2012:732634. doi:10.1155/2012/732634

197. Ronald J, Chen Y, Bernas L, et al. Clinical field-strength MRI of amyloid plaques induced by low-level cholesterol feeding in rabbits. *Brain.* 2009;132(5):1346-1354. doi:10.1093/brain/awp031

198. Belinson H, Michaelson DM. Pathological synergism between amyloid-beta and apolipoprotein E4--the most prevalent yet understudied genetic risk factor for Alzheimer's disease. *J Alzheimers Dis.* 2009;17(3):469-481. doi:10.3233/JAD-2009-1065

199. Hamanaka H, Y K-F, Suzuki K, et al. Altered cholesterol metabolism in human apolipoprotein E4 knock-in mice. *Hum Mol Genet.* 2000;9(3):353-361.

200. Knouff C, Hinsdale ME, Mezdour H, et al. Apo E structure determines VLDL clearance and atherosclerosis risk in mice. *J Clin Invest.* 1999;103(11):1579-1586. doi:10.1172/JCI6172

201. Sullivan PM, Mezdour H, Aratani Y, et al. Targeted replacement of the mouse apolipoprotein E gene with the common human APOE3 allele enhances diet-induced hypercholesterolemia and atherosclerosis. *J Biol Chem.* 1997;272(29):17972-17980. doi:10.1074/jbc.272.29.17972

202. Kariv-Inbal Z, Yacobson S, Berkecz R, et al. The isoform-specific pathological effects of apoE4 in vivo are prevented by a fish oil (DHA) diet and are modified by cholesterol. *J Alzheimers Dis.* 2012;28(3):667-683. doi:10.3233/JAD-2011-111265

203. Cole S, Vassar R. Isoprenoids and Alzheimer's disease: a complex relationship. *Neurobiology of disease.* 2006;22(2):209-222. doi:10.1016/j.nbd.2005.11.007

204. Ostrowski SM, Wilkinson BL, Golde TE, Landreth G. Statins reduce amyloid-beta production through inhibition of protein isoprenylation. *J Biol Chem.* 2007;282(37):26832-26844. doi:10.1074/jbc.M702640200

205. Cole SL, Grudzien A, Manhart IO, Kelly BL, Oakley H, Vassar R. Statins cause intracellular accumulation of amyloid precursor protein, beta-secretase-cleaved fragments, and amyloid beta-peptide via an isoprenoid-dependent mechanism. *J Biol Chem.* 2005;280(19):18755-18770. doi:10.1074/jbc.M413895200

206. Pedrini S, Carter TL, Prendergast G, Petanceska S, Ehrlich ME, Gandy S. Modulation of statin-activated shedding of Alzheimer APP ectodomain by ROCK. *PLoS medicine*. 2005;2(1):e18. doi:10.1371/journal.pmed.0020018

207. Zhou Y, Suram A, Venugopal C, et al. Geranylgeranyl pyrophosphate stimulates gamma-secretase to increase the generation of Abeta and APP-CTFgamma. *Faseb J.* 2008;22(1):47-54. doi:10.1096/fj.07-8175com

208. Pooler AM, Xi SC, Wurtman RJ. The 3-hydroxy-3-methylglutaryl co-enzyme A reductase inhibitor pravastatin enhances neurite outgrowth in hippocampal neurons. *Journal of neurochemistry*. 2006;97(3):716-723. doi:10.1111/j.1471-4159.2006.03763.x

209. Meske V, Albert F, Richter D, Schwarze J, Ohm T. Blockade of HMG-CoA reductase activity causes changes in microtubule-stabilizing protein tau via suppression of geranylgeranylpyrophosphate formation: implications for Alzheimer's disease. *Eur J Neurosci.* 2003;17(1):93-102.

210. Fan QW, Yu W, Senda T, Yanagisawa K, Michikawa M. Cholesterol-dependent modulation of tau phosphorylation in cultured neurons. *J Neurochem.* 2001;76(2):391-400. doi:10.1046/j.1471-4159.2001.00063.x

211. Poirier J, Miron J, Picard C, et al. Apolipoprotein E and lipid homeostasis in the etiology and treatment of sporadic Alzheimer's disease. *Neurobiology of aging*. 2014;35 Suppl 2:S3-10. doi:10.1016/j.neurobiolaging.2014.03.037

212. Darst BF, Koscik RL, Racine AM, et al. Pathway-specific polygenic risk scores as predictors of β-amyloid deposition and cognitive function in a sample at increased risk for Alzheimer's disease. *J Alzheimers Dis.* 2017;55(2):473-484. doi:10.3233/JAD-160195

213. Proitsi P, Lupton MK, Velayudhan L, et al. Genetic Predisposition to Increased Blood Cholesterol and Triglyceride Lipid Levels and Risk of Alzheimer Disease: A Mendelian

Randomization Analysis. George DS, ed. *PLoS Medicine*. 2014;11(9):e1001713. doi:10.1371/journal.pmed.1001713

214. Wollmer MA, Sleegers K, Ingelsson M, et al. Association study of cholesterol-related genes in Alzheimer's disease. *Neurogenetics*. 2007;8(3):179-188. doi:10.1007/s10048-007-0087-z

215. Kölsch H, Lütjohann D, Ludwig M, et al. Polymorphism in the cholesterol 24S-hydroxylase gene is associated with Alzheimer's disease. *Mol Psychiatry*. 2002;7(8):899-902. doi:10.1038/sj.mp.4001109

216. Papassotiropoulos A, Streffer JR, Tsolaki M, et al. Increased brain beta-amyloid load, phosphorylated tau, and risk of Alzheimer disease associated with an intronic CYP46 polymorphism. *Arch Neurol*. 2003;60(1):29-35. doi:10.1001/archneur.60.1.29

217. Benn M, Nordestgaard BG, Frikke-Schmidt R, Tybjærg-Hansen A. Low LDL cholesterol, PCSK9 and HMGCR genetic variation, and risk of Alzheimer's disease and Parkinson's disease: Mendelian randomisation study. *BMJ*. 2017;357. doi:10.1136/bmj.j1648

218. Wavrant-DeVrièze F, Pérez-Tur J, Lambert J-C, et al. Association between the low density lipoprotein receptor-related protein (LRP) and Alzheimer's disease. *Neuroscience Letters*. 1997;227(1):68-70. doi:10.1016/S0304-3940(97)00304-2

219. Recuero M, Vicente MCC, Ana M-G, et al. A free radical-generating system induces the cholesterol biosynthesis pathway: a role in Alzheimer's disease. *Aging Cell*. 2009;8(2):128-139. doi:10.1111/j.1474-9726.2009.00457.x

220. Patterson MC, Walkley SU. Niemann-Pick disease, type C and Roscoe Brady. *Mol Genet Metab*. 2017;120(1-2):34-37. doi:10.1016/j.ymgme.2016.11.008

221. Auer IA, Schmidt ML, Lee VM, et al. Paired helical filament tau (PHFtau) in Niemann-Pick type C disease is similar to PHFtau in Alzheimer's disease. *Acta Neuropathol*. 1995;90(6):547-551. doi:10.1007/BF00318566

222. Distl R, Treiber-Held S, Albert F, Meske V, Harzer K, Ohm TG. Cholesterol storage and tau pathology in Niemann-Pick type C disease in the brain. *J Pathol*. 2003;200(1):104-111. doi:10.1002/path.1320

223. Distl R, Meske V, Ohm TG. Tangle-bearing neurons contain more free cholesterol than adjacent tangle-free neurons. *Acta Neuropathol*. 2001;101(6):547-554. doi:10.1007/s004010000314

224. Mattsson N, Zetterberg H, Bianconi S, et al. Gamma-secretase-dependent amyloid-beta is increased in Niemann-Pick type C: a cross-sectional study. *Neurology*. 2011;76(4):366-372. doi:10.1212/WNL.0b013e318208f4ab

225. Mendoza-Oliva A, Ferrera P, Fragoso-Medina J, Arias C. Lovastatin Differentially Affects Neuronal Cholesterol and Amyloid-β Production in vivo and in vitro. *CNS Neuroscience & Therapeutics*. 2015;21(8):631-641. doi:10.1111/cns.12420

226. Willer CJ, Schmidt EM, Sengupta S, et al. Discovery and refinement of loci associated with lipid levels. *Nature Genetics*. 2013;45(11):1274-1283. doi:10.1038/ng.2797

227. Teslovich TM, Musunuru K, Smith AV, et al. Biological, clinical and population relevance of 95 loci for blood lipids. *Nature*. 2010;466(7307):707-713. doi:10.1038/nature09270

228. Breitner JCS, Poirier J, Etienne PE, Leoutsakos JM. Rationale and Structure for a New Center for Studies on Prevention of Alzheimer's Disease (StoP-AD). *J Prev Alzheimers Dis*. 2016;3(4):236-242. doi:10.14283/jpad.2016.121

229. Scriver CR. Human genetics: lessons from Quebec populations. *Annu Rev Genomics Hum Genet*. 2001;2:69-101. doi:10.1146/annurev.genom.2.1.69

230. Tremblay M, Vézina H. Genealogical Analysis of Maternal and Paternal Lineages in the Quebec Population. *Human Biology*. 2010;82(2):179-198.

231. Laberge A-M, Michaud J, Richter A, et al. Population history and its impact on medical genetics in Quebec. *Clin Genet*. 2005;68(4):287-301. doi:10.1111/j.1399-0004.2005.00497.x

232. Purcell S, Neale B, Kathe T-B, et al. PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am J Hum Genet*. 2007;81(3):559-575. doi:10.1086/519795

233. Chang CC, Chow CC, Tellier LC, Vattikuti S, Purcell SM, Lee JJ. Second-generation PLINK: rising to the challenge of larger and richer datasets. *Gigascience*. 2015;4(1). doi:10.1186/s13742-015-0047-8

234. Saykin AJ, Shen L, Yao X, et al. Genetic studies of quantitative MCI and AD phenotypes in ADNI: Progress, opportunities, and plans. *Alzheimer's & Dementia*. 2015;11(7):792-814. doi:10.1016/j.jalz.2015.05.009

235. Miron J, Picard C, Nilsson N, et al. CDK5RAP2 gene and tau pathophysiology in late-onset sporadic Alzheimer's disease. *Alzheimers Dement*. 2018;14(6):787-796. doi:10.1016/j.jalz.2017.12.004

236. The 1000 Genomes Project Consortium. A global reference for human genetic variation. *Nature*. 2015;526(7571):68-74. doi:10.1038/nature15393

237. McCarthy S, Das S, Kretzschmar W, et al. A reference panel of 64,976 haplotypes for genotype imputation. *Nat Genet*. 2016;48(10):1279-1283. doi:10.1038/ng.3643

238. Delaneau O, Marchini J, Zagury J-F. A linear complexity phasing method for thousands of genomes. *Nature Methods*. 2012;9(2):179-181. doi:10.1038/nmeth.1785

239. Durbin R. Efficient haplotype matching and storage using the positional Burrows-Wheeler transform (PBWT). *Bioinformatics*. 2014;30(9):1266-1272. doi:10.1093/bioinformatics/btu014

240. Sudmant PH, Rausch T, Gardner EJ, et al. An integrated map of structural variation in 2,504 human genomes. *Nature*. 2015;526(7571):75-81. doi:10.1038/nature15394

241. Leduc V, Théroux L, Dea D, Dufour R, Poirier J. Effects of rs3846662 Variants on HMGCR mRNA and Protein Levels and on Markers of Alzheimer's Disease Pathology. *J Mol Neurosci*. 2016;58(1):109-119. doi:10.1007/s12031-015-0666-7

242. Meyer P-F, Tremblay-Mercier J, Leoutsakos J, et al. INTREPAD. *Neurology*. 2019;92(18):e2070-e2080. doi:10.1212/WNL.0000000000007232

243. Bittner T, Zetterberg H, Teunissen CE, et al. Technical performance of a novel, fully automated electrochemiluminescence immunoassay for the quantitation of β-amyloid (1–42) in human cerebrospinal fluid. *Alzheimer's & Dementia*. 2016;12(5):517-526. doi:10.1016/j.jalz.2015.09.009

244. Shaw LM, Hansson O, Manuilova E, et al. Method comparison study of the Elecsys® β-Amyloid (1–42) CSF assay versus comparator assays and LC-MS/MS. *Clinical Biochemistry*. 2019;72:7-14. doi:10.1016/j.clinbiochem.2019.05.006

245. Randolph C, Tierney MC, Mohr E, Chase TN. The Repeatable Battery for the Assessment of Neuropsychological Status (RBANS): preliminary clinical validity. *J Clin Exp Neuropsychol*. 1998;20(3):310-319. doi:10.1076/jcen.20.3.310.823

246. Khachaturian ZS. Diagnosis of Alzheimer's Disease. *Arch Neurol*. 1985;42(11):1097-1105. doi:10.1001/archneur.1985.04060100083029

247. Leduc V, Théroux L, Dea D, Robitaille Y, Poirier J. Involvement of paraoxonase 1 genetic variants in Alzheimer's disease neuropathology. *The European journal of neuroscience*. 2009;30(9):1823-1830. doi:10.1111/j.1460-9568.2009.06983.x

248. R Core Team. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing; 2019. https://www.R-project.org/

249. Wickham H, Averick M, Bryan J, et al. Welcome to the Tidyverse. *Journal of Open Source Software*. 2019;4(43):1686. doi:10.21105/joss.01686

250. Wickham H. *Ggplot2: Elegant Graphics for Data Analysis*. Springer-Verlag New York; 2016. https://ggplot2.tidyverse.org

251. Fox J, Weisberg S. *An {R} Companion to Applied Regression*. Third. Sage; 2019. https://socialsciences.mcmaster.ca/jfox/Books/Companion/

252. Revelle W. *Psych: Procedures for Psychological, Psychometric, and Personality Research*. Northwestern University; 2019. https://CRAN.R-project.org/package=psych

253. Mangiafico S. *Rcompanion: Functions to Support Extension Education Program Evaluation.*; 2020. https://CRAN.R-project.org/package=rcompanion

254. Lenth R. *Emmeans: Estimated Marginal Means, Aka Least-Square Means.*; 2020. https://CRAN.R-project.org/package=emmeans

255. Rich J T, Neely G J, Paniello R C, Voelker C CJ, Nussenbaum B, Wang E W. A practical guide to understanding Kaplan-Meier curves. *Otolaryngol Head Neck Surg*. 2010;143(3):331-336. doi:10.1016/j.otohns.2010.05.007

256. Therneau TM. *A Package for Survival Analysis in R.*; 2020. https://CRAN.R-project.org/package=survival

257. Therneau TM, Grambsch PM. *Modeling Survival Data: Extending the Cox Model*. Springer; 2000.

258. Leduc V, Bourque L, Poirier J, Dufour R. Role of rs3846662 and HMGCR alternative splicing in statin efficacy and baseline lipid levels in familial hypercholesterolemia. *Pharmacogenet Genomics*. 2016;26(1):1-11. doi:10.1097/FPC.0000000000000178

259. Price AL, Patterson NJ, Plenge RM, Weinblatt ME, Shadick NA, Reich D. Principal components analysis corrects for stratification in genome-wide association studies. *Nat Genet*. 2006;38(8):904-909. doi:10.1038/ng1847

260. Bherer C, Labuda D, Roy-Gagnon M-H, Houde L, Tremblay M, Vézina H. Admixed ancestry and stratification of Quebec regional populations. *American Journal of Physical Anthropology*. 2011;144(3):432-441. doi:10.1002/ajpa.21424

261. Chang X-L, Tan L, Tan M-S, et al. Association of HMGCR polymorphism with late-onset Alzheimer's disease in Han Chinese. *Oncotarget*. 2016;7(16):22746-22751. doi:10.18632/oncotarget.8176

262. Burkhardt R, Kenny EE, Lowe JK, et al. Common SNPs in HMGCR in micronesians and whites associated with LDL-cholesterol levels affect alternative splicing of exon13. *Arterioscler Thromb Vasc Biol*. 2008;28(11):2078-2084. doi:10.1161/ATVBAHA.108.172288

263. Yu C-YY, Theusch E, Lo K, et al. HNRNPA1 regulates HMGCR alternative splicing and modulates cellular cholesterol metabolism. *Human molecular genetics*. 2014;23(2):319-332. doi:10.1093/hmg/ddt422

264. Chatterjee S, Krishnamoorthy P, Ranjan P, et al. Statins and cognitive function: an updated review. *Curr Cardiol Rep*. 2015;17(2):4. doi:10.1007/s11886-014-0559-3

265. Aulchenko YS, Ripatti S, Lindqvist I, et al. Loci influencing lipid levels and coronary heart disease risk in 16 European population cohorts. *Nature Genetics*. 2009;41(1):47-55. doi:10.1038/ng.269

266. Hall K, Murrell J, Ogunniyi A, et al. Cholesterol, APOE genotype, and Alzheimer disease: An epidemiologic study of Nigerian Yoruba. *Neurology*. 2006;66(2):223-227. doi:10.1212/01.wnl.0000194507.39504.17

267. Zhou T, Benda C, Dunzinger S, et al. Generation of human induced pluripotent stem cells from urine samples. *Nat Protoc*. 2012;7(12):2080-2089. doi:10.1038/nprot.2012.115

268. Takahashi K, Yamanaka S. Induction of Pluripotent Stem Cells from Mouse Embryonic and Adult Fibroblast Cultures by Defined Factors. *Cell*. 2006;126(4):663-676. doi:10.1016/j.cell.2006.07.024

269. Bell S, Peng H, Crapper L, et al. A Rapid Pipeline to Model Rare Neurodevelopmental Disorders with Simultaneous CRISPR/Cas9 Gene Editing. *STEM CELLS Translational Medicine*. 2017;6(3):886-896. doi:10.1002/sctm.16-0158

270. Schindelin J, Arganda-Carreras I, Frise E, et al. Fiji: an open-source platform for biological-image analysis. *Nature Methods*. 2012;9(7):676-682. doi:10.1038/nmeth.2019

271. Rueden CT, Schindelin J, Hiner MC, et al. ImageJ2: ImageJ for the next generation of scientific image data. *BMC Bioinformatics*. 2017;18(1):529. doi:10.1186/s12859-017-1934-z

272. Honda A, Mizokami Y, Matsuzaki Y, Ikegami T, Doy M, Miyazaki H. Highly sensitive assay of HMG-CoA reductase activity by LC-ESI-MS/MS. *J Lipid Res.* 2007;48(5):1212-1220. doi:10.1194/jlr.D600049-JLR200

273. Kuzaj P, Kuhn J, Faust I, Knabbe C, Hendig D. Measurement of HMG CoA reductase activity in different human cell lines by ultra-performance liquid chromatography tandem mass spectrometry. *Biochemical and Biophysical Research Communications.* 2014;443(2):641-645. doi:10.1016/j.bbrc.2013.12.013

274. Mertens J, Reid D, Lau S, Kim Y, Gage FH. Aging in a Dish: iPSC-Derived and Directly Induced Neurons for Studying Brain Aging and Age-Related Neurodegenerative Diseases. *Annu Rev Genet.* 2018;52:271-293. doi:10.1146/annurev-genet-120417-031534

275. Mertens J, Paquola ACM, Ku M, et al. Directly Reprogrammed Human Neurons Retain Aging-Associated Transcriptomic Signatures and Reveal Age-Related Nucleocytoplasmic Defects. *Cell Stem Cell.* 2015;17(6):705-718. doi:10.1016/j.stem.2015.09.001

276. Cao L, Wang H-F, Tan L, et al. Effect of HMGCR genetic variation on neuroimaging biomarkers in healthy, mild cognitive impairment and Alzheimer's disease cohorts. *Oncotarget.* 2016;7(12):13319-13327. doi:10.18632/oncotarget.7797

277. Zhu Z, Zheng Z, Zhang F, et al. Causal associations between risk factors and common diseases inferred from GWAS summary data. *Nature Communications.* 2018;9(1). doi:10.1038/s41467-017-02317-2

278. Han Z, Qu J, Zhao J, Zou X. Analyzing 74,248 Samples Confirms the Association Between CLU rs11136000 Polymorphism and Alzheimer's Disease in Caucasian But Not Chinese population. *Scientific Reports.* 2018;8(1). doi:10.1038/s41598-018-29450-2

279. Liu G, Zhang L, Feng R, et al. Lack of association between PICALM rs3851179 polymorphism and Alzheimer's disease in Chinese population and APOEε4-negative subgroup. *Neurobiology of Aging.* 2013;34(4):1310.e9-1310.e10. doi:10.1016/j.neurobiolaging.2012.08.015

280. Global Lipids Genetics Consortium, Willer CJ, Schmidt EM, et al. Discovery and refinement of loci associated with lipid levels. *Nature Genetics.* 2013;45(11):1274-1283. doi:10.1038/ng.2797

281. the ENGAGE Consortium, Aulchenko YS, Ripatti S, et al. Loci influencing lipid levels and coronary heart disease risk in 16 European population cohorts. *Nature Genetics.* 2009;41(1):47-55. doi:10.1038/ng.269

282. Teslovich TM, Musunuru K, Smith AV, et al. Biological, clinical and population relevance of 95 loci for blood lipids. *Nature*. 2010;466(7307):707-713. doi:10.1038/nature09270

283. Heller DA, de Faire U, Pedersen NL, Dahlen G, McClearn GE. Genetic and Environmental Influences on Serum Lipid Levels in Twins. *New England Journal of Medicine*. 1993;328(16):1150-1156. doi:10.1056/NEJM199304223281603

284. Center for Statistical Genetics. Global Lipids Genetics Consortium Results. Accessed June 20, 2018. http://csg.sph.umich.edu/willer/public/lipids2013/

285. Bennett DA, Buchman AS, Boyle PA, Barnes LL, Wilson RS, Schneider JA. Religious Orders Study and Rush Memory and Aging Project. *J Alzheimers Dis*. 2018;64(Suppl 1):S161-S189. doi:10.3233/JAD-179939

286. Weale ME. Quality Control for Genome-Wide Association Studies. In: Barnes MR, Breen G, eds. *Genetic Variation: Methods and Protocols*. Methods in Molecular Biology. Humana Press; 2010:341-372. doi:10.1007/978-1-60327-367-1_19

287. Verma SS, de Andrade M, Tromp G, et al. Imputation and quality control steps for combining multiple genome-wide datasets. *Front Genet*. 2014;5:370. doi:10.3389/fgene.2014.00370

288. Canty A, Ripley BD. *Boot: Bootstrap R (S-Plus) Functions.*; 2020.

289. Davison AC, Hinkley DV. *Bootstrap Methods and Their Applications*. Cambridge University Press; 1997. http://statwww.epfl.ch/davison/BMA/

290. Robin X, Turck N, Hainard A, et al. pROC: an open-source package for R and S+ to analyze and compare ROC curves. *BMC Bioinformatics*. 2011;12(1):77. doi:10.1186/1471-2105-12-77

291. McSweeney M, Pichet Binette A, Meyer P-F, et al. Intermediate flortaucipir uptake is associated with Aβ-PET and CSF tau in asymptomatic adults. *Neurology*. 2020;94(11):e1190-e1200. doi:10.1212/WNL.0000000000008905

292. Horikoshi M, Tang Y. *Ggfortify: Data Visualization Tools for Statistical Analysis Results.*; 2016. https://CRAN.R-project.org/package=ggfortify

293. Tang Y, Horikoshi M, Li W. ggfortify: Unified Interface to Visualize Statistical Result of Popular R Packages. *The R Journal*. Published online 2016. https://journal.r-project.org/

294. Kassambara A, Kosinski M. *Survminer: Drawing Survival Curves Using "Ggplot2."*; 2016. https://CRAN.R-project.org/package=survminer

295. Langsted A, Nordestgaard BG. Nonfasting versus fasting lipid profile for cardiovascular risk prediction. *Pathology*. 2019;51(2):131-141. doi:10.1016/j.pathol.2018.09.062

296. Grundy SM, Stone NJ, Bailey AL, et al. 2018 AHA/ACC/AACVPR/AAPA/ABC/ACPM/ADA/AGS/APhA/ASPC/NLA/PCNA Guideline on the Management of Blood Cholesterol: A Report of the American College of Cardiology/American Heart Association Task Force on Clinical Practice Guidelines. *Journal of the American College of Cardiology*. 2019;73(24):e285-e350. doi:10.1016/j.jacc.2018.11.003

297. Hoffmann TJ, Theusch E, Haldar T, et al. A large electronic-health-record-based genome-wide study of serum lipids. *Nature Genetics*. 2018;50(3):401-413. doi:10.1038/s41588-018-0064-5

298. Kanai M, Akiyama M, Takahashi A, et al. Genetic analysis of quantitative traits in the Japanese population links cell types to complex human diseases. *Nature Genetics*. 2018;50(3):390-400. doi:10.1038/s41588-018-0047-6

299. Morgan AE, Mooney KM, Wilkinson SJ, Pickles NA, Mc Auley MT. Investigating cholesterol metabolism and ageing using a systems biology approach. *Proceedings of the Nutrition Society*. 2017;76(3):378-391. doi:10.1017/S0029665116002822

300. Mielke MM, Zandi PP, Sjogren M, et al. High total cholesterol levels in late life associated with a reduced risk of dementia. *Neurology*. 2005;64(10):1689-1695. doi:10.1212/01.WNL.0000161870.78572.A5

301. Jacobs JM, Cohen A, Ein-Mor E, Stessman J. Cholesterol, Statins, and Longevity From Age 70 to 90 Years. *Journal of the American Medical Directors Association*. 2013;14(12):883-888. doi:10.1016/j.jamda.2013.08.012

302. Ambikairajah A, Walsh E, Cherbuin N. Lipid profile differences during menopause: a review with meta-analysis. *Menopause*. 2019;26(11):1327-1333. doi:10.1097/GME.0000000000001403

303. Inaraja V, Thuissard I, Andreu-Vazquez C, Jodar E. Lipid profile changes during the menopausal transition. *Menopause*. 2020;27(7):780-787. doi:10.1097/GME.0000000000001532

304. Notelovitz M. The effect of long-term oestrogen replacement therapy on glucose and lipid metabolism in postmenopausal women. *S Afr Med J*. 1976;50(50):2001-2003.

305. Kivipelto M, Ngandu T, Laatikainen T, Winblad B, Soininen H, Tuomilehto J. Risk score for the prediction of dementia risk in 20 years among middle aged people: a longitudinal,

population-based study. *The Lancet Neurology*. 2006;5(9):735-741. doi:10.1016/S1474-4422(06)70537-3

306. Exalto LG, Quesenberry CP, Barnes D, Kivipelto M, Biessels GJ, Whitmer RA. Midlife risk score for the prediction of dementia four decades later. *Alzheimers Dement.* 2014;10(5):562-570. doi:10.1016/j.jalz.2013.05.1772

307. DeCarli C, Villeneuve S, Maillard P, et al. Vascular Burden Score Impacts Cognition Independent of Amyloid PET and MRI Measures of Alzheimer's Disease and Vascular Brain Injury. *J Alzheimers Dis.* 2019;68(1):187-196. doi:10.3233/JAD-180965

308. Kamboh MI, Aston CE, Ferrell RE, Hamman RF. Impact of apolipoprotein E polymorphism in determining interindividual variation in total cholesterol and low density lipoprotein cholesterol in Hispanics and non-Hispanic whites. *Atherosclerosis.* 1993;98(2):201-211. doi:10.1016/0021-9150(93)90129-I

309. Luo S, Lu JY, Liu L, et al. Divergent lncRNAs Regulate Gene Expression and Lineage Differentiation in Pluripotent Cells. *Cell Stem Cell.* 2016;18(5):637-652. doi:10.1016/j.stem.2016.01.024

310. Dufouil C, Richard F, Fievet N, et al. APOE genotype, cholesterol level, lipid-lowering treatment, and dementia: The Three-City Study. *Neurology*. 2005;64(9):1531-1538. doi:10.1212/01.WNL.0000160114.42643.31

311. Tao Q-Q, Chen Y, Liu Z-J, et al. Associations between apolipoprotein E genotypes and serum levels of glucose, cholesterol, and triglycerides in a cognitively normal aging Han Chinese population. *Clin Interv Aging*. 2014;9:1063-1067. doi:10.2147/CIA.S62554

312. de-Andrade FM, Larrandaburu M, Callegari-Jacques SM, Gastaldo G, Hutz MH. Association of apolipoprotein E polymorphism with plasma lipids and Alzheimer's disease in a Southern Brazilian population. *Brazilian Journal of Medical and Biological Research.* 2000;33(5):529-537. doi:10.1590/S0100-879X2000000500007

313. Evans RM, Hui S, Perkins A, Lahiri DK, Poirier J, Farlow MR. Cholesterol and APOE genotype interact to influence Alzheimer disease progression. *Neurology*. 2004;62(10):1869. doi:10.1212/01.WNL.0000125323.15458.3F

314. Evans RM, Emsley CL, Gao S, et al. Serum cholesterol, APOE genotype, and the risk of Alzheimer's disease: a population-based study of African Americans. *Neurology*. 2000;54(1):240-242. doi:10.1212/wnl.54.1.240

315. Sáiz-Vazquez O, Puente-Martínez A, Ubillos-Landa S, Pacheco-Bonrostro J, Santabárbara J. Cholesterol and Alzheimer's Disease Risk: A Meta-Meta-Analysis. *Brain Sci.* 2020;10(6). doi:10.3390/brainsci10060386

316. Marcum ZA, Walker R, Bobb JF, et al. Serum Cholesterol and Incident Alzheimer's Disease: Findings from the Adult Changes in Thought Study. *J Am Geriatr Soc.* 2018;66(12):2344-2352. doi:10.1111/jgs.15581

317. Eckert GP, Hooff GP, Strandjord DM, et al. Regulation of the brain isoprenoids farnesyl- and geranylgeranylpyrophosphate is altered in male Alzheimer patients. *Neurobiology of Disease.* 2009;35(2):251-257. doi:10.1016/j.nbd.2009.05.005

318. Pelleieux S, Picard C, Lamarre-Théroux L, et al. Isoprenoids and tau pathology in sporadic Alzheimer's disease. *Neurobiol Aging.* 2018;65:132-139. doi:10.1016/j.neurobiolaging.2018.01.012

319. Tian C, Gregersen PK, Seldin MF. Accounting for ancestry: population substructure and genome-wide association studies. *Hum Mol Genet.* 2008;17(R2):R143-R150. doi:10.1093/hmg/ddn268

320. Reitz C, Jun G, Naj A, et al. Variants in the ATP-binding cassette transporter (ABCA7), apolipoprotein E ϵ4,and the risk of late-onset Alzheimer disease in African Americans. *JAMA.* 2013;309(14):1483-1492. doi:10.1001/jama.2013.2973

321. Wu Y, Wang Z, Jia X, et al. Prediction of Alzheimer's disease with serum lipid levels in Asian individuals: a meta-analysis. *Biomarkers.* 2019;24(4):341-351. doi:10.1080/1354750X.2019.1571633

322. Novembre J, Johnson T, Bryc K, et al. Genes mirror geography within Europe. *Nature.* 2008;456(7218):98-101. doi:10.1038/nature07331

323. Panza F, Solfrizzi V, Torres F, et al. Decreased frequency of apolipoprotein E epsilon4 allele from Northern to Southern Europe in Alzheimer's disease patients and centenarians. *Neurosci Lett.* 1999;277(1):53-56. doi:10.1016/s0304-3940(99)00860-5

324. Panza F, Solfrizzi V, D'Introno A, et al. Shifts in angiotensin I converting enzyme insertion allele frequency across Europe: implications for Alzheimer's disease risk. *J Neurol Neurosurg Psychiatry.* 2003;74(8):1159-1161. doi:10.1136/jnnp.74.8.1159

325. Beecham GW, Hamilton K, Naj AC, et al. Genome-wide association meta-analysis of neuropathologic features of Alzheimer's disease and related dementias. *PLoS Genet.* 2014;10(9):e1004606. doi:10.1371/journal.pgen.1004606

326. Rabinovici GD, Carrillo MC, Forman M, et al. Multiple comorbid neuropathologies in the setting of Alzheimer's disease neuropathology and implications for drug development. *Alzheimer's & Dementia: Translational Research & Clinical Interventions.* 2017;3(1):83-91. doi:10.1016/j.trci.2016.09.002

327. Bentley AR, Sung YJ, Brown MR, et al. Multi-ancestry genome-wide gene-smoking interaction study of 387,272 individuals identifies new loci associated with serum lipids. *Nat Genet.* 2019;51(4):636-648. doi:10.1038/s41588-019-0378-y

328. Lifelines Cohort Study, Kilpeläinen TO, Bentley AR, et al. Multi-ancestry study of blood lipid levels identifies four loci interacting with physical activity. *Nature Communications.* 2019;10(1). doi:10.1038/s41467-018-08008-w

329. Stark R, Grzelak M, Hadfield J. RNA sequencing: the teenage years. *Nature Reviews Genetics.* 2019;20(11):631-656. doi:10.1038/s41576-019-0150-2

330. Franzén O, Gan L-M, Björkegren JLM. PanglaoDB: a web server for exploration of mouse and human single-cell RNA sequencing data. *Database (Oxford).* 2019;2019. doi:10.1093/database/baz046

331. Ramdane S, Daoudi-Gueddah D. Mild Hypercholesterolemia, Normal Plasma Triglycerides, and Normal Glucose Levels Across Dementia Staging in Alzheimer's Disease: A Clinical Setting-Based Retrospective Study. *American Journal of Alzheimer's Disease & Other Dementiasr.* 2011;26(5):399-405. doi:10.1177/1533317511414552

332. Wilke C O. *Cowplot: Streamlined Plot Theme and Plot Annotations for "Ggplot2."*; 2019. https://CRAN.R-project.org/package=cowplot

333. Dowle M, Srinivasan A. *Data.Table: Extension of `data.Frame`.*; 2019. https://CRAN.R-project.org/package=data.table

Appendix 1 Supplementary material for Chapter 1

sTable 1.1 Hypercholesterolemia and AD risk

	Study	Cohort	Age @ TC [years]	TC [mM]	Age @ follow up [years]	OR/HR groups	OR/HR[d]	95% CI	P	N	model
Midlife											
	21	7 countries study (***men only***)	40-74[b]	6.6	70-104[b]	high TC>=6.5 mM	**3.1**	1.2-8.5	0.024	471	LR
	18	the North Karelia project & FINMONICA study	50.2/54.0 (CTRL/AD)	6.7/7.2 (CTRL/AD)	71.1/74.7 (CTRL/AD)	high TC>=6.5 mM	**2.2**	1.0-4.7	< 0.05	1400	LR
	17	the North Karelia project & FINMONICA study	50.4/50.0 (APOE-ε4 neg/ε4 pos)	6.68/6.85 (APOE-ε4 neg/ε4 pos)	71.3/70.9 (APOE-ε4 neg/ε4 pos)	high TC>=6.5 mM	**2.8**	1.2-6.7	< 0.05	1287	LR
	16	CAIDE study	50.6	~6.7	71.6	high TC>6.5 mM	**2.12**	1.05-4.30	< 0.05	1409	LR
	19	The Kaiser Permanente Northern California Medical Group	40–45	5.8/5.9 (CTRL/AD)	68.6/69.9 (CTRL/AD)	high TC>=6.2 mM	**1.57**	1.23-2.01	< 0.05	9717	CPH
	169	The Prospective Population Study of Women (***women only***)	38-60[a]	6.1-7.3[a]	32 years after TC	highest quartile	**2.82**	0.94-8.43	n.s.	1462	CPH
	170	ULSAM	50	6.9	<=90	high TC>7.0 mM	**1.0**	0.9-1.2	n.s.	2268	CPH
	20	ILSE	61-65	6.4/6.0 (AD/CTRL)	73-78	highest quartile	**2.64**	1.12-6.23	< 0.05	222	LR
Late midlife											
	168	Framingham heart study	<65/76.1 (before BL/BL)[c]	5.6 (BL)	4-12 after BL	continuous	**0.95/0.97**	0.87-1.04/ 0.90-1.05		1026	CPH
	331		55-94	4.4/5.1 (CTRL/AD)	same as TC	high TC>=4.9 mM	**7.77**	4.96-12.21	< 0.0001	470	LR
Late life											
	175	cross-sectional	74.2/80.4 (CTRL/AD)	5.4/5.2/5.1 (White/Black/Hispanic)	same as TC	lowest quartile	**1.3**	0.8-2.1	n.s.	987	LR
		prospective	74.2/77.6 (CTRL/AD)	5.4/5.2/5.1 (White/Black/Hispanic)	2.5 year follow up	lowest quartile	**1.6**	1.0-2.7		987	CPH
	174		82.0/87.0 (NoAD-D/AD)	5.0/5.4 (NoAD-D/AD)	85.0/90.5 (NoAD-D/AD)	continuous, OR refers to a 0.3 mM increase	**1.22**	1.05-1.41	0.009	106	LR: ***CERAD AD***
	176	cross-sectional	77.2	5.1	same as TC	highest quartile	**0.94**	0.58-1.52	0.56	2820	LR
		prospective	78.4	5.3	82.6 (age of onset for AD)	highest quartile	**0.48**	0.26-0.86	0.04	975	CPH

sTable 1.1 cont.

	Study	Cohort	Age @ TC [years]	TC [mM]	Age @ follow up [years]	OR/HR groups	OR/HR[d]	95% CI	P	N	model
Late life											
	[177]	ACT	<= 8 years before enrollment and AD	6.0/6.0 (CTRL/dementia)	74.9 (enrollment)	highest quartile	1.0	0.61-1.62	n.s.	2112	CPH
	[170]	ULSAM	70	5.8	<=90	high TC>7.0 mM	1.1	0.9-1.3	n.s.	1174	CPH
	[178]	3C study	73.8	5.8	13 year follow up	continuous, HR refers to 1 SD increase	1.12	1.02-1.23	0.017	7369	CPH

Unless otherwise stated, AD refers to clinical diagnoses of AD (i.e. possible or probable AD).

[a] means over birth cohort's range.

[b] estimated, TC levels averaged over a 15-year period with baseline age 40-59.

[c] age before baseline (<65) is an estimation. In article TC levels before baseline are averaged over a time period ending at least 10 years before baseline. Age for at this time point is not reported.

[d] OR for logistic regressions, HR for cox.

Abbreviations: BL, baseline; CI, confidence interval; CPH, Cox proportional hazards model; CTRL, control group; HR, hazards ratio; LR, logistic regression; NoAD-D, dementia other than AD; OR, odds ratio; TC, total cholesterol.

Appendix 2 Supplementary material for Chapter 2

sTable 2.1 Data sets, software, and R packages

Data	
1000 Genomes	https://www.internationalgenome.org/ [236]
ADNI	http://adni.loni.usc.edu/
PREVENT-AD	https://openpreventad.loris.ca/ [228]
StoP-AD centre	https://douglas.research.mcgill.ca/stop-ad-centre [228]
Software	
PLINK	https://www.cog-genomics.org/plink2 [232,233]
R	https://www.r-project.org/ [248]
Sanger Imputation Service	https://imputation.sanger.ac.uk/ [237]
R packages	
car	https://socialsciences.mcmaster.ca/jfox/Books/Companion/ [251]
cowplot	https://CRAN.R-project.org/package=cowplot/ [332]
emmeans	https://github.com/rvlenth/emmeans [254]
psych	https://CRAN.R-project.org/package=psych/ [252]
rcompanion	https://CRAN.R-project.org/package=rcompanion/ [253]
survival	[256,257]
survminer	https://CRAN.R-project.org/package=survminer/ [294]
tidyverse	https://www.tidyverse.org/ [249]

sTable 2.2 Kaplan-Meier cohort characteristics

	***APOE*-ε4 non-carriers**					***APOE*-ε4 carriers**				
	rs72633963 A non-carriers		**rs72633963 A carriers**			**rs72633963 A non-carriers**		**rs72633963 A carriers**		
	N	**Mean**	**N**	**Mean**	**p**	**N**	**Mean**	**N**	**Mean**	**p**
ADNI										
Age [years][#]	163	74.5 (0.56)	31	76.6 (1.28)	0.135	102	73.4 (0.67)	33	72.1 (0.90)	0.255
Females [%]	163	44.8 (3.9)	31	41.9 (9)	0.924	102	38.3 (4.8)	33	36.4 (8.5)	1.000
ext-QFP										
Age [years][#]	33	73.4 (1.72)	6	74.7 (4.61)	0.799	54	71.2 (1.12)	9	72.8 (2.01)	0.511
Females [%]	33	57.6 (8.7)	6	66.7 (21.1)	1.000	54	68.5 (6.4)	9	55.6 (17.6)	0.703

[#] baseline age for ADNI and age at death for ext-QFP

sTable 2.3 ANOVA statistics

	APOE-ε4 non-carriers				*APOE*-ε4 carriers							
	rs72633963 A non-carriers		rs72633963 A carriers		rs72633963 A non-carriers		rs72633963 A carriers		Interaction		Main rs72633963	
variable	**N**	**Mean (SE)**	**N**	**Mean (SE)**	**N**	**Mean (SE)**	**N**	**Mean (SE)**	**F statistic**	**P**	**F statistic**	**P**
PREVENT-AD												
TC	52	5.33 (0.09)	15	5.02 (0.22)	31	5.61 (0.16)	7	4.73 (0.22)	2.703	0.103	11.392	**0.001**
HDL-C	51	1.53 (0.05)	15	1.55 (0.07)	30	1.55 (0.07)	8	1.45 (0.20)	0.969	0.327	0.328	0.568
LDL-C	49	3.03 (0.09)	15	2.59 (0.21)	28	3.22 (0.14)	7	2.70 (0.30)	0.235	0.629	8.352	**0.005**
CSF Aβ-42/p-TAU	49	26.51 (0.94)	16	25.77 (2.31)	27	22.12 (2.01)	6	25.80 (3.21)	0.988	0.323	0.436	0.511
CSF Aβ-42	49	1147.78 (35.52)	15	1123.69 (66.61)	27	921.26 (51.71)	6	1120.37 (92.69)	2.641	0.108	1.624	0.206
CSF p-TAU	47	44.33 (2.04)	15	42.98 (3.53)	26	46.91 (3.26)	6	45.35 (4.72)	0.001	0.979	0.136	0.714
CSF TAU	48	246.33 (12.78)	15	237.80 (21.83)	26	280.60 (24.48)	6	255.87 (25.53)	0.081	0.777	0.480	0.490
ADNI												
HMGCR RNA	316	4.96 (0.02)	84	5.06 (0.04)	206	4.98 (0.02)	62	5.05 (0.04)	0.193	0.660	6.844	**0.009**
log10(CSF Aβ-42 @ BL)	192	2.98 (0.01)	47	2.97 (0.03)	156	2.86 (0.01)	53	2.82 (0.02)	0.412	0.521	1.951	0.163
log10(CSF p-TAU/TAU @ BL)	266	-1.04 (0.00)	71	-1.05 (0.00)	168	-1.01 (0.00)	56	-1.01 (0.01)	0.051	0.821	0.145	0.704
log10(CSF TAU @ BL)	266	2.36 (0.01)	72	2.36 (0.02)	169	2.48 (0.01)	56	2.46 (0.02)	0.279	0.597	0.485	0.487
log10(TC)	344	2.27 (0.00)	88	2.26 (0.01)	227	2.28 (0.01)	72	2.29 (0.01)	2.053	0.152	0.000	0.996
ext-QFP												
Fr.Ctx. HMGCR RNA	32	4.78 (0.35)	9	6.51 (1.16)	41	4.28 (0.26)	9	4.00 (0.38)	3.538	**0.063**	3.775	**0.055**
Fr.Ctx. FL-HMGCR	33	1.01 (0.03)	9	1.11 (0.06)	40	1.03 (0.02)	9	0.97 (0.06)	5.315	**0.024**	0.001	0.972
log10(Fr.Ctx. Δ13-HMGCR)	32	0.14 (0.05)	9	-0.09 (0.06)	43	0.30 (0.06)	9	0.26 (0.13)	1.163	0.284	2.610	0.110
NFTs	36	108.28 (13.35)	9	112.67 (28.26)	41	195.93 (15.46)	7	155.71 (48.29)	0.652	0.422	0.400	0.529
NPs	36	109.33 (13.30)	9	99.89 (27.70)	41	234.80 (17.87)	7	132.71 (23.14)	2.891	**0.093**	4.244	**0.042**

Abbreviations: Δ13-HMGCR, HMGCR transcript lacking exon 13; BL, baseline; CSF, cerebrospinal fluid; FL-HMGCR, full length HMGCR; Fr.Ctx., frontal cortex; NFTs, neurofibrillary tangles; NP, neuritic plaque; TC, total cholesterol.

sTable 2.4 Mann-Whitney Wilcoxon statistics

	APOE-ε4 non-carriers						*APOE*-ε4 carriers					
	rs72633963 A non-carriers		rs72633963 A carriers				rs72633963 A non-carriers		rs72633963 A carriers			
variable	N	Median	N	Median	W statistic	P	N	Median	N	Median	W statistic	P
PREVENT-AD												
CSF p-TAU/Aβ-42	50	0.039	16	0.043	362	0.575	27	0.045	6	0.037	97	0.479
ADNI												
CSF p-TAU/Aβ-42 @ BL	193	0.019	47	0.017	4781	0.566	156	0.046	53	0.044	4021	0.767
CSF p-TAU @ BL	266	19.78	72	19.34	9753	0.811	169	28.25	56	25.81	5126	0.351
TG	345	122	88	125	15296	0.912	228	121.5	72	126.5	8179	0.965

Abbreviations: BL, baseline; CSF, cerebrospinal fluid.

sTable 2.5 Additional logistic regression in ADNI

	***APOE*-ε4 non-carrier**			***APOE*-ε4 carrier**		
	N	**OR**	**p**	**N**	**OR**	**p**
Aβ (+) status[#]	240	0.756	0.431	209	0.917	0.814
"Biological" AD[$]	97	0.622	0.416	113	1.851	0.300

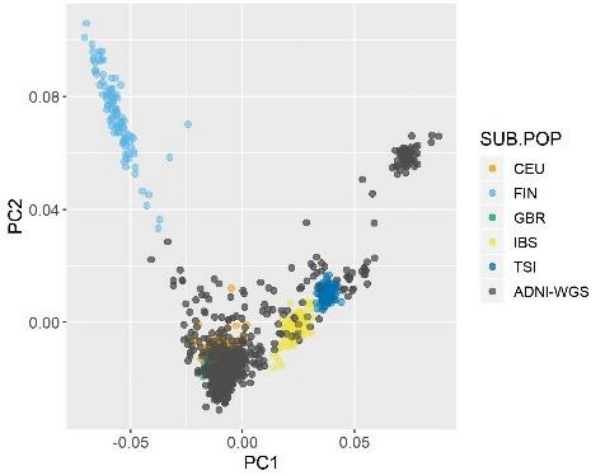
Abbreviations: OR, odds ratio.

[#] defined as a CSF p-TAU/Aβ-42 ratio > 0.028 (according to [104])

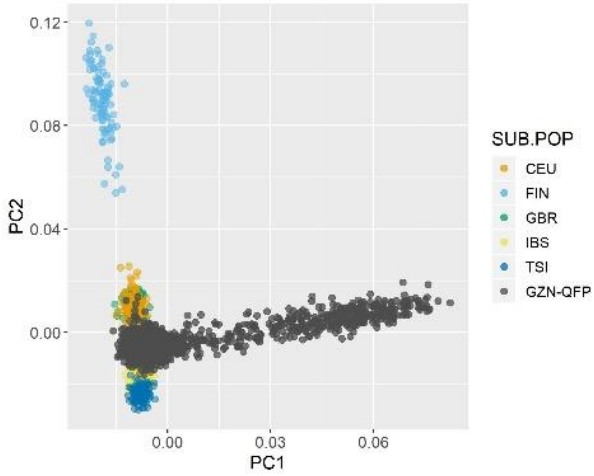
[$] controls defined as amyloid (-) and cognitively unimpaired, AD defined as Aβ (+) and with a clinical diagnosis of AD.

sFigure 2.1 Population stratification in ADNI and ext-QFP

(A) ADNI



(B) ext-QFP



Population stratification in individuals with European ancestry in ADNI (A) and ext-QFP (B) after mapping with the 1000 Genomes.
Abbreviations: PCA, principle component analysis; PC, principal component; CEU, Utah residents with Northern and Western European ancestry; FIN, Finnish in Finland; GBR, British in England and Scotland; IBS, Iberian populations in Spain; TSI, Toscani in Italy.

Appendix 3 Supplementary material for Chapter 3

sTable 3.1 REC buffers and media

Reagent	Cat#	Company	Volume		Final Conc.
Wash buffer (WB) [V = 50 ml]:					
PBS	10010-023	ThermoFisher/Gibco	50	ml	
Antibiotic-Antimycotic [100x]	15240062	ThermoFisher/Gibco	0.5	ml	1x
Renal epithelial primary medium (RE primary medium) [V = 50 ml]:					
DMEM/F-12, GlutaMAX™ supplement	10565018	ThermoFisher/Gibco	450	ml	
REGM™ SingleQuots™	CC-4127	Lonza			
FBS	12483-020	ThermoFisher/Gibco	50	ml	10%
Antibiotic-Antimycotic [100x]	15240062	ThermoFisher/Gibco	5	ml	1x
10% FBS DMEM/F12 [V = 500 ml]					
DMEM/F-12, GlutaMAX™ supplement	10565018	ThermoFisher/Gibco	450	ml	
FBS	12483-020	ThermoFisher/Gibco	50	ml	10%

sTable 3.2 NPC buffers and media

Product	Cat#	Company	Volume		Conc.
Neural induction medium 1 [V = 50 ml]					
DMEM/F-12, GlutaMAX™ supplement	10565018	Thermofisher/Gibco	48	ml	
N-2 supplement (100x)	17502001	ThermoFisher/Gibco	0.5	ml	
B-27® Supplement (50X), serum free	17504044	ThermoFisher/Gibco	1	ml	1x
BSA [7.5%]	15-260 037	ThermoFisher/Gibco	0.7	ml	
MEM Non-Essential Amino Acids Solution (100X)	11140050	ThermoFisher/Gibco	0.5	ml	1x
SB431542	72232	STEMCELL Technologies	50	ul	10 µM
Human Recombinant Noggin	78060.1	STEMCELL Technologies	50	ul	200 ng/ml
Laminin from Engelbreth-Holm-Swarm murine sarcoma basement membrane	L2020	Sigma-Aldrich	50	ul	1ug/ml
Neural induction medium 2 [V = 50 ml]					
DMEM/F-12, GlutaMAX™ supplement	10565018	Thermofisher/Gibco	48	ml	
N-2 supplement (100x)	17502001	ThermoFisher/Gibco	0.5	ml	1x
B-27® Supplement (50X), serum free	17504044	ThermoFisher/Gibco	1	ml	1x
BSA [7.5%]	15-260 037	ThermoFisher/Gibco	0.7	ml	
MEM Non-Essential Amino Acids Solution (100X)	11140050	ThermoFisher/Gibco	0.5	ml	1x
Laminin from Engelbreth-Holm-Swarm murine sarcoma basement membrane	L2020	Sigma-Aldrich	50	ul	1ug/ml
Neural expansion medium [V = 50 ml]					
DMEM/F-12, GlutaMAX™ supplement	10565018	Thermofisher/Gibco	48	ml	
N-2 supplement (100x)	17502001	ThermoFisher/Gibco	0.5	ml	1x
B-27® Supplement (50X), serum free	17504044	ThermoFisher/Gibco	1	ml	1x
MEM Non-Essential Amino Acids Solution (100X)	11140050	ThermoFisher/Gibco	0.5	ml	1x
Laminin from Engelbreth-Holm-Swarm murine sarcoma basement membrane	L2020	Sigma-Aldrich	50	ul	1ug/ml
Human Recombinant bFGF	78003.1	STEMCELL Technologies			20 ng/ml
Human Recombinant EGF	78006.1	STEMCELL Technologies			20 ng/ml
Neuronal differentiation medium [V = 50 ml]					
BrainPhys™ Neuronal Medium N2-A & SM1 kit	5793	STEMCELL Technologies	50	ml	
BDNF, human [20 ug/ml, 1000x]	Z03208-25	GenScript	50	ul	20 ng/ml
GDNF, human [20 ug/ml, 1000x]	Z02927-50	GenScript	50	ul	20 ng/ml

sTable 3.3 IF antibodies

Product	cat#	Company	Dilution
For iPSC QC:			
Human Embryonic Stem Cell Marker Panel	ab109884	Abcam	
SSEA4 mouse monoclonal			1:50
OCT4 rabbit polyclonal			1:100
TRA-1-60 mouse monoclonal			1:100
Nanog rabbit polyclonal			1:50
For NPC QC:			
STEMdiff™ Human Neural Progenitor Antibody Panel	69001	STEMCELL Technologies	
Nestin mouse monoclonal			1:1000
SOX1 rabbit monoclonal			1:1000
PAX6 rabbit polyclonal			1:500
OCT4 mouse monoclonal			1:1000
For neuron QC:			
Anti-beta III Tubulin antibody	ab18207	Abcam	1:2000
Anti-Synaptotagmin antibody [ASV30]	ab13259	Abcam	1:1000
Anti-MAP2 antibody [HM-2]	ab11267	Abcam	1:100
Secondary antibodies for all stages:			
Goat anti-Rabbit IgG (H+L) Cross-Adsorbed Secondary Antibody, Alexa Fluor 488	A-11008	ThermoFisher/Invitrogen	1:1000
Goat anti-Mouse IgG (H+L) Cross-Adsorbed Secondary Antibody, Alexa Fluor 555	A-21422	ThermoFisher/Invitrogen	1:1000

sTable 3.4 WiCell microarray analysis results

A CL2

Call Table (GRCh37/hg19):

Chr	Start Cyto	End Cyto	Variant Type (% mosaic)	Copy Number	Start	End	Size (bp)	Number of Genes
1	1p36.11	1p36.11	Loss	1	25,599,331	25,647,782	48,452	3
4	4q22.1	4q22.1	Loss	1	90,862,372	90,947,793	85,422	1
10	10p15.3	10p15.3	Loss	1	2,415,439	2,547,022	131,584	0
15	15q11.2	15q11.2	Gain (20%)	3	25,103,660	25,126,013	22,354	1
17	17q25.3	17q25.3	Gain	3	77,367,698	77,392,001	24,304	1
19	19q13.41	19q13.41	Gain	3	53,522,100	53,552,296	30,197	2
X	**Xp22.33**	**Xq28**	**Loss**	**1**	**60,814**	**155,236,747**	**155,175,934**	**1512**

B CL2

Call Table (GRCh37/hg19):

Chr	Start Cyto	End Cyto	Variant Type (% mosaic)	Copy Number	Start	End	Size (bp)	Number of Genes
5	5p15.2	5p15.2	Loss	1	9,902,403	9,924,597	22,195	0
10	10q21.1	10q21.1	Gain	3	56,447,525	56,467,563	20,039	1
13	**13q11**	**13q34**	**Gain (50%)**	**3**	**19,058,717**	**114,889,967**	**95,831,250**	**1017**
15	15q11.2	15q26.3	Gain (<5%)	3	25,030,374	102,252,852	77,222,479	977
17	17q12	17q12	Loss	1	33,677,495	33,768,199	90,705	3
17	17q21.31	17q21.31	Gain (50%)	3	44,165,803	44,293,546	127,744	3
17	17q25.3	17q25.3	Gain	3	77,367,698	77,392,001	24,304	1
X	**Xp22.33**	**Xq28**	**Gain (30%)**	**3**	**60,814**	**155,236,747**	**155,175,934**	**1512**

C CL1

Call Table (GRCh37/hg19):

Chr	Start Cyto	End Cyto	Variant Type (% mosaic)	Copy Number	Start	End	Size (bp)	Number of Genes
7	**7q11.23**	**7q21.11**	**Gain**	**3**	**76,637,441**	**77,635,119**	**997,679**	**16**
9	9q34.2	9q34.2	Gain	3	136,762,458	136,865,879	103,422	1
12	12p12.3	12p12.3	Loss	1	16,934,292	16,971,926	37,635	0
15	15q11.2	15q11.2	Gain (20%)	3	25,035,515	25,080,892	45,378	1
X	**Xp22.33**	**Xq28**	**Gain (30%)**	**3**	**60,814**	**155,236,747**	**155,175,934**	**1512**

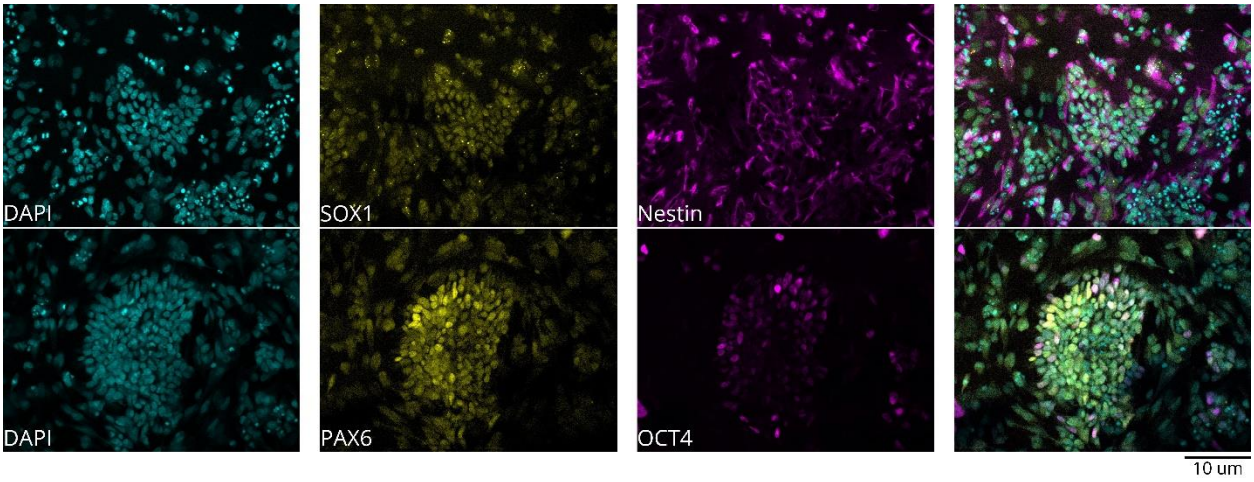
F CL1

Call Table (GRCh37/hg19):

Chr	Start Cyto	End Cyto	Variant Type (% mosaic)	Copy Number	Start	End	Size (bp)	Number of Genes
12	12p13.31	12p13.31	Gain	3	8,000,912	8,123,306	122,395	3
15	15q13.1	15q13.1	Gain	3	29,431,765	29,451,613	19,849	1
16	16q23.1	16q23.1	Gain	3	75,539,436	75,575,410	35,975	2
17	17q21.31	17q21.31	Gain	3	44,165,803	44,277,825	112,023	3
X	Xp22.33	Xp22.33	Gain	3	1,985,554	2,250,212	264,658	1

Reported are all copy number changes identified by the software. Reportable copy number changes are gains or losses > 400 kb. Reportable regions of LOH are > 5Mb. Reportable change are indicated in bold.

sFigure 3.1 NPC IF staining of B CL2



NPC cell line from individual B CL2 were stained for NPC markers SOX1, Nestin, and PAX6 as well as iPSC marker OCT4. Note the cluster like growth characteristic of iPSCs and the presence of OCT4 signal in these cells.

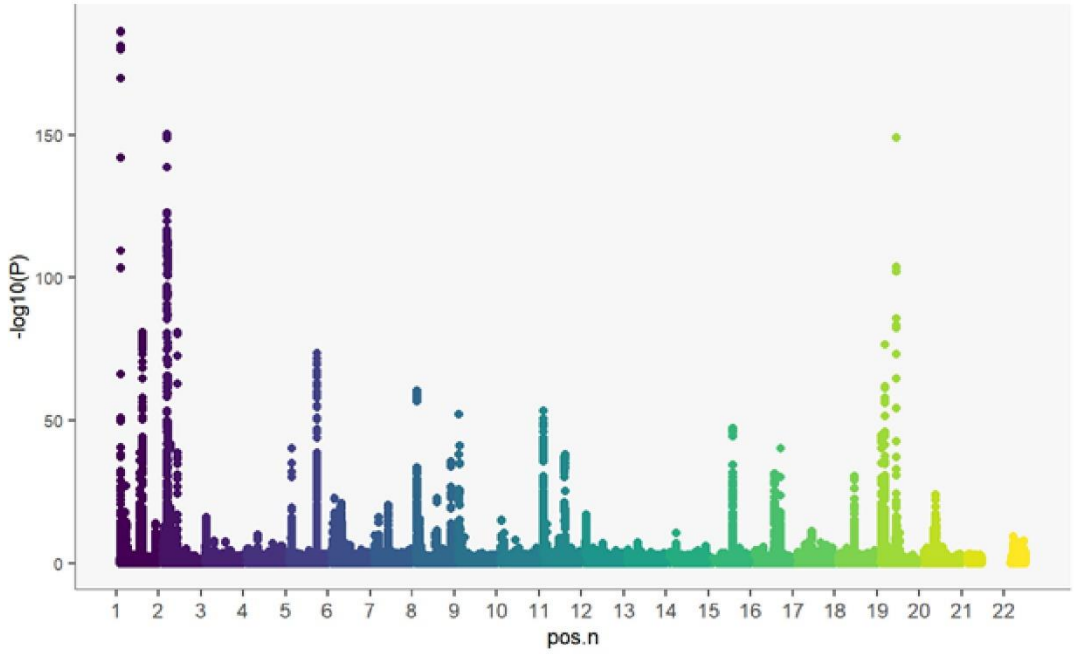
Appendix 4 Supplementary material for Chapter 4

sTable 4.1 Data sets, software and R packages

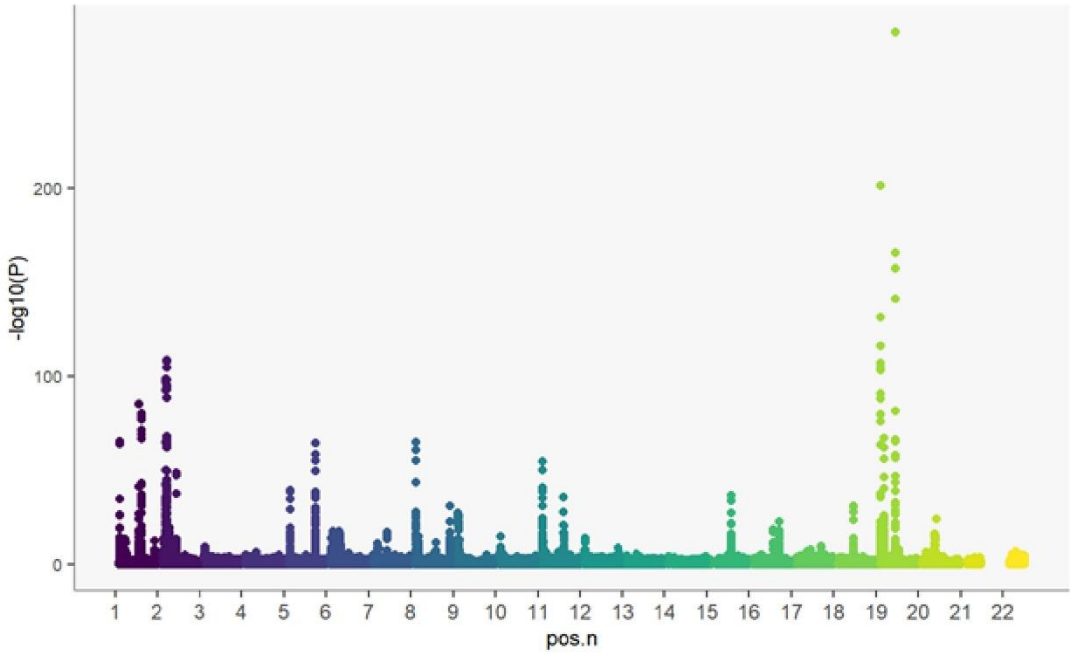
Data	
1000 Genomes	https://www.internationalgenome.org/ [236]
ADNI	http://adni.loni.usc.edu/
Global Lipids Genetics Consortium	http://csg.sph.umich.edu/willer/public/lipids2013/ [280]
GRCh37 genome	https://www.ncbi.nlm.nih.gov/assembly/GCF_000001405.13/#/st
PREVENT-AD	https://openpreventad.loris.ca/ [228]
ROSMAP	https://www.radc.rush.edu/ [285]
StoP-AD centre	https://douglas.research.mcgill.ca/stop-ad-centre [228]
Software	
PLINK	https://www.cog-genomics.org/plink2 [232,233]
R	https://www.r-project.org/ [248]
Sanger Imputation Service	https://imputation.sanger.ac.uk/ [237]
R packages	
boot	[288,289]
cowplot	https://CRAN.R-project.org/package=cowplot/ [332]
data.table	https://CRAN.R-project.org/package=data.table [333]
ggfortify	https://CRAN.R-project.org/package=ggfortify/ [292,293]
pROC	[290]
psych	https://CRAN.R-project.org/package=psych/ [252]
rcompanion	https://CRAN.R-project.org/package=rcompanion/ [253]
survival	[256,257]
survminer	https://CRAN.R-project.org/package=survminer/ [294]
tidyverse	https://www.tidyverse.org/ [249]

sFigure 4.1 Manhattan plots of included/excluded SNPs

(A) Included SNPs



(B) Excluded SNPs



Manhattan plots of SNPs evaluated for the TC-PGS (A) and excluded (B) after matching with the three target cohorts (i.e. SNPs before pruning).