Dissertation

Predictors of brain aging
Prädiktoren der Gehirnalterung

Submitted by

Mag.rer.nat.

Paul Freudenberger

For the Academic Degree of

Doctor of Philosophy

(Ph.D.)

At the

Medical University of Graz

Department of Molecular Biology and Biochemistry

Under supervision of

Prof. Dr. Dr. Helena Schmidt

2015

Eidesstattliche Erklärung

Ich erkläre ehrenwörtlich, dass ich die vorliegende Arbeit selbständig angefertigt und abgefasst, und jene Personen und Institutionen, die am Zustandekommen der Forschungsdaten beteiligt waren, namentlich genannt habe. Andere als die angegebenen Quellen habe ich nicht verwendet und die den benutzten Quellen wörtlich oder inhaltlich entnommenen Stellen habe ich als solche kenntlich gemacht. Die Arbeit an der Dissertation und daraus entstandener Publikationen wurde gemäß den Regeln der „Good Scientific Practice“ durchgeführt.

Graz, am 01.12.15

Declaration

I hereby declare that this thesis is my own original work and that I have fully acknowledged by name all of those individuals and organizations that have contributed to the research for this thesis. Due acknowledgement has been made in the text to all other material used. Throughout this thesis and in all related publications I followed the guidelines of "Good Scientific Practice.

Graz, 1st of December 2015

Acknowledgement

I am grateful for my mentor Prof. Helena Schmidt giving me the opportunity to carry out my research and her guidance over the past years. She was closely involved in my work and her insightful suggestions and support guided me to the completion of my thesis and publications.
She gave me the opportunity to contribute to internationally recognized research articles, introduced me to a host of researchers in the field, and gave me the opportunity to participate in international teaching courses and conferences.
Secondly I also want to thank the additional members of my thesis committee Prof. Reinhold Schmidt and Prof Christian Windpassinger for his valuable feedback and contribution to my thesis.
At this point I also would to extend my thankfulness to my PhD colleagues Dr. Abhijit Sen, Msc Anna Töglhofer and Dr. Edith Hofer their support and insight, as well as for many interesting discussions.
Additional thanks go to Dipl.Ing. Johannes Semmler, Irmgard Pölzl and Michael Schallert for teaching me many useful wet lab techniques. I would also like to thank my colleagues André Fixa, Anita Harb, Christoph Murg, Elfriede Hofer, Philip Spechtl and Veronika Schmidt for their help.
I would like to dedicate my thesis to my parents who supported and encouraged me in all my endeavors. Without their devotion and support I would have never been able to achieve this.
In addition I thank my friends for their thorough proof reading of the manuscript.
Last but not least I want to express my gratitude to my girlfriend Katharina for being incredibly supportive and patient and helping me to complete my thesis.

List of Manuscripts

First authorships

- Freudenberger P, Hofer E, Petrovic K, et al. Fitness and Cognition in the Elderly: the Austrian Stroke Prevention Study. Accepted in *Neurology* 2015

- Freudenberger P, Schmidt R, Schmidt H. Genetics of age-related white matter lesions from linkage to genome wide association studies. *J Neurol Sci* 2012;322(1-2):82-6.

Co-authorships

- Hofer E, Cavalieri M, Bis J, et al. White matter lesion progression: A genome-wide search for genetic influences. Accepted in *Stroke* 2015

- Verhaaren BF, Debette S, Bis JC, et al. Multi-ethnic Genome-Wide Association Study of Cerebral White Matter Hyperintensities on MRI. *Circ Cardiovasc Genet* 2015;8(2):398-409.

- Gorski M, Tin A, Garnaas M, et al. Genome-wide study of kidney function decline in individuals of European descent. *Kidney Int* 2015;87(5):1017-29.

- Woo D, Falcone GJ, Devan WJ, et al. Meta analysis of Genome-Wide Association Studies identifies 1q22 as a Susceptibility Locus for Intracerebral Hemorrhage. *Am J Hum Genet* 2014;94(4):511-21.

- Sen A, Marsche G, Freudenberger P, et al. Higher concentration of lutein+zeaxanthin and vitamin C associated with longer telomere length: results of the Austrian Stroke Prevention Study. *J Am Geriatr Soc* 2014;62(2):222-9.

- Devan WJ, Falcone GJ, Anderson CD, et al. Heritability estimates identify a substantial genetic contribution to risk and outcome of intracerebral hemorrhage. *Stroke*

2013;44(6):1578-83.

- Schmidt H, Freudenberger P, Seiler S, Schmidt R. Genetics of subcortical vascular dementia. *Exp Gerontol* 2012;47(11):873-7.

- Sen A, Gider P, Cavalieri M, et al. Association of cardiorespiratory fitness and morphological brain changes in the elderly: results of the Austrian Stroke Prevention Study. *Neurodegener Dis* 2012;10(1-4):135-7.

- Schmidt H, Zeginigg M, Wiltgen M, et al. Genetic variants of the NOTCH3 gene in the elderly and magnetic resonance imaging correlates of age-related cerebral small vessel disease. *Brain* 2011;134(Pt 11):3384-97.

Collaborations

- Jun G, Ibrahim-Verbaas CA, Vronskaya M, et al. A novel Alzheimer disease locus located near the gene encoding tau protein. *Mol Psychiatry* 2015 doi: 10.1038/mp.2015.23. [Epub ahead of print]

- Lambert JC, Ibrahim-Verbaas C, Harold D, et al. Meta-analysis of 74,046 individuals identifies 11 new susceptibility loci for Alzheimer's disease. *Nat Genet* 2013;45(12):1452-8.

- Seiler S, Schmidt H, Lechner A, et al. Driving cessation and dementia: results of the prospective registry on dementia in Austria (PRODEM). *PLoS One* 2012;7(12):e52710.

Selected presentation at conferences

Oral presentation

- 26th Meeting of the Austrian Alzheimer's Society, Vienna, Austria (23-25 Oct, 2013): Freudenberger P, Petrovic K, Sen A, et al. VO_2max and Brain Aging: Cognitive and MRI Findings in the Austrian Stroke Prevention Study

Poster presentation

- 21st meeting of the World Congress of Neurology Vienna, Austria (21-26 Sept, 2013): (Abstract published; Journal of Neurological Science; doi:10.1016/j.jns.2013.07.1158) Freudenberger P, Petrovic K, Sen A, et al. VO_2max and Brain Aging: Cognitive and MRI Findings in the Austrian Stroke Prevention Study

- European Society of Human Genetics (ESHG), Nuremberg (23-26 June, 2012): Freudenberger P, Petrovic K, Sen A, et al. VO_2max and Brain Aging: Cognitive and MRI Findings in the Austrian Stroke Prevention Study

- 25th Meeting of the Austrian Alzheimer's Society (24-25 June, 2012) Innsbruck, Austria Freudenberger P, Petrovic K, Sen A, et al. VO_2max and Brain Aging: Cognitive and MRI Findings in the Austrian Stroke Prevention Study
  at the

- 24th Meeting of the Austrian Alzheimer's Society (24-25 June,2011) Horn, Austria Freudenberger P, Petrovic K, Sen A, et al. VO_2max and Brain Aging: Cognitive and MRI Findings in the Austrian Stroke Prevention Study

Table of contents

9.1 Course participation

9.2 References

List of abbreviations

Chapter I

ASPS	Austrian Stroke Prevention Study
BPF	brain parenchymal fraction
FLAIR	fluid attenuated inversion recovery
LOD	logarithm of odds
MAF	minor allele frequency
MRI	magnetic resonance imaging
SPSS	Statistics Package for Social Sciences
VaD	vascular dementia
WML	white matter lesions
WMH	white matter hyperintensities
VO_2max	maximum oxygen consumption

Chapter II

ASPS	Austrian Stroke Prevention Study
BPF	brain parenchymal fraction
CADASIL	Cerebral autosomal dominant Ateriopathy with Subcortical Infarcts and leukoencephalopathy
cM	centimorgan
GWAS	genome wide association studies
LOD	Logarithm of odds
MAF	minor allele frequency
OR	odds ratio
SCI	silent cerebral infarct
SNP	single nucleotide polymorphism
SPSS	Statistics Package for Social Sciences
WML	white matter lesions

Chapter III

ASPS	Austrian Stroke Prevention Study
GWAS	genome wide association studies
ICH	intracerebral hemorrhage
LOAD	Late onset Alzheimer's disease
OR	Odds ratio
SNP	single nucleotide polymorphism

Zusammenfassung

Der Rückgang der Kognitiven Fähigkeiten ist ein natürliches Phänomen der Hirnalterung. Sehr oft ist dieser ein Vorbote schwererer Krankheiten und führt zu Demenz. Durch die globale Erhöhung der Lebensdauer und Alterung der Gesellschaften leiden Millionen von Menschen an Krankheiten, die mit Gehirn Alterung assoziiert sind. Die zugrundeliegenden Mechanismen, sind komplex und werden sowohl von genetischen Faktoren, als auch Umgebung und Lebensstil beinflusst. Diese These stellt neue und eigene Forschung zum Thema Hirnalterung in drei separaten Kapiteln vor.

In Kapitel I wird der Einfluss von kardiorespiratorischer Fitness im Rahmen der gesunden Hirnalterung untersucht. Unseren Daten zufolge ist die kardiorespiratorische Fitness mit weniger Läsionen der weißen Hirnmasse, die Markenzeichen der strukturellen Hirnalterung vaskulären Ursprungs bei gesunden älteren Menschen darstellt, verbunden. Außerdem beobachteten wir bei höherer Fitness eine bessere Leistung in allen kognitiven Domänen sowie Kognition insgesamt in derselben Kohorte.

Kapitel II fasst den aktuellen Stand der Literatur zur Genetik von Läsionen der weißen Substanz und vaskulärer Demenz zusammen. Zusätzlich wird Original-Forschung hinsichtlich genetischen Polymorphismen im *NOTCH3* Gen vorgestellt. Seltene Mutationen im Gen sind ursächlich für CADASIL, einer familiären Erkrankung, die mit frühzeitigem Schlaganfall einhergeht. Wir präsentieren Daten, denen zufolge vier häufigen Polymorphismen in *NOTCH3* mit dem Volumen von Läsionen der weißen Hirnmasse bei hypertensiven älteren Menschen assoziiert sind.

Kapitel III repräsentiert meinen Beitrag zur internationalen Untersuchungen des genetischen Hintergrundes der Gehirn Alterung. Im ersten Artikel untersuchen wir die Heritabilität, i.e. die Menge der Varianz in einem Phänotyp, der durch die Genetik erklärt werden kann für intrazerebrale Blutungen, einem Untertyp von Schlaganfall. Unsere Ergebnisse zeigen, dass das Risiko einer Hirnblutung, das Hämatom Volumen und Progression zeigen hohe Vererbbarkeit und signifikante genetische Basiswerte. In einer nachfolgenden Untersuchung führen wir eine Meta-Analyse von genomweiten Assoziationsstudie zu Hirnblutung bei Europäern durch. Wir berichten über eine bisher unbekannte Assoziation auf Chromosom 12q21.1 (rs11179580) für lobare Hirnblutung und auf Chromosom 1q22 (rs2984613) für nicht lobare Hirnblutung. Der letzte Abschnitt des Kapitels III beschreibt eine Meta-Analyse für spät einsetzenden Alzheimer mit über 70.000 Personen. Wir identifizierten 11 neue Loci, die neue Einblicke in eine komplexe Krankheit ermöglichen.

Zusammenfassend sind sowohl Genetik und Umwelt maßgeblich entscheidend für gesunde Hirnalterung. Genetische Forschung ist entscheidend um die Risikofaktoren zu ermitteln und molekulare Pfade zu ermitteln, führen allerdings in den seltensten Fällen zu Behandlungen. Lifestyle-Entscheidungen, wie zum Beispiel Aufrechterhaltung der Fitness sind leicht adaptierbar und obwohl sie Demenz wahrscheinlich nicht vollständig verhindern, haben sie dennoch erhebliche Auswirkungen auf gesundes Altern insgesamt.

Abstract

The decline in a one's mental capacities are a natural phenomenon of brain aging. Very often the cognitive decline precedes more serious diseases and results in dementia.

Due to global increases in life span and aging of societies millions of people suffer from diseases associated with brain aging. The diseases underlying brain aging are complex, and influence by genetic factors, as well as many life-style choices. This thesis provides novel and original research on the topic of brain aging in three separate chapters.

In chapter I the influence of maintaining a high cardiorespiratory fitness is investigated in the context of the healthy brain aging. According to our data, cardiorespiratory fitness is associated with less white matter lesions, which are hallmark of structural brain aging of vascular origin in healthy elderly. In addition, we observed higher levels of fitness to be associated with and better performance in all cognitive domains and overall cognition in the same cohort.

Chapter II summarizes the current state of literature on the genetic underlying of white matter lesions and vascular dementia. In addition original research is presented on genetic polymorphisms in the *NOTCH3* gene. Rare mutations in the gene are causative for CADASIL, a disease associated with early vascular brain aging and vascular stroke. We demonstrate data on four common polymorphisms in *NOTCH* that were associated with increase amounts of white matter lesions in hypertensive elderly.

Chapter III represents my contribution to international investigations on the genetic background of brain aging. In the initial article we investigate the heritability, i.e. the amount of variance in a phenotype that is explained by genetics for intracerebral hemorrhage, a subtype of stroke. Our findings indicate that the risk of intracerebral hemorrhage, the hematoma volume and progression display high heritability and significant genetic underlying.

In a follow up investigation, we present a meta-analysis of genome wide association study with intracerebral hemorrhage in Europeans. We report a novel susceptibility locus finding on chromosome 12q21.1 (rs11179580) for lobar intracerebral hemorrhage and on chromosome 1q22 (rs2984613) for non-lobar intracerebral hemorrhage.

The last section of chapter III we provide a meta-analysis to identify susceptibility loci for late onset Alzheimer's disease in over 70.000 people. We successfully identify 11 new susceptibility loci, which provide novel insight into a complex disease.

In summary, both genetics and life-style influence healthy brain aging. Genetic research is crucial to determine risk factors and pinpoint molecular pathways, but they do not necessarily lead to treatments. Life-style choices, such as maintaining fitness are easily adaptable and although they might not prevent dementia, have significant impact on healthy aging overall.

1. Introduction

1.1 Aging process

The aging process can be described as an inevitable and progressive decline of cellular function, causing increased vulnerability to death (*López-Otin C et al. 2013*). It directly influences a plethora of diseases such as cancer, diabetes, cardiovascular and neurodegenerative diseases.
Many cellular hallmarks have been identified so far, including accumulation of genetic damage (*Moskalev AA et al. 2013*), replication errors (*Hoeijmakers JH 2009*), mitochondrial mutations or dysfunction (*Wallace DC 2010*) (*Green DR 2011*), telomere attrition (*Blackburn EH et al. 2006*), epigenetic alterations (*Talens RP et al. 2012*), loss of protein homeostasis (*Powers ET et al. 2009*), deregulated sensing of hormones i.e. insulin and IGF-1 (*Barzilai N et al. 2012*), and depletion of stem cells (*Jones DL & Rando TA 2011*).

1.2 Brain aging

Due to increase in life-span and aging population in the Western world, aging of the brain has attracted a lot of interest by public and science over the last years. Similar to other tissues the brain undergoes several changes in function and structure that are highly variable across individuals. A certain amount of physiological deterioration is inevitable as age progresses. These changes may also progress and turn pathological, causing mental dysfunction (dementia) and loss of motor functions (ataxia) in its final state.

1.3 Aging related functional and morphological changes in the brain

Brain function is described by the term cognition stemming from the Greek words “con” and “gnsoco” meaning “I perceive” and refers to the entirety of mental capacities. It is assessed by oral or written tests and can be grouped into domains, namely: motor skills, memory, executive function and attention. Not all of these domains are equally affected by age-related

cognitive decline.
Differences are apparent in selective (*Quigley C & Müller MM 2014*) and divided attention (*Tsang PS & Shaner TL 1998*); as well as episodic (*Kinugawa K et al. 2013*), working (*Brockmole JR & Logie RH 2013*) and autobiographical memory (*Levine B et al. 2002*).
The aging brain also exhibits several structural changes, discoverable by post mortem investigations or brain imaging methods. With ongoing age brains display expansion of ventricles, (*Scahill RI et al. 2003)* thinning of the cortex *(Salat DH et al. 2004*) and overall atrophy. Longitudinal studies have revealed that not all brain regions age at the same speed. In fact the hippocampus, cerebellum, prefrontal cortex, caudate nucleus are most affected by atrophic changes (*Raz NN et al. 2005).*

1.3.1 Vascular changes

Aging is also accompanied by changes to vasculature, which may also effect the brain starting in the 3rd or 4th decade of life and can cause cerebrovascular disease. Both small as well as large vessels are affected by age-related changes, but through different mechanisms.
Vascular aging may promote the formation of vessel occlusion and ischemia through a thrombus. If these events occur in small vessels they are referred to as lacunar infarcts, or ischemic stroke if larger supplying arteries are affected. A different manifestation of aging vasculature it the occurrence of brain aneurysms, causing vessel rapture and subsequent hemorrhagic bleeding, and account for~15% of stroke cases.
A stroke is the most devastating vascular incident in brain aging and it is often preceded by subclinical morphological brain changes, such as whiter matter lesions (WML) also referred to as leukoaraiosis, which are visible on T2-weighted MRI scans (*Pantoni L & Garcia JH 1997*). These WML progress and grow with ongoing age, and although they appear asymptotic, they contribute to the formation of stroke (*Schmidt et al. 2003; Prins et al. 2004*).

1.3.2 Neurodegenerative changes

The term neurodegeneration describes the pathological and progressive deterioration of nerve

cell function and structure, which usually manifest in mid to late life. Multiple diseases are characterized as neurodegenerative, among which Alzheimer's, Parkinson's, Huntington's and amyotrophic lateral sclerosis. Hallmarks of these diseases are often accumulation of excess misfolded proteins (proteopathies), such as β amyloid (*Karran A et al. 2011*) and tau protein (*Small SA & Duff K 2008*) in Alzheimer's disease and α-Synuclein in Parkinson's disease (*Stefanis L 2012*).

1.3.3 Pathology, histopathology and pathomechanisms of brain aging

Both neurodegenerative, as well as vascular diseases contribute to pathological brain aging through different mechanisms. In the case of Alzheimer's disease, is characterized by progressive atrophic cell death and gliosis in the hippocampus, temporal and parietal lobe (*Wenk GL 2003*). Histological findings include both amorphous extracellular deposits of A*β* protein (amyloid plaques) and intra-neuronal accumulation of unfolded protein tau (neurofibrillary tangles).

Vascular dementia is the result of accumulated vessel damage, due to progressive cerebral small vessel damage or a traumatic brain infarct. Atherosclerosis of the large cerebral arteries is caused by deposits of lipid macrophages and progressive calcification and ulceration (*Lusis AJ 2000*). In small vessels, the vessel lumen is constricted by smooth muscle hyperplasia and deposition of hyaline material, in arteriolosclerosis (*Munoz DG 2003*).

In case of Parkinson's disease the histological hallmark are so called Lewy bodies, which are intracellular protein aggregates which can occur in different brain regions and predominantly consist of α-Synuclein. In addition the diseases is characterized by death of dopaminergic neurons in the midbrain substantia nigra, causing shortages and the characteristic movement disorders (*Davie CA 2008*).

1.3.4 Epidemiology of Dementia

The risks of age-related neurologic diseases are enormous, with an estimated life-time risk of 1 in 3 of developing either stroke or dementia (*Seshadri S et al. 2006*).

The prevalence of dementia, i.e. the proportion of a population that a have a certain condition, is estimated at 36 million people were affected in 2010, and the number is expected to double every 10 years (*Prince M et al. 2013*). In Austria 100,000 are currently suffering from dementia *(Alf C 2006*). The probability to develop the condition is referred to as incidence, which is highly age dependent. In the case of dementia the incidence rate increases from 3.5 affected in the age group of 65-69 to 72.8 per thousand in elderly of 85 years and older (*Kokmen E et al. 1993*). The most frequent types of dementia are Alzheimer's disease (40-70%) and vascular dementia (15-30%) (*Lobo A et al. 2000*; *Plassmann BL et al 2007*).
Well accepted risk factors for dementia and for cognitive decline, besides age, are low education (*Launer LJ et al. 1999*), hypertension (*Skoog I et al. 1996*), white matter lesions (*Prins ND, et al. 2004*) diabetes mellitus (*Ott A et al.1999),* current smoking (*Launer LJ et al. 1999*), obesity (*Whitmer et al. 2008*; *Elias MF et al. 2005*), or metabolic syndrome (*Yaffe K et al. 2004*).
Less established protective effects are exerted by antioxidants-rich foods (*Otsuka M et al. 2002*), overall caloric restriction (*Mattson MP 2003*), consumption of fish (*He K et al. 2004*), and low to moderate consumption of alcohol (*Ruitenberg A et al. 2002*).
Cognitive or cerebral reserve is a frequently discussed term, describing education and challenging occupation (*Staff RT et al. 2004*), which has been indicated to reduce incidence of AD incidence (*Stern Y et al. 1994*) and less cerebrovascular lesions (*Del Ser T et al. 1999*).
Exercise may represent an important aspect in successful aging. Reports indicate, that physical activity is associated with lower risks of cognitive impairment and dementia (*Laurin D et al. 2001*). Regular exercise plays an important role in both prevention of AD in healthy elderly (*Rolland Y et al 2008*), as well as in people at high risk of developing AD (*Lautenschlager NT et al. 2008*).

1.3.5 Genetic architecture of brain aging related phenotypes

1.3.5.1 Monogenic forms

In some cases, pathological brain aging can be caused by mutations in a single gene.

"CADASIL" is the single most common form of hereditary stroke disorders, and results in cerebral small vessel diseases and stroke at middle age. The causal genetic mutation was discovered to be on chromosome 19, (*Tournier Lasserve E et al. 1993*) more specifically on ch19p13.1 in the *NOTCH3* gene (*Joutel A et al. 1996*), which is involved in signal transduction and differentiation of cells. The prevalence of CADASIL has been estimated at 1.98 affected per 100000 in the Scottish population (*Ravzi SS et al. 2005*).
The early onset form of Alzheimer's disease accounts for a small fraction (1-5%) of AD cases, with a similar disease progression but much earlier onset. It is caused by dominant mutations one of three genes, namely *APP*, *PSEN1*, or *PSEN2* (*Bertram L &Tanzi RE 2008*).

1.3.5.2 Multifactorial forms with previously identified genetic determinants

The proportion in the variability of any phenotype that can be explained by genetic risk factors is referred to as the heritability, while the remainder reflects the influence of the environment. Heritability is estimated by investigating relatives of different grades for concordance of phenotypes (dichotomized traits) or correlations (quantitative traits), such as monozygotic vs dizygotic. These estimates range from 0 to 1 and importantly are always relative to genetic and environmental factors in the relative population. Many hallmarks of morphological brain aging, such as the volume of brain ventricles (*Kremen WS et al. 2012*), individual lobes (*Geschwind DH et al. 2002*) and the brain overall brain (*Batouli SA et al. 2014*) are highly heritable. Pathological lesions of white matter, a hallmark for the cerebrovascular diseases, are highly heritable as well and have been estimated at 0.73 in world war II two veteran twins (*Carmelli D et al. 1998*) and in women at 0.78 in a family study *(Atwood LD et al. 2004*).
Most of the diseases involved in brain aging have a complex genetic architecture and many genetic polymorphisms that contribute to a small percentage in the development are referred to as genetic risk factors. These most commonly investigated genetic risk factors are single nucleotide polymorphisms (SNPs), which represent changes in a single nucleotide (A G C or T), and make up for a large part of the inter-individual variability, but there are other genetic variations such as copy number variations, and large insertions, deletions or rearrangements of DNA. SNPs are distributed unevenly all over the genome and are usually characterized by

their location and their frequency across a given population which is referred to as the minor allele frequency. There are two hypothesis in the field of genetic association, namely that a large amount of common variants (MAF ≥5%) explain any common multigenic disease with a low penetrance (probability the carrier will develop the disease), opposed to the common disease - rare variant hypothesis in which very few rare alleles with low MAF cause the disease. These genetic variants especially important when identifying individuals at risk and determining genetic susceptibility and can act on the disease on their own, or over other risk factors such as blood pressure, diabetes, or obesity.

Elevated blood pressure is a risk factor for white matter lesions, stroke and Alzheimer's disease (*Feldstein CA 2012*), which is regulated by a wide range of genes and their products. A frequent polymorphism in the *AGT* gene (M235T) has for example been identified as risk factor for WML in a candidate gene investigation (*Schmidt R et al. 2001*), which was recently replicated in recent genome wide association (*Fornage M et al. 2011*).

Polymorphisms in the *APOE* Gene, involved in lipid metabolism, also carry a substantial risk for brain aging. Depending on the dosage of the detrimental ε 4 allele, the risk of developing white matter lesions (*Schmidt R et al. 1997*), late onset AD (*Corder EH et al. 1993*) and vascular dementia (*Chuang YF et al. 2010*) is significantly increased.

1.3.6 Clinical consequences

The advent of dementia is usually preceded by preclinical phase followed by mild cognitive impairment (MCI) or insipient dementia (*Petersen RC et al. 1999*). If memory is the only domain affected referred to as prodromal AD (*Grundman M et al. 2004*). Already manifested cognitive decline and MCI increases the likelihood of developing dementia (*Dawe B et al. 1992*). Assessment of cognitive status is frequently done by using mini mental state (MMSE) (*Folstein MF et al. 1975*), or Mattis dementia scale (DRS) (*Pedraza O et al. 2010*).

Dementia itself has several stages and the experienced symptoms depend on the cause of dementia.

In the late stages, patients require constant supervision and develop a range of symptoms such loss of speech, tremor, inability to recognize family members, restlessness, delusion,

depression and loss of basic body functions. Currently there is no cure for neurodegenerative diseases and dementia and the functional changes are progressive and irreversible. The late stages of dementia are accompanied by loss of autonomy and complete dependence, creating an enormous demand for caregivers that also represents drain on the health system. Given the overall aging of the society and the increasing incidence rates of dementia with ongoing age, research on dementia and neurodegenerative diseases will remain a topic of intense research in the following years.

1.4 Current focus in research

Currently the focus lies on identifying common genetic risk factors involved in complex diseases using genome wide association studies (GWAS) and meta-analyses of GWAS datasets. Collaboration between many clinical investigations allow creation of larger cohorts to create unprecedented statistical power. Using these datasets it is possible to identify large amounts of common genetic variations associated with a phenotype, with little effect on the disease. These genetic variations are not causal in most cases but lie in close vicinity to these variations on the genome and ideally yield genes that do not play roles in previously identified pathways involved in a disease. Using subsequent pathway analysis of the associated genes in the regions may yield a deeper understanding of the complex biochemical pathways involved in the pathogenesis. GWAS studies have replaced candidate gene investigations, using genome hybridization chips to screen for a large number of polymorphisms. The next step will likely be a whole exome or genome sequencing approach and even larger cohorts, which allow detection and research on rare variations (<1%MAF) with large effects, gene-gene interaction effects or analysis of larger polymorphisms, such as copy number variations.

2. Aims of the thesis

This works extends previous knowledge on brain aging in several ways, as both genetic as well as the environmental components of brain aging are investigated and novel insight can be presented for both topics in three separate chapters. This is achieved using state of the art scientific conduct and epidemiological and genetic methods.

In the first chapter I aim to ascertain the effect of cardiorespiratory fitness on lesions of white matter, which are a hallmark of structural brain aging. In a follow up study I determine whether the beneficial effect of fitness is also present in functional brain measures and whether it is specific to cognitive domains.

Chapter II contains review articles on white matter lesions and vascular dementia. These articles aim to summarize the current state of genetic research and previously identified risk factors. They provide readers with an overview of topic and different methodologies used involved in dissecting genetic traits. In second part of chapter II and article is presented in which I investigate genetic polymorphisms in the *NOTCH3* gene. Mutations in *NOTCH3* can cause a hereditary disease associated with early onset strokes. The article aims to identify the role of common variants in *NOTCH3* in the development of age-related structural brain aging.

Lastly this thesis also consists of three articles that highlight the author's contribution to modern genome wide association studies and meta-analysis of their results.

The first article aims to estimate the heritability of intracerebral hemorrhage, a subtype of stroke. The last two articles describe meta-analyses on intracerebral hemorrhage and late onset Alzheimer's disease in which we aim to validate existing genetic risk factors and identify novel genes and pathways involved in the development of the disease.

3. Methods section

3.1 Cohorts relevant to the thesis

ASPS

The Austrian stroke prevention study (ASPS) is a single center community based cohort study on the effects of vascular risk factors on brain structure and function in healthy elderly (50-75 years of age, mean age 65±7.7) in Graz, Austria (*Schmidt R et al. 1999; Schmidt R et al. 2005*). It is a longitudinal study, which includes 3 year follow up investigations of the participants.
The enrollment began between 1991 and 1994, during which written invitations were sent out. The cohort was enlarged in a second panel between 1999 and 2003.
The response rate was 32.4%, but interviews with non-responders yielded no significant differences to responders regarding demographics and frequencies of vascular risk factors. (*Schmidt R et al 1997*).
Out of 2794 replies, 2008 were included in total in two separate panels.
The Inclusion criteria prohibited former history of neuropsychiatric disease such as stroke or signs of dementia. Evaluations included clinical history, laboratory evaluation, cognitive testing, and assessment of risk factors. Every 4th participant was additionally invited to undergo Doppler sonography and brain MRI. In total, 996 genotyped and 829 participants have 1000Genomes imputation data available using Illumina Human Quad 610 bead chip.
The official website can be found at http://www.gehirnforschung.at/project/austrian-stroke-prevention-study/ .

ASPS-Fam

The ASPS-Fam is a single center community based family cohort including the initial members of the ASPS cohort and their healthy relatives and offspring in Graz, Austria.
The Recruitment phase took place between 2006 and 2013, and the same inclusion criteria as

in the ASPS were applied. In total 381 individuals from 169 families, with 2-6 participants per family were included. Currently 125 participants have also completed follow up investigations.
The test protocol is similar to the ASPS, but includes more sophisticated MRI methodology. 388 participants have been genotyped and have 1000Genomes data available, using the Affymetrix Genome-Wide Human SNP Array 6.0.

PRODEM

PRODEM is an ongoing longitudinal multi-center cohort study, in 12 different neurological clinics in Austria on demented elderly (*Seiler S et al. 2012*).
Recruitment started in 2009, 437 subjects were initially included and currently over 700 participants have contribute to the study. Inclusion criteria were, (1) dementia diagnosis according to DSM-IV criteria (*American Psychiatric Association 2000*), (2) non-institutionalization and no need for 24 hour supervision; (3) caregiver had to be willing to provide information on the patient on her /his own condition. Subjects not able to sign informed consent or with co-morbidities that could interfere with study completion were excluded.
The investigations includes baseline measurement and 6-months interval follow up examination for 2 years or until institutionalization, withdrawal, loss of follow up, or death occurs. The test protocol includes a range of clinical, behavioral and functional evaluation, including DNA, RNA, plasma, serum sampling, but also MRI scans, medication, caregiver burden, cognitive tests.
280 participants have been genotyped and imputed 1000Genomes data are available using Affymetrix Genome-Wide Human SNP Array 6.0.
The official website can be found at http://www.alzheimer-gesellschaft.at/index.php?id=27.

3.2 Used techniques

3.2.1 Statistical analyses using SPSS

Statistical analysis was performed using IBM software statistics package for social sciences (SPSS) v20 for descriptive statistics, linear regression and correlation analysis of clinical datasets which is available at http://www--01.ibm.com/software/at/analytics/spss/products/statistics/.

3.2.1.1 Descriptive statistics

Normal distribution was tested with the Kolmogorov-Smirnov test, comparison of mean with median, and assessment of QQ-plots, in order to be able to perform parametric tests. Normally distributed variables are reported as mean ± standard deviation and non-normally distributed variables as median and interquartile range. Variables with a skewed distribution were log transformed. In cases of variables containing zero values, f.e. WML volume, the value 1 was added to the volumes before natural log transformation.
Outliers, deviating more than 1.5 interquartile ranges (IQR) were excluded from the analysis.

3.2.1.2 Linear regression

In order to describe the statistical relationship between a dependent variable and a predictor or explanatory variable linear regression was performed. Additional variables (so called covariates), that are possibly predictive of the outcome, were added to the model to evaluate their effect on the predictor variable. Effect size and direction was given as B plus confidence intervals and strength of the association as p. An association was considered significant at p-values<0.5.
Linear regression was also used upon stratification of the dataset (f.e. age groups, sex and APOE ε carrier status). To conclude that effects are restricted to a subgroup, interaction terms between the predictor and the variable used for stratification were calculated by multiplying them. Upon addition of the interaction term to the regression model, modulating effects can be become evident and were considered significant at p-values<0.5.
To determine dose-effect relationships linear regression was performed using a quartile

distribution of the continuous predictor variable, in which the lowest quartile served as a reference. To determine significance, the dose-response was tested using a trend test, which was considered significant at p-values<0.5.

3.2.1.3 Mediation analysis

Simple mediation models were used in order to estimate indirect effect sizes.
In other word, the influence of an independent variable X on the dependent variable Y is accounted for by a mediator M. This is the case if the value of the direct path coefficient between X and Y is reduced by the inclusion of M. Perfect mediation occurs when reduction decreases the value of the direct path to zero, partial mediation occurs if the effect decreases the value to a considerable amount and no evidence for mediation is present when the reduction of the direct path coefficient is not significant. Estimation of mediator effect sizes was calculated using methods with bootstrapped models using a freely available script for SPSS called PROCESS, which is freely available at http://www.processmacro.org/ (*Preacher KF & Hayes AF 2008*). Output was given as total, direct and indirect effect size and bootstrapped confidence intervals.

3.2.2 Statistical analysis involving genomic data

3.2.2.1 Genotyping

The Illumina Human 610 Quad Bead Chip http://support.illumina.com/array/array_kits/human610-quad_beadchip_kit.html were used for the ASPS cohort and the Affymetrix Genome-Wide Human SNP Array 6.0 http://www.affymetrix.com/catalog/131533/AFFY/Genome-Wide+Human+SNP+Array+6.0#1_1 were used for the ASPS-Fam and PRODEM cohorts. The data files for ASPS-Fam and PRODEM using the Affymetrix chip data were used for genotyped in the course of this thesis.
906600 SNPs including SNPS from chromosome X and Y, mitochondrial SNPS, SNPS in

recombination hotspots and also 946000 copy number variations, which are a larger type of sequence variation not investigated in the course of this thesis.

The wet lab work which includes preparation of DNA samples and all subsequent steps until creating chip data was performed by Mag. Anna Maria Töglhofer and Anita Harb.

In brief it is a 5 day protocol, in which a total of 250 ng DNA is digested by two restriction enzymes, namely NSP and Sty in separate reactions. Secondly a specific adaptor is added to ligate the separate reactions. After a subsequent PCR step, an aliquot of the product is placed on an agarose gel for confirmation. In the next step, the NSP and Sty reactions are pooled and purified using DNA beads. Consecutively, the DNA is fragmented, labeled and loaded in the CHIP.

Following an overnight incubation, the CHIP data can be read as a fluorescent signal using the CHIP reader. A more detailed description of the process can be found at http://media.affymetrix.com/support/downloads/manuals/genomewidesnp6_manual_rev7.pdf.

Genotyping was performed using the Affymetrix Genome console software version 4.1 http://www.affymetrix.com/estore/browse/level_seven_software_products_only.jsp?productId=131535#1_1. The software uses the birdseed[3] v1 and v2 algorithm.

According to software guidelines, genotyping was run in a large batch (min 50 people, half of which female), and samples on the same plate were genotyped together. Due to differences in processing of lab samples, namely a step up from low to high-throughput schedule, these samples were genotyped together. Genotyping was performed for the ASPS-F n=376, PRODEM n=283 cohort.

Input files created by CHIP reader were *.CEL and *.ARR files.

A pedigree batch file had to be created using the .ARR files and then uploaded using the AGCC portal in order to receive PLINK genotype files *.map and *.ped.

The batch edit file contained the following columns “Sample File Path, Project, Sample File, Array Name, Probe Array Type, Barcode, Family ID, Individual ID, Father ID, Mother ID, Sex, Affection to be filled out. In case of unrelated individuals (all but ASPS-F) a fake family ID, as well as mother or father ID had to be created.

Before the genotyping step, samples had to fulfill the contrast QC (CQC) step. A CQC value

of >0.4 was considered inbound by manual and the genotyping console. Outbound samples were subjected to repeated CHIP analysis. Average of all samples in a batch had to be >=1.7. After genotyping samples with a birdseed call rate of <97% were discarded.
The *.map output file consisted of columns describing the genetic marker, namely chromosome (1-22, X, Y or 0 if unplaced) rs# or SNP identifier, genetic distance (morgans), and base-pair position (bp units). Pedigree files or *.ped give information about the individual, in particular family ID, individual ID, paternal ID, maternal ID, Sex (1=male; 2=female; other=unknown), as well as phenotype.

3.2.2.2 Quality control of genotype data

Raw genotyped data had to be subjected to quality control in order to be used.
This was done using the free command based software PLINK for analysis of genome-wide data called PLINK http://pngu.mgh.harvard.edu/~purcell/plink/ (*Purcell S et al. 2007*). The guidelines for these procedures were taken from (*Turner S et al. 2011*).
In the first step the palindromic SNPs, f.e. A/T, C/G for which the strand information cannot clearly be determined are removed, to avoid reduction of statistical power in meta-analyses, which was done by Dr. Hofer (*Winkler CJ et al. 2014*).
SNPs lying on the opposite side of the DNA strand were flipped using the --flip command. The information which SNPs needed flipping was created using the annotation file provided by Affymetrix.
plink --noweb --file /home/o_freudenb/ASPSF/ASPSF_genotyped --flip /home/ASPSF/SNPsToFlip.txt --make-bed --out /home/ASPSF/ASPS_flipped
This script would flip the SNPs in the original file and create a binary output which is easier and faster to use for subsequent analysis.
Subsequently SNPs with a low SNP call rate (--geno) of < 98% (i.e. low call rate of this marker across all samples) were removed. In similar fashion to specific markers, individuals with a low call rate of SNPs (--mind) of < 98% were handled the same way.
SNPs with a low minor allele frequency (--maf) (<5%) were also excluded in this quality control step. Lastly, a check for hardy Weinberg equilibrium was performed (--hwe). Hardy

Weinberg equilibrium is a model in population genetics, describing allele frequency in parent and offspring generation. In this case, a departure from the equilibrium would indicate errors in genotyping or population stratification (*Wittke-Thompson JK et al. 2008*).
A sample script would look accordingly:
plink --noweb --/home/ASPSF/ASPS_flipped --mind 0.02 --geno 0.02 --maf 0.05 --hwe 0.000005 --make-bed --out plink_files/Prodem_flipped_QCed
The next step included a sex check, to confirm if the gender of a subject matches with their number of X-chromosomes. The output compares the information about the given gender in the original pedigree files to number to the actual number of computed X chromosomes and states discrepancies.
plink --noweb –bfile ASPS_QCed --check-sex --out ASPS_sexchecked
Lastly, it is important to investigate the dataset for sample relatedness using the --genome command. PLINK will then calculate the proportion of loci in which individuals share zero, one or two alleles that are identical by descent (IBD). Two samples that have two alleles in common for all loci are identical twins, or, more likely, a sample that has been used twice erroneously and should be investigated.
plink --noweb --bfile plink_files/Prodem_flipped_QCed --genome --min 0.05

3.2.2.3 Preparing genotype data for analysis

Handling large amounts of genotyped data can only be performed using command line tools in UNIX environment. It is necessary to extract files from datasets, merge files, rename and rearrange files and columns, or change column separators to provide the required input that is needed for subsequent analysis for GWAS.
In order to perform these actions basic text editing programs are required, such as awk, gawk, sed or grep. https://www.gnu.org/software/awk/manual/awk.html
https://www.gnu.org/software/sed/manual/sed.html,
https://www.gnu.org/software/grep/manual/grep.html

3.2.2.4 Meta-analysis

Results GWASs from different cohorts can be combined in a meta-analysis (also called meta-GWAS), which increases statistical power, often creating combined cohorts with more than 10000 participants. However it is especially important to impute the used data of all contributing cohorts to the same sample panel. In contrast to regular association studies the significance threshold for GWAS is set a p-values$\leq 8*10^{\wedge -5}$.

Meta-analysis was performed using the free software command line tool METAL, which is available at http://www.sph.umich.edu/csg/abecasis/Metal/download/ (*Willer CJ et al. 2010*). First of all, a script summarizing the input files had to be created that METAL would use to identify the input. SEPERATOR was specified as WHITESPACE (default option), COMMA or TAB in order to be able to separate the columns of the input files.

The filename and path was given, as PROCESS <filename>, as well as the required names of the input file columns, namely MARKER <name of marker column>, ALLELE <name of allele column>, PVALUE <name of p-value column>, EFFECT <name of effect column>.

P-values were weighted by the number of individuals contributing to it using

WEIGHTLABEL N

Additional filtering was performed by ADDFILTER and then a requirement for a value of a previously specified column in particular imputation quality and MAF a. f.e. ADDFILTER MAF > 0.01 would only allow SNPS with a MAF of at least 1% to be processed.

Using REMOVEFILTERS in the script would discard above mentioned filters, so the next file could be described in detail. Upon reading in all files, the commands ANALYZE HETEROGENEITY would process the input files, allowing for heterogeneity. Addition of GENOMICCONTROL ON would automatically correct the statistics for population stratification and or unaccounted relatedness.

3.2.2.5 Visualization and interpretation of output

In order to check for inflation and visualize genome data sets, the free command line interface software package R version 3.11 was used http://www.r-project.org/. The program allows handling of large datasets and can create i.e. QQ-plots and Manhattan plots of genome wide

analyses. QQ plots were used to visualize deviations of distributed of p-values from the null distribution.

Outcome from GWAS or Meta-GWAS is usually displayed as a Manhattan plot, which is a scatter plot, which plots $-\log_{10}$ of p-values on y-axis and genomic position (Chromosomes 1-22) on the x-axis. Often an abline is added at the level of GWAS significance, for better overview.
Ideally "hits" above the threshold are accompanied by a number of hits of lower significance on the same region, (i.e. SNPs in high LD to the top hit), ruling out a chance finding.
Regional Manhattan plots were drawn using the free tool Locuszoom (*Pruim et al. 2010*) available at http://csg.sph.umich.edu/locuszoom/ , which allows plotting of data online, but can also be downloaded to perform plotting in batch locally. Regional Manhattan plots are similar to regular Manhattan plots, but "zoomed in" over a region of interest, namely top hits. In addition to genomic position on the x-axis they also display genes in the region, and show overlays f.e. with linkage disequilibrium (LD) or recombination rates.

3.2.2.6 Annotation of genotype data

Annotation of genotype data is used to predict effects of individual SNPs based on the available databases of DNA and protein sequences. In addition, it is also required due to changes in version of human genome reference panels, such as hg (human genome) 18 to hg19 of the UCSC browse rhttps://genome.ucsc.edu/ and the genome reference panel GRCH38 http://www.ncbi.nlm.nih.gov/projects/genome/assembly/grc/human/.
In our analyses, the annotation was performed using the free command line software ANNOVAR (*Wang K et al. 2010*) available at http://www.openbioinformatics.org/annovar.
Outcomes of GWAS and meta-GWAS can subsequently be subjected to pathway analysis using databases such as KEGG (*Kanehisa M & Goto S 2000*), which is available at http://www.genome.jp/kegg/pathway.html.

3.2.2.7 Prediction of functional effects of non-synonymous SNPS

In rare polymorphisms, a non-synonymous single nucleotide polymorphism (nsSNP) may affect the function of a protein provided the SNP is located inside of a coding region. The worst case scenario would be an affected cysteine residue that would affect a disulfide bond and result in a nonfunctioning protein. In case the amino acid is replaced with a one of a similar chemical structure, f.e. serine with threonine, the effects would be much less pronounced. It is also important to point out, that the position of the polymorphism is very important, the binding site of an enzyme may easily affect its function. Two web based tools were used, sorting intolerant from tolerant (SIFT) (*Ng PC & Henikof S 2001*) available at http://sift.jcvi.org/ and polyphen2 (*Adzhubei IA et al. 2010*) available at http://genetics.bwh.harvard.edu/pph2/.

The prediction was either given as affected or tolerated, in case of SIFT, or benign, possibly damaging, probably damaging when Polyphen2 was used.

4. Chapter I - Cardiorespiratory fitness as a predictor of brain aging

4.1 VO_2max and Brain Aging: Cognitive and MRI Findings in the Austrian Stroke Prevention Study

Introduction: Cognitive decline in the elderly is closely related to quality of life and disability and may precede more severe neurological deficits. The gradual loss of mental abilities is influenced by genetics as well as environmental factors, such as life-style.
In the presented study we investigate the association between age-related cognitive changes and cardiorespiratory fitness represented by maximum oxygen consumption (VO_2max) in an elderly population based cohort. Furthermore we also test if the association is influenced by age-related morphological brain changes of the vascular or atrophic type. We hypothesize that people in good cardiorespiratory condition perform better in demanding cognitive tasks.
Methods: Our cohort consists of 887 participants with a mean age of 65 ± 7 years and 55% females from the Austrian Stroke Prevention Study. Cardiorespiratory fitness was assessed by an exercise stress test on a treadmill. VO_2max was then calculated based on weight, maximum and resting heart rate. Cognition measurements included single tests, composite scores for memory, executive function and motor skills, and the g-factor, as a measure of global cognition.
Presence of lacunes, log transformed volume of white matter lesions and overall brain atrophy represented by brain parenchymal fraction were assessed by magnetic resonance imaging (MRI).
Results: We found a significant and positive association between VO_2max and the cognitive domains of memory and executive function, as well as overall cognition (g-factor) that is also dose dependent manner. Our observations were not influenced by age, sex, education, vascular risk factors and were not mediated by structural brain changes. Stratification of the results on by age, weight and APOEε4 genotype weakened the beneficial effect of VO_2max and revealed a confinement of the effect to low risk groups.
Conclusions: Our results support a protective role of VO_2max on cognition, regardless of vascular risk factors. The underlying mechanisms are not related to MRI-detectable

morphological brain changes. Interventional studies are needed to investigate the potential of improving cardiorespiratory fitness to ameliorate cognitive decline in the elderly.

4.2 Association of cardiorespiratory fitness and morphological brain changes in the elderly

Introduction: Regular physical activity and cardiorespiratory fitness improve cognition in the elderly. The interaction between fitness and morphological brain changes in the elderly is less commonly investigated. Therefor we investigate the association between fitness represented by maximum oxygen consumption (VO_2max) and white matter lesion (WML) volume for vascular changes, as well as brain parenchymal fraction (BPF) for atrophic brain changes, in an elderly population based cohort. We hypothesize that people in good cardiorespiratory condition display less morphological brain changes during aging.

Methods: Our study consisted of 715 participants of the Austrian Stroke Prevention Study with a mean age of 65 ± 8 years and 54% females. For all participants brain MRI with semi-automated measurement of WML volume (cm^3) and automated assessment of BPF (%) was available.

A maximum exercise stress test was done on a bicycle ergometer. VO_2max was calculated based on body weight, maximum and resting heart rate.

Results: We found a significant negative association between VO_2max and WML volume, but not with BPF. Additional stratification on gender revealed that the observation was only statistically significant in men.

Conclusion: Our findings indicate a protective role of fitness on age-related morphological brain changes of the vascular, rather than atrophic type. This might have clinical implication, since WML are considered a determinant of cognitive decline and disability.

5. Chapter II - Genetics of physiological and pathological vascular brain changes

5.1 Genetics of age-related white matter lesions from linkage to genome wide association studies

White matter lesions (WML) are a common MRI finding in the elderly population and contribute to the development of more severe neurological disorders including stroke and dementia.

WML are evidently highly heritable, and age and hypertension are recognized as established risk factors. Candidate gene studies have discovered variants in the Renin-angiotensin and NOTCH3 pathways to be important in the development of WML. Novel genome wide methods have confirmed some of these findings, and discovered a locus in the chromosome region 17q25, in which the *TIMR65* and *TRIM47* genes are located, to be highly associated with the disease.

In the article we review the current literature on WML and the different methods that are used to dissect the genetic background of the disease.

5.2 Genetics of subcortical vascular dementia

Dementia causes disability and death in the elderly and affects millions of elderly world-wide. Vascular dementia represents the most frequent causes, second only to Alzheimer's disease. Other than protein deposition, a series of small ischemic or hemorrhagic strokes, often subclinical progressively impairs circulation in the brain consecutively impairs function. Hallmarks of the disease include small lacunar infarcts, micro bleeds and lesions of white matter (WML), which can be observed by brain MRI.

Previous findings have reported a high heritability off white matter lesions across different cohorts, suggesting a substantial contribution of genetic variants in the variance of the phenotype.

Linkage studies investigate the tendencies of alleles in close proximity on a chromosome to be inherited together and the probability is given in centimorgans (cM). Significant observations have been made for WML at 4cM on chromosome 4 in healthy elderly and at chromosome 1q24 in hypertensive siblings. Genetic association studies using a set of candidate genes have found several pathways involved in the disease, namely lipid metabolism *APOE*, vascular function *NOTCH3* and blood pressure (genes of the renin angiotensin pathway f.e. AGT, *AGTR1, AGTR 2* and *ACE*) for WML. Micro bleeds were associated with *SORL1* and lacunes with *APOE*.

In a comprehensive genome wide association approach, new genes involved in the etiology of WML were discovered on chromosome 17, most prominently *TRIM65* and *TRIM47*.

After reviewing the literature it is evident that micro bleeds and lacunes are investigated less often than white matter lesions and often in studies of relatively small size. At least three distinct pathways are involved in the development of vascular dementia, but these discoveries explain only part of the heritability, which would indicate that other genetic factors, such as rare variants or epigenetic modifications have yet to be discovered.

5.3 Genetic variants of the NOTCH3 gene in the elderly and magnetic resonance imaging correlates of age-related cerebral small vessel diseases

Objective: White matter lesions (MWL) are a frequent MRI finding in the elderly that contributes to the development of disability and dementia. The clinical, radiological and histopathological appearance bears a high resemblance to the cerebral autosomal dominant ateriopathy with subcortical infarcts and leukoencephalopathy (CADASIL), a hereditary disease, which is caused by mutations in the *NOTCH3* gene. We hypothesized that the frequent genetic variations in the gene might also contribute to common form of age-related WML and not only the rare hereditary CADASIL.

Methods: We sequenced *NOTCH3* and its promotor and 3'untranslated region in 195 study participants with prevalent white matter lesions and in 82 participants without white matter changes.

We discovered 9 common and 33 rare polymorphisms (20 of which novel). The common polymorphisms were then genotyped in the entire cohort of 888 elderly participants with a mean age of 65.2 ± 8.0 years and (56.9%) females.

Results: For four of the common polymorphisms (rs1043994, rs10404382, rs10423702 and rs1043997) we found significant associations with WML volume as well as progression in follow ups. This finding is restricted to hypertensive participants of our cohort, and we could successfully replicate the finding in an independent cohort of 4773 stroke free elderly.

Conclusion: Our results are among the first to sequence *NOTCH3* variations in a community dwelling elderly cohort and indicate that the *NOTCH3* gene is more variable than anticipated. At least four common polymorphisms in the gene contribute to the development of age-related white matter lesions.

6. Chapter III - International collaborations and GWAS on brain aging

6.1 Heritability estimates identify a substantial genetic contribution to risk and outcome of Intracerebral hemorrhage

Objective: Intracerebral hemorrhage (ICH) represents roughly a fifth of all stroke cases and is especially hard to treat, which commonly leads to chronic disability or death. The heritability of ICH is not commonly investigated because the required large pedigree based family studies are unavailable. The advent of method allows calculation of heritability in cohort of unrelated individuals. We hypothesize, that genetic factors explain a considerable part of the variation of ICH.

Methods: 791 ICH cases and 876 controls from four different studies with mean ages of 73.9 ±10.1 and 72.4 ±7.8 contributed to our study and had genotype data available. Heritability was separately estimated for APOE, a recognized risk factor and for the remainder of the genome. The investigated phenotypes were ICH risk, admission hematoma volume and 90-day mortality.

Results: Heritability for risk of ICH was estimated at 29% for non *APOE* loci and 15% for *APOE.*

For 90-day mortality the estimation 29% for non *APOE* loci and 15% for *APOE loci.*

Lastly, we estimated the heritability of admission hematoma volume at 60% for non *APOE* loci and 12% for *APOE* loci.

Conclusion: Genetic variation accounts for a large amount of the variation in all three investigated phenotypes (risk, hematoma volume, 90-day mortality). Established risk factors such as APOE explain part of the heritability, but there are other variants that have yet to be discovered.

6.2 Meta-analysis of genome-wide association studies identifies 1q22 as a susceptibility locus for intracerebral hemorrhage

Objective: Intracerebral hemorrhage (ICH) is considered a particularly deadly manifestation of stroke. Depending on the location of the hemorrhage ICH can be classified as lobar and non-lobar ICH and might potentially have different pathologies. According to previous reports both forms show a high amount of heritability. We therefore conducted a meta-analysis of previous genome wide association studies of ICH cases and controls to identify novel susceptibility loci.

Methods: 6 different studies with participants of European ancestry contributed a total of 1545 ICH cases (664 lobar and 881 non lobar) and a control group of participants. The observations were replicated in another ICH case cohort of 1,681 (484 lobar and 1,194 non-lobar) and a control group of 2,261 of multiethnic background.

Results: We found two loci for susceptibility, the chromosomal region 12q21.1 (rs11179580, odds ratio [OR] = 1.56, $p = 7.0 \times 10^{-8}$) for lobar ICH, and chromosomal region 1q22 (rs2984613, OR = 1.44, $p = 1.6 \times 10^{-8}$) for non-lobar ICH. Our findings could be replicated for the chromosomal region 1q22 ($p = 6.5 \times 10^{-4}$; meta-analysis $p = 2.2 \times 10^{-10}$) but not for 12q21.1 ($p = 0.55$; meta-analysis $p = 2.6 \times 10^{-5}$).

Conclusion: Our findings suggest a difference between lobar of non-lobar ICH regarding susceptibility loci, and underline the heterogeneity in vascular pathology. We emphasize that ICH cases should therefore be categorized accordingly.

6.3 Meta-analysis of 74046 individuals identifies 11 new susceptibility loci for Alzheimer's disease

Objective: Previous investigations have discovered 11 susceptibility loci for late-onset Alzheimer's disease (LOAD). Nevertheless these do not explain the entire risk conveyed by genetic factors for the development of LOAD. Therefore we performed a large, two-stage meta-analysis of genome-wide association studies (GWAS) in individuals of European ancestry.

Methods: Stage 1 consisted of genotyped and imputed data (7,055,881 SNPs) from 4 previously published GWAS data sets. A total of 17,008 Alzheimer's disease cases and 37,154 controls were used to perform our initial meta-analysis on.

Consecutively we used the obtained results (11,632 SNPs) to genotype and test our observation in an independent set of 8,572 Alzheimer's disease cases and 11,312 controls.

Results: Apart from the previous the widely established *APOE locus,* 19 other loci reached genome-wide significance ($P< 5 \times 10^{-8}$) when we combined both stages of analysis. Elven of these loci represent novel associations with Alzheimer's disease.

Conclusion: Our results replicate previous findings such as the amyloid precursor protein pathway (APP) and SORL1 and CASS4 gene, as well as tau protein (FERMT2). In addition, our results point towards other pathways being relevant to the disease etiology, namely hippocampal synaptic function (*MEF2C* and *PTK2B*), cytoskeletal function and axonal transport (*CELF1*, *NME8* and *CASS4*), regulation of gene expression, post-translational modification of proteins, and microglial and myeloid cell function (*INPP5D*).

7. Discussion

7.1 Main findings

7.1.1 Chapter I

Chapter I focuses on cardiorespiratory fitness, which is assessed as the maximum oxygen consumption (VO_2max) as an important lifestyle factor influencing the brain aging process.

VO_2max and Brain Aging: Cognitive and MRI Findings in the Austrian Stroke Prevention Study

We hypothesized that participants in overall better fitness, and thus higher VO_2max display less morphological brain aging and better cognition.
To investigate morphological correlates of the brain aging we quantified white matter lesions and determined the extent of brain atrophy by using MRI. Our study consisted of 715 healthy elderly of the Austrian Stroke prevention study with a mean age of 65 ± 8 years and 54% female participants. We found a protective effect of cardiorespiratory fitness on age related vascular brain changes in the elderly. Importantly the effect was confined to male study participants.

Association of cardiorespiratory fitness and morphological brain changes in the elderly

In a next presented study we investigated if VO_2max can also be associated with better cognitive functions in the normal elderly. Indeed we found a highly significant and positive association between VO_2max and the cognitive domains of memory, motor skills and executive function, as well as overall cognition (g-factor). The effect was dose dependent and the highest vs lowest quartiles of VO_2max corresponded to test results obtained memory, executive function and global cognition results were typical for participants being 6, 7 and 4 years younger. Our observations were not influenced by age, sex, years of education, or use of

Calcium-channel antagonist and β-blockers. We observed no influence of vascular risk factors, with the exception of motor skills. No mediation process on the association between VO_2max and cognition by structural brain changes such as brain parenchymal fraction, white matter lesions and lacunes could be observed. Stratification of the results on by age, weight and APOEε4 genotype revealed a confinement of the effect to low risk groups. Interaction terms revealed a nominally significant modulating effect of BMI on VO_2max in the memory domain.

Summary

We conclude that high VO_2max is beneficial for brain aging and has a widespread effect on cognition and brain aging. High performers achieve cognitive test scores as if they were 5 years younger in all domains. According to our results life both genes and fitness plays an important role in determining brain aging, however the effect sizes of genetic factors are much smaller than that of fitness. Several genetic pathways have been implicated to influence brain aging, but the direct clinical usage is presently small. In addition these discoveries are seldom applied in clinical treatments aside from prediction through established risk factors. Recommendations for life style and fitness still requires guidelines, but the beneficial effects are much more easily apparent and applicable. Importantly we observe a beneficial effect of fitness is still present in the in the elderly population.

7.1.2 Chapter II

In the first part of chapter II, we review the literature on the genetics of age-related vascular brain changes such as white matter lesions and vascular dementia.
We also provide original research on the topic of WML in the second part of chapter II by investigating common polymorphisms of *NOTCH3* in healthy elderly.

Genetics of age-related white matter lesions from linkage to genome wide association studies

The first article reviews current literature on white matter lesions (WML), which are a common finding in elderly brains of vascular pathology that contribute to severe brain pathologies such as stroke and dementia.
In the first manuscript we describe heritability of white matter lesions to describe the variability in the phenotype explained by genetic influences in several populations. Consecutively we describe the two main hypotheses used in the explanation of common diseases. The common disease - common variant hypothesis assumes that many common polymorphisms with small effect sizes (minor allele frequency MAF ≥ 5%) explain the majority of the heritability and this hypothesis is applied in the genome wide association studies, which investigate millions of polymorphisms in unrelated subjects. This is opposed by the common diseases - rare variant hypothesis according to which rare variants (MAF ≤ 1-5%) with large effect sizes are predominantly responsible for common diseases and rare variants are usually investigated using linkage studies in families.
We list findings for significant logarithm of odds scores (LOD) obtained from linkage studies on WML (4cM chromosome 4 in healthy elderly whites and on chromosome 1q24 in hypertensive sib ships).
In the next section we describe the traditional approach of genetic association, namely candidate gene studies. The genes that could be identified in the etiology of WML so far can be grouped in different groups that govern blood pressure (*ATG, ACE, AGTR1* and *AGTR2*) and vascular function (*MTHFR, NOS1, ET1* and *NOTCH3*) and are described in detail. Lastly we list findings from novel genome wide investigations on the topic (*TRIM47, TRIM 65, WBP2, MRPL38, FBF1, UNC13D, ACOX1, MTHFD1* and *COL25A1*).

Genetics of subcortical vascular dementia

In the second article we widened (the) scope to also include small lacunar infarcts and micro bleeds, which represent other manifestations of vascular dementia in the brains of elderly. We present the available date in a similar way to aforementioned article, but pay special attention to APOE, which was the only gene investigated all three morphological hallmarks of vascular dementia.

Genetic variants of the *NOTCH3* gene in the elderly and magnetic resonance imaging correlates of age-related cerebral small vessel diseases

Dominant mutations in *NOTCH3* are causative for cerebral autosomal dominant ateriopathy with subcortical infarcts and leukoencephalopathy (CADASIL), whose clinical and radiological presentation is remarkably similar to age-related cerebral small vessel disease. Therefor we hypothesized that variations in the *NOTCH3* gene are also involved in age-related cerebral small disease. Thereupon we sequenced all 33 exons, the untranslated region and promotor of the *NOTCH3* gene in a group of 195 elderly with prevalent cerebral small vessel disease and compared the results to 82 random controls. In the course of the investigation we found 9 common and 33 rare polymorphisms, 20 of which have not been described in the literature so far. The common polymorphisms were then sequenced in the entire cohort of 888 elderly. Four polymorphisms were significantly associated with presence and progression of WML rs1043994, rs10404382, rs10423702 and rs1043997, but the observation was confined to hypertensives. We could successfully replicate the finding in a cohort of 4773 stroke free hypertensive elderly of European descent. The rare polymorphisms did not affect the CADASIL specific cysteine residues and presented a radiological profile distinct from CADASIL as well.

Summary

In summary chapter II provides a solid overview of the genetic backgrounds of white matter lesions and vascular dementia. We review the current state of knowledge and describe both the available methodology and identified genetic key players.
In the second part of chapter II we present our results from the first study investigating the effect of *NOTCH3* variation in the elderly on age-related cerebral small vessel disease unrelated to CADASIL. We successfully demonstrate the high variability of the gene, and establish an association between common variants and an increased risk of age-related white matter lesions in the hypertensives participants of our cohort.

7.1.3 Chapter III

Chapter III represents my contribution to international collaborations on brain aging. The investigated phenotypes include intracerebral hemorrhage and late onset type Alzheimer's disease

Heritability estimates identify a substantial genetic contribution to risk and outcome of intracerebral hemorrhage

In the initial article of chapter III we estimate the heritability of intracerebral hemorrhage using a novel program called GTCA, which is freely available at http://cnsgenomics.com/software/gcta/. Using this method heritability estimates can be calculated based on genotypes of common SNPs, rather than family structure (*Yang J et al 2011*). Without this approach it is very difficult to assemble a sizeable cohort, due to the high age of onset and lethality of the disease. Our cohort consists of a set of 791 unrelated cases and 876 controls once by excluding and once by including APOE4 carrier status in the model. In the models we also investigated 3 different phenotypes such as intracerebral hemorrhage as dichotomized variable, and hematoma volume as well as 90-day mortality, as quantitative variables. We report a heritability estimate of intracerebral hemorrhage risk of 29% and 15% for non APOE and for APOE loci. Hematoma volume heritability was estimated to be 60% and 12 for non APOE and for APOE loci. Lastly heritability of 90-day mortality was estimated at 41 and 10% for non APOE and for APOE loci. In summary, ICH risk, volume and progression display high heritability in our cohort, but previously identified risk factors do not fully explain the variation, which may be due to undiscovered risk loci.

Meta-analysis of genome-wide association studies identifies 1q22 as a susceptibility locus for intracerebral hemorrhage:

In the next article we try to unravel some of the genetic architecture of intracerebral

hemorrhage by identify risk factors. We perform a GWAS meta-analysis of six studies investigating intracerebral hemorrhage in European individuals. We further distinguish phenotypes based on the location of hemorrhage, as the pathology is different (lobar vs non-lobar).The cohort consisted of 1545 intracerebral cases (664 lobar, 881 non-lobar) and 1481 controls. We report a novel finding on chromosome 12q21.1 (rs11179580) for lobar intracerebral hemorrhage and on chromosome 1q22 (rs2984613) for non-lobar intracerebral hemorrhage.

We furthermore tried to replicate these findings in a second cohort of multi-ethnic origin of 1681 cases (484 lobar 1194 non-lobar) and 2261 controls. In the second cohort 12q21.1 was no longer significantly associated with lobar ICH. These results underline the assumption that the location of the intracerebral hemorrhage is highly important and that the two lobar vs. non-lobar might be two distinct subtypes. Furthermore we provide novel insight in the genetic architecture of intracerebral hemorrhage, which might be of interest to subsequent investigations.

Meta-analysis of 74046 individuals identifies 11 new susceptibility loci for Alzheimer's disease

The last section of chapter III we provide a meta-analysis to identify susceptibility loci for late onset Alzheimer's disease. The analysis was performed in two consecutive steps. In the first step, data from four consortia, namely the Alzheimer's disease Genetic Consortium (ADGC), the Cohorts for Heart and Aging Research in Genomic Epidemiology (CHARGE) Consortium, the European Alzheimer's Disease Initiative (EADI) and the Genetic and Environmental Risk in Alzheimer's disease (GERAD), was collected. These consortia previously published GWAS on late onset Alzheimer's disease in individuals of European ancestry in a total of 17008 cases and 37154 controls. This lead to the discovery of 14 regions with genome wide significance ($p < 5 \times 10^{-8}$), five of which represent novel associations (*HLA-DRB5–HLA-DRB1, PTK2B, SORL1, SLC24A4-RIN3 and DSG2*).

The second step consisted of genotyping the discovered SNPs (n= 211632) that reached a threshold of $p < 1 \times 10^{-3}$ in the first step. These were used in second association in a sample of

European descent of 8572 cases and 11312 controls.
All but two regions implicated in the first step (in vicinity of *CD33* and *DSG2*) reached genome wide significance in the second cohort as well and were successfully replicated.
The combined stage 1 and 2 additionally lead to the discovery of 7 new loci (*INPP5D, MEF2C, NME8, ZCWPW1, CELF1, FERMT2, and CASS4).* These findings not only confirm previous findings from GWAS and candidate gene studies (*SORL1*), but also implicate completely new pathways involved in the etiology of Alzheimer's disease of the late onset type.

Summary

We demonstrate that the risk for ICH its severity and progression are highly heritable. Current knowledge of risk factors does not fully explain our findings which points towards unidentified variants that potentially operate in unknown biochemical pathways.
In a joined effort we also discover two novel genetic regions involved in the etiology of ICH. These are located on chromosome 12q21.1 (rs11179580) for lobar ICH and on chromosome 1q22 (rs2984613) for deep ICH.
In the last article we replicate previous findings and successfully identify 11 novel loci involved in late onset Alzheimer disease. Some of our results point towards pathways the have not yet been suspected to influence Alzheimer's diseases.

7.1.4 Summary

In this thesis I present original research on the environmental aspect of brain aging through the pursuit of an active healthy life-style as well the genetic background of brain aging phenotypes.
In summary the results displayed in chapter I, II, and III clearly highlight the fact that both genetic and environmental effects are important in brain aging. We also observe the effect sizes of common genetic risk factors on brain aging is much smaller when compared to the large effect of environmental factors such as cardiorespiratory fitness. Clearly subsequent

genetic investigations might yield more genes involved in the etiology of age related diseases such as dementia and stroke. The identification of new genes may point towards the involvement of unsuspected or even undiscovered pathways in the pathogenesis of these diseases. Nevertheless the discoveries in genetic research so far have little application beyond the prediction of risk profile. Even undisputed findings of common variants such as APOE genotype on Alzheimer's disease have not yet lead to the development of medical compounds. It is therefore particularly important to develop life-style recommendations and minimize the environmental risk, which will likely include regular exercise to maintain a physically active life-style.

7.2 Methodological issues

7.2.1 Definition of phenotypes

Physical activity, exercise and fitness

- Physical activity (PA)

PA is defined as bodily movement by skeletal muscle in order to perform everyday activities. Depending on life style choices the amount of PA often varies heavily between individuals. It can be categorized in a number of ways, for example whether PA is performed during work or leisure time, or by evaluation of the time distribution (weekends or seasonal changes). Another distinction can be made based on the intensity of the activity (low, moderate or high f.e) or using the scale of metabolic equivalents (MET), which categorizes energy cost of physical activities based on energy consumption.
PA can be assessed subjectively by self-reports using either questionnaires or diary entries. In epidemiological studies wearable activity monitors that count steps (pedometer) or acceleration (accelerometer) are frequently used to get less biased and more objective assessment of PA. More precise measures for physical activity can be obtained by using as direct or indirect calorimetry or doubly labeled water (*Freedson P et al.2012*), a method which is however not practicable in epidemiological setting. It comes as no surprise that

the objective measures of PA often only reveal low to moderate correlation with self-reported questionnaires (*Prince SA et al. 2008*).

- Exercise

Similar to physical activity, exercise uses the same muscular activity to generate movements. Exercise can therefore be described as a subcategory of physical activity, which explains its frequent synonymous use with physical activity. In contrast to PA, exercise is planned, structured and repetitive with the intention of maintaining or improving skeletal movement and physical fitness. Exercise is usually categorized in type, (aerobic, anaerobic or flexing), intensity, duration and frequency.
Both physical activity and exercise correlate with physical fitness to some extent. Physical fitness itself, however, does not describe movements itself, but rather the capability of an individual to perform everyday tasks and activities. In addition to the bodily function, physical fitness also describes resistance to diseases associated with sedentary life-style (such as cardiovascular disease, obesity, type 2 diabetes), which are also referred to as hypokinetic diseases.

- Cardiorespiratory fitness (CRF)

CRF in particular refers to the circulatory and respiratory fitness capabilities to supply sufficient amounts of oxygen to the skeletal musculature and regulate blood pressure during exercise. Cardiorespiratory fitness is most often assessed as the maximum oxygen consumption, also known as peak oxygen uptake or VO_2max, which is usually obtained during incremental exercise tests on an ergometer or a treadmill. It can be directly measured by analysis of the inhaled and exhaled air during maximal tests. If direct measurements or maximal tests are not applicable and would be detrimental to patient's health VO_2max can be indirectly estimated from submaximal exercise tests using the Uth–Sørensen–Overgaard–Pedersen formula, which takes body mass, maximum and resting heart rate into account. Regardless of the assessment VO_2max is either given in absolute ($ml^{-1*}min^{-1}$) or relative ($ml^{-1*}min^{-1*}kg^{-1}$) values.

Correlation between PA and CRF often yield conflicting results, either claiming no correlation (*Tager IB et al. 1998*) or moderate correlation (*Bowles HR et al. 2004*), which can have several reasons that will be discussed. As described previously, cardiorespiratory fitness can be regarded as an outcome of frequent physical activity and exercise. In addition, it may also reflect long term dedication to regular bodily exercise, as well as a healthier life-style (f.e. smoking abstinence).
Secondly the type and intensity of physical activity and exercise seems to be a key determinant of improving VO_2max (*Wenger HA & Bell GJ 1986*). High interval training and other forms of aerobic exercise are more effective at raising VO_2max in clinical trials than moderate continuous training in case of heart failure patients (*Wisløff U et al 2007*).
In general, the evaluation of CRF using VO_2max is a gold standard in determining fitness and allows easy comparison with other studies due to the wide spread of its application.
In addition, due to submaximal testing protocols the method can also be applied to elderly and frail study participants.

Structural correlates of the aging brain

- White matter lesions(WML)

The most frequent finding in brain magnetic resonance imaging of elderly population are WML. They appear as that appear as white matter hyperintensities (WMH) on FLAIR and T2 weighted scan sequences.
These lesions can be assessed qualitatively based on their severity using the Fazekas scale or their volume can be quantified. Quantification can either be performed visually by operators, automatically or semi- automatically, with operators outlining and evaluating computer-generated images. Especially at low volumes of WML the different methods yield incomparable results (*Olsson E et al. 2013*). Additionally, grading by personal investigators also inevitably introduces bias and inter-rater variability, which needs adjustment. Recent reports claim, that correlation between qualitative grading using the Fazekas scale and quantitative assessments of white matter lesions is high (*Valdés*

Hernández Mdel C et al. 2013). Secondly, the assessment of white matter lesions is also dependent on the MRI sequence that is applied. The results can be comparable in case of FLAIR and T2-weighted images (*Piguet O et al. 2005*).
Depending on their origin in the brain WML are often dichotomized as periventricular or deep (subcortical) lesions, depending on their location (*KIM W et al. 2008*).
The underlying assumption is that the two types have different pathomechanisms. Deep WML are considered to be more closely related to small vessel disease and high blood pressure, whereas periventricular lesions are more determined by chronic hemodynamic insufficiency, smoking and hypercholesterolemia (*Khan U, et al. 2007*). How deep and periventricular WML are classified varies between studies and whether the classification is suitable is subject to some debate among current studies.

- Lacunes

Lacunes are an entity of cerebral small vessel disease and appear as cavities filled with cerebrospinal fluid in FLAIR imaging. This is the result of a previous small ischemic incident in the region. Lacunes are not as commonly investigated in MRI investigations, which has several reasons. The terms lacune, lacunar stroke and lacunar infarct and lacunar infarct syndrome are used interchangingly, but have different meanings, which are being debated in current literature (*Wardlaw JM 2008*). Lacunar stroke refers to a subset of ischemic stroke caused by occlusion of the brain arteries with the typical stroke presentation. Lacunar acute syndrome refers to patients that display the same symptoms, which were not yet confirmed by imaging methods. Lacunar infarcts, which are also likely to be clinically silent, are generally thought to precede the development of lacunes. These lacunar infarcts show high amounts of similarity in appearance with WML in FLAIR imaging. This can potentially lead to misclassification unless diffusion tensor imaging is applied. The initial infarct can progress into the liquid filled cavity referred to as lacune, but this is not necessarily the case (*Potter GM et al. 2008*).
Naturally the location of the lacune is important in order to predict cognitive outcomes (*Benjamin P et al. 2014*), but this is not always taken into account. Instead presence of

lacunes or number lacunes are often used in epidemiological investigation.

- Intracerebral hemorrhage (ICH)

ICH is a subtype of stroke that accounts for approximately 15% of the total stroke cases and shows particularly dire outcomes for patient recovery (*Ikram MA et al. 2012*). It is usually detected using computer tomography (CT), but magnetic resonance imaging can also be applied (*Fiebach et al 2004*). Before the advent of novel imaging techniques ICH was often difficult to classify and often identified as ischemic strokes or subarachnoid bleedings.

Two forms of ICH are usually distinguished depending on location of the hemorrhage (*Biffi A, et al. 2011*). Lobar and non lobar (also referred to as deep) ICH differ in terms of etiology and risk factors. Lobar ICH is strongly associated with cerebral amyloid angiopathy, age and alcohol consumption (*Matsukawa H et al. 2012*). Deep ICH is associated with high blood pressure vasculopathy *(Ikram MA et al. 2012*).

ICH is additionally classified as primary or secondary depending on the causes (*Sahni R & Weinberger J 2007*). The more common primary form is caused by hypertensive arteriolosclerosis and amyloid angiopathy in over 80% of the cases (*Sutherland GR & Auer RN 2006*). Secondary causes for ICH include vascular malformation, previous ischemic stroke that transforms or presence of intracranial tumors.

Cognition

Cognitive assessment in the elderly is important for identifying impairments, as well as assessing their severity and predicting their outcome (*Woodford HJ & George J 2007*).

- Single tests

There is a wide range of tests that can be applied. These tests can take anywhere from a minute to several hours to complete. The outcomes of cognitive tests can be

dichotomous or continuous outcomes. Clinicians often tend to rely on single tests in clinical settings, rather than specialized ones for reasons of feasibility (*Cullen B et al. 2007*). One such method is the mini mental state exam (MMSE), which allows diagnosis of cognitive impairment in minutes.

In general, there is no uniform single test to evaluate every cognitive aspect. To accurately determine the cognitive status of an individual an array of single tests is required. Widely used examples include the Wisconsin card sorting test (WCST) or the digit span backwards (DSB), or Stroop test. Many cognitive tests are language-based, which causes differences in results across cultures (*Reilly D 2012*). Several cognitive tests, such as the cognitive abilities screening instrument (CASI) (*Teng EL et al. 1994*) or the Rowland Universal Dementia Assessment Scale (RUDAS) (*Naqvi RM et al. 2015*) are available that allow cross-culture comparison. This is not sufficient to fully reduce bias created by participants with low education, and non-native English speakers (*Jones RN 2001*).

Due to cultural and regional differences in the performances in cognitive test, z-scores are often used in standardized testing. It is a statistical method in which the mean values of the test results is subtracted from the obtained test value, and then divided by the standard deviation.

$$z = \frac{x - \mu}{\sigma}$$

- Domain scores

  Individual tests scores from single tests are often grouped into domains for specific functions, since they fall into same categories of cognitive function. There are different ways to group cognitive abilities into domains. In the research presented in this thesis a system with three domains, namely memory, executive function and motor skills was used using the z scores from individual tests.

- Global cognition, g factor

At the highest level, a single common factor of cognitive ability can be calculated using all domain scores. This summary measure is referred to as the general factor, g-factor or g (*Deary IJ et al. 2010*).

In summary, definitions of phenotypes investigated in this thesis are demanding and their assessment are highly heterogeneous among studies. International efforts to harmonize these phenotypes are of major importance in the field and would allow cross study comparison.

7.2.2 Genotyping

Human sequence variation

Human inter-individual genomic variation is in the range of 0.5% (*Levy S et al. 2007*). There are different types of genetic polymorphisms that require different approaches and are not equally well represented in current research.

- Types of polymorphisms

There are two main types of sequence variation.
Single nucleotide polymorphisms (SNP) are the most commonly investigated type. They are biallelic variations in DNA that affect a single nucleotide (A, T, C, or G). For each genetic locus two alleles are present, one from each parent. The less frequent allele in any given population is called the minor allele and its frequency is referred to as the minor allele frequency (MAF). Large databases, such as dbSNP of human in which collective data from sequencing is stored report more than 140 million in the human genome so far (http://www.ncbi.nlm.nih.gov/SNP/ ; *Sherry ST et al. 2001*).

Structural variation refer to polymorphic DNA regions the size of 1kb and larger (*Freeman JL et al. 2006*). These variations can be copy number variants (CNVs) such as insertions or

deletions of DNA segments. In contrast structural variations can be copy neutral variation, which includes inversions and translocations of entire sections.
Structural variations are very common (*Itsara A et al. 2009*), but nevertheless underrepresented in current research. This was caused by the inability of commercial genotyping arrays to assess structural polymorphisms. Nevertheless databases on structural variation exist and more than 10000 are known to date (*Conrad DF et al. 2010*).

Genotypes can already be very accurately determined by products from companies such as Illumina, Affymetrix, Life Sciences or Agilent. For studies on genetic epidemiology, gneotptyping is usually performed using genome hybridization chips, which target a predetermined set number of polymorphisms. These polymorphisms (usually SNPs), are easy to genotype and frequent enough in the genome making them an ideal research target.
Whole exome (WES) or genome sequencing (WGS) in contrast allows complete sequencing of a person's exome or genome. These methods capture every polymorphism in the genome, both SNPs and structural variations. Although the benefit might be immediately apparent the data can easily be of use for later analyses. In any case the method should be chosen based on research question, storage capacity and cost efficiency. The genotyping technology that was used in the conduct of the research presented in this thesis are further explained below and in the methods section in chapter **3.2.2.1.**

Genome Hybridization chips

Genome hybridization chips use a strategic set of single DNA strands corresponding to the SNPs that are mounted on a surface. Labeled DNA from the sample will then hybridize to a complementary strand and can be detected. The Affymetrix Genome-Wide Human SNP Array 6.0 was used in the conduct of this thesis and allowed detection of ~900000 SNPs and 90000 CNVs. Genome hybridization chip technology typically use variations that span the entire genome, can also be customized to fit specific investigations.
One such example is the exome chip by Illumina, which is available at http://www.illumina.com/products/infinium_humanexome_beadchip_kit.html. It allows the

investigation of exonic (protein coding) markers that have been established as risk factors for common diseases, such as cancer, diabetes type 2, cancer and metabolic and psychiatric disorders.

Imputation of genotypes

A common method to increase statistical power of genome hybridization chips is to impute the dataset. SNPs are inherited in blocks, so called haplotypes, with high linkage disequilibrium (LD) which can be interpreted as likelihood to be jointly inherited together. The imputation process allows statistical "bootstrapping" of genotypes from already known haplotypes in a given population. As an example, 1-2Million SNPs are genotyped using a hybridization chip and ~ 30 Million SNPs are imputed (*Scheet P & Stephens M 2006*). The requirement for the imputation are reference haplotype for the respective population. These haplotypes or haplotypes maps are available for many different populations from the two comprehensive catalogues of human sequence variation.
These catalogues are called the Hapmap Project, and the 1000Genomes Project which can be found at http://hapmap.ncbi.nlm.nih.gov/ (*International Hapmap Consortium et al. 2005***)**, at http://www.1000genomes.org/ (*1000Genomes Project Consortium et al. 2010*). Both catalogues are free and are frequently used in genome wide association studies. They allow easy comparison similar sets of polymorphisms and overall increase the power and likelihood to observe causal associations.

7.2.3 Genetic association studies

Genetic association is used to investigate associations between one or more genetic polymorphisms and a trait (*Cordell HJ & Clayton DG 2005*). The trait or phenotype can be a continuous variable (f.e. height) or a dichotomous outcome (diseases vs non diseased). Importantly, genetic associations are possible because human populations share common ancestry. There are several possibilities in the conduct of genetic associations, namely indirect and direct associations. Association can occur directly with a causal variant or with a surrogate

of the causal variant. The two approaches used in this thesis are explained below.

Candidate gene studies

Candidate gene approach in genetic association investigates association between genetic variants in specified set of genes with a specific phenotype. The set of genes is usually selected based on available knowledge of biochemical pathways the phenotype is related to. Candidate gene approach is subject to some criticism, because the selection of genes also introduces a certain amount of inevitable bias. In addition the results from candidate gene studies do not necessarily replicate well in follow up studies (*Patnala R et al. 2013*).

Genome wide association studies

Current research focus lies on the genome wide association studies (GWAS), in which millions of variants all over the genome are associated with a certain phenotype. In contrast to candidate gene approaches, GWAS are not driven by hypothesis and represent an unbiased approach. Since their advent in 2005 GWAS could successfully identify thousands of risk factors for many diseases. (*Haines JL et al. 2005*). There are two distinct hypothesis regarding the genetic architecture of multifactorial diseases.

- The common disease–common variant hypothesis suggests that large numbers of common genetic variants additively contribute to disease susceptibility of complex polygenic diseases (*Botstein D & Risch N 2002; Pritchard JK & Cox NJ 2002*). Alleles are referred to as common, when the frequency of the least common allele in a population, the minor allele frequency (MAF) is ≥5%.
- The common disease-rare variant hypothesis predicts that very few rare (MAF of ≤1%) with larger effect size represent the major contributors to the disease (*Schork et al. 2009*). Variants implicated in GWAS usually fall in the category of common disease – common variant. Association of common diseases with rare variants demands sufficient statistical power, which is only achieved by very large sample sizes (*Peloso*

GM et al. 2014).

Both of these hypothesis are illustrated in figure 1. Common alleles are easy to genotype and detect through statistical methods on the right side. Common alleles with low effect size represent the current interest of genome wide association studies in the bottom right corner. Common alleles with high impact on common diseases in the top right are very rare by themselves, but represent ideal targets for research of pharmaceutical compounds. Rare alleles with low effect, which are displayed on the left side of rare which are hard to pinpoint, but ultimately not particularly interesting. Whereas the rare alleles with high effect size typically follow a Mendelian type of inheritance, similar to hereditary diseases. The low frequency (MAF<5%) with intermediate effect in the middle of the figure are likely a future research topic.

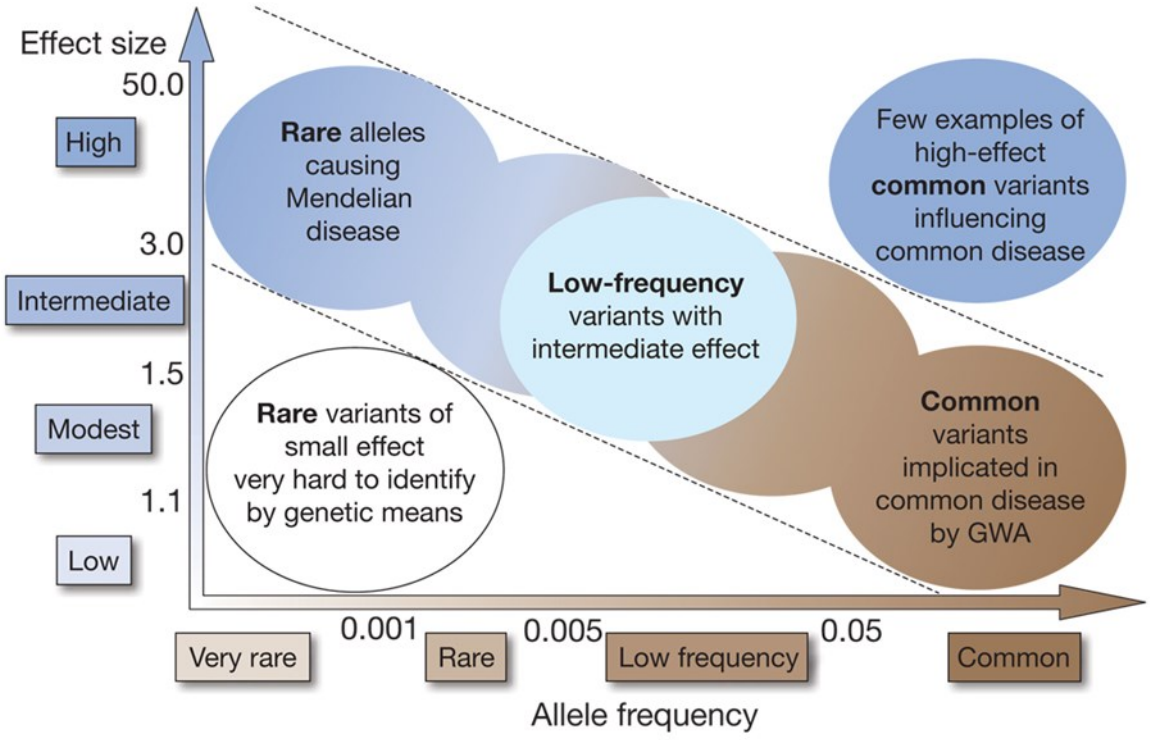

Figure 1 Relationship of allele frequency and effect sizes (relative risk) of risk variants in genetic association

studies. Image reproduced from Manolio et al. 2009

It is important to emphasize that candidate gene approach and genome wide association studies are both equally valuable and complement each other. GWAS can be used as a means to replicate previous findings from candidate gene studies. This is the case for a common variant in *AGT*, which was identified as a risk factor for WML in a candidate gene investigation (*Schmidt R et al. 2001*) and was recently replicated in GWAS. (*Fornage M et al. 2011*). Vice versa, if a specific chromosomal region or gene has been implicated in a GWAS, candidate approach can be used to sequence the region or gene of interest to potentially find a causal variant.

Missing heritability

Despite the valuable insight into many diseases provided by these genetic association studies on common variants, only a small amount of the heritability is explained by them. The remainder of genetic contribution to the variation in an observed phenotype is referred to as the missing heritability (*Eichler EE et al. 2010*). Several factors are suspected to contribute to missing heritability in genetic epidemiology.

- Rare variants (Frazer KA et al 2009; *Lee S et al. 2014*)
  Genetic association of rare variants requires large sample sizes and are very often specific to populations.
- Structural variations (*Stankiewicz P & Lupski JR 2010*)
  Research on structural variation was hindered for several years because they used to be difficult to genotype with previous genome hybridization chip technology.
- Epigenetic inheritance (*Slatkin M 2009*).
  Epigenetics represents a relatively novel field of research. Current epidemiological studies mainly focus on histone methylation, not acetylation patterns.
- Interactions between genes and environment (*Kaprio J 2006*).
  The sensivity to an environmental risk can be heritable as well and individuals might

not be similarly afftected to the exposure.

- Epistatic effects, gene-gene interaction(*Zuk O et al. 2012*)
  Epistatic effects occur when effects of one gene depend on one or more modifier genes. It can occur between different loci but also at the same locus between the two alleles of a gene. Epistasis is expected to lead to inflation of the heritability estimates.

7.2.4 Interpretation of findings

Genetic association studies must be carefully interpreted (*Campbell H & Rudan I 2002*). A positive association may implicate that the variant by itself is causally involved in the disease. It could also mean that it is in LD with a yet unidentified causal variant (indirect association), the association is biased due to population admixture or substructure or it is a false positive chance finding.

Causal Association

The most valuable and least likely finding in genetic association studies is the discovery of a causal mutation. In this case, the variant is directly responsible for the development of the disease and alters the functionality or expression of associated protein. Replication of the findings in a different unrelated cohort provide further support for a true association, however it can be also due to LD. Failure to replicate might be caused by differences in the phenotype definitions or genetic and environmental factors that interact with the causal variant, which are simply absent in the follow up investigation. Pinpointing the causal variant is difficult and possibly expensive and usually done by fine-mapping experiments using custom arrays or sequencing (*Wang K, et al. 2010*). Nevertheless, the final proof for causality requires experimental studies. These are preferably performed in small animal models that allow high throughput screening. Possible candidates include the rodent *Mus musculus*, the nematode *Caenorhabditis elegans*, the fruit fly *Drosophila melanogaster*, and the fish *Danio rerio*.

Marker due to linkage disequilibrium

Much more commonly, an observed significant association between the polymorphism and the phenotype is due to in LD and are very likely to be jointly inherited in blocks of haplotypes. The physical range of LD is subject to debate from a few kilobases (kb) to more than 100 kb, but it differs from one population to another and from the chromosomal region as well as between the sexes (*Reich DE et al. 2001*).

False positive or negative due to bias

Genotype phenotype association can also yield false positive findings due to population stratification or population admixture. In case of population stratification systematic differences in the frequencies of genetic variants can be observed due to different ancestry (*Cardon LR & Palmer LJ 2003*). Population admixture occurs when individuals from formerly separated populations start mating. In both cases, this may can lead to false positive findings or masking of genuine associations (false negative) especially in large GWAS (*Marchini J et al. 2004*). Since stratification cannot be completely ruled out by careful planning of studies (*Freedman ML et al. 2004*), it is important to adjust accordingly. This can be done using statistical methods by investigating genomic controls (*Cardon LR & Freedman ML 2004*) or performing principal component analysis (*Price AL et al. 2006*).

False positive due to chance

Finally a positive finding can occur simply by chance, as vast amounts of polymorphisms in large cohorts are investigated at the same time. Due to that reason GWAS are adjusted for multiple testing using a Bonferroni correction, and generally p-values below 5x10^-8 are considered significant due to an assumed million independent variants (*Johnson RC et al. 2010*).

Guidelines for genetic epidemiological studies

Several findings from genetic association studies could not be replicated due to unaccounted population stratification (*Colhoun HM et al. 2003*). However, successful replication of the genotype-phenotype associations is needed in order to avoid reporting of false positive findings in the (*NCI-NHGRI Working Group on Replication in Association Studies et al. 2007*). Therefore Guidelines on the conduct of studies have been established for candidate gene studies (*Tabor HK et al. 2002*) and GWAS (*Barsh GS et al. 2012*). Candidate gene association studies require evaluation of biological plausibility and dose-response relationships.
For GWAS controlling for multiple comparisons, population stratification, relatedness by statistical methods are required (*Bouaziz M et al. 2011*). In both cases critical selection of cases and controls and overall homogeneity is a requirement to ensure replication of results.

Another key factor in epidemiological research is the inference of causality from established associations. A cross sectional study observing elderly citizen will likely find an association between disability an depression, but from this observation it is impossible to determine if one is causal to the other. Longitudinal observations can help understanding the temporal relationship between outbreak and progression of diseases and can provide further evidence. Clinical trials on the other hand can establish causality but findings from observational studies are required for their design and conduct, which highlights their complementary role (*Guralnik JM & Kritchevsky SB 2010*).

7.2.5 Bias in epidemiology

In any epidemiologic study several sources of bias have to be considered and are discussed below (*Rothman J 1998; Delgado-Rodriguez M & Llorca J 2004)*.

Selection bias

Selection bias is introduced into study through subject selection in a way that influences both the exposure and outcome. In case control studies the control groups might not be represent

healthy counterparts of the case group. This can be considered as a systematic failure to equally represent all classes of cases or people that are supposed to be represented in a sample (some are more likely to be included than others). In prospective studies loss to follow up bias can occur if one group is more likely to develop the outcome. Another example of selection bias is related to the response rate of the invitations. Participants often differ from non-participants due to the "healthy volunteer or healthy worker" effect. People who respond to invitations are, in general, more concerned about their well-being and lead different life-styles and are more likely to be female. A study can be referred to as community dwelling if the response rate to the randomized invitations is >50%. In the case of the ASPS this was not achieved and the correct term would be population based.
Selection bias is not be confused with representativeness. If the selected study population is not representative for a general population but instead represents a certain subgroup the results cannot be generalized. They are nevertheless valid in the observed subpopulation.

Information bias

Secondly, information bias can occur when measurement flaws influence outcome, covariates or exposures between comparison groups. It occurs when sensitivity and specificity of the measurements are not insufficient. This can result in misclassification of exposed/affected/diseased participants. Misclassification can occur non-differential if the error rate is similar in all observed groups. In case of differential misclassification the error frequency differs among the observed groups. Examples for information bias are underreport of information, f.e because the use is considered morally condemnable (alcohol consumption or drug use). Other examples include lack of memory to remember disease or exposition history in recall bias. Study participants that willingly collaborate with investigators in a way they perceive to be helpful can also cause reporting bias. Opposed to this, observer bias can occur in situations, in which the interviewer is not blinded and unwillingly influences data recording. A certain amount of information bias can also occur by random error and is ultimately unavoidable.

Confounding

Lastly, confounding is an attribute that refers to variables that correlate with both dependent and independent variable. Ideally this is taken care of at the design stage of research. When performing analysis, confounders have to be taken into consideration and added to statistical models to minimize their effect. The remainder of bias after confounding has been taken into account as much as possible by adjustment is referred to as residual confounding (*Rothman KJ et al 2008*). Some investigations further differentiate between and refer to confounding caused by errors in measurement as residual confounding, and confounding caused by variables not considered or omitted in a model as unmeasured confounding (*Fewell Z et al 2007*).

7.3 Implications

7.3.1 Chapter I

VO_2max and Brain Aging: Cognitive and MRI Findings in the Austrian Stroke Prevention Study

Our obtained results might have important implications in non-pharmaceutical treatment and preventive medicine in the elderly.

- We demonstrate the importance of having a high VO_2max, which was associated with higher cognitive scores in memory, motor skills and executive function, as well as overall cognition (*Colcombe S & Kramer AF 2003*), which is in line with literature.
- We also report a dose dependent association between VO_2max and cognition, which translates to participants from the highest quartile of VO_2max achieving results similar to being 4, 6, and 7 years younger for global cognition, memory and executive function. Taking the rising incidence of Alzheimer's disease, which is expected to double every 5-6 years, into consideration, our observed effects are large and potentially relevant for preventing or delaying dementia (*Mayeux R & Stern Y 2012*).

- To our knowledge, we are the first study to report mediation effects of white matter lesions, lacunes and brain parenchymal fraction on cognitive measures. Other age-related brain changes such as caudate nucleus volume have been identified as mediators between fitness and cognitive function in older adults (*Verstynen TD et al. 2012*).
- These findings underline the importance to maintain a healthy and active life-style, including maintenance of low BMI, being physical active and not smoking, which are associated with higher cardiorespiratory fitness throughout life (*Jackson AS et al. 2009*)
- A recent longitudinal study in children found no protective effect of cardiorespiratory in midlife. The authors suggest that children with better cognitive functioning instead choose a healthier life-style (*Belsky DW et al. 2015*). In healthy older adults cardiorespiratory fitness was associated with preservation of cognitive function over a 6-year period (*Barnes DE et al. 2003*).
- According to a recent study, episodic memory and executive function are attenuated by cardiorespiratory fitness in older, but not younger adults (*Hayes SM et al. 2014*). Taking these findings, as well as our own into consideration there seems to be no diminishing effect of fitness interventions in the elderly. The type intensity and frequency of exercise that is required to ameliorate cognitive decline and brain aging should be studied in future investigations would benefit from guidelines.

<u>Association of cardiorespiratory fitness and morphological brain changes in the elderly</u>

- We demonstrate the importance of having a high VO_2max, which is associated with reduced volume of WML in the elderly, which is in line with previous observations (*Tseng BY et al 2013*).
- Interestingly our findings were confined to male participants. Sex differences in cardiovascular aging and adaptive responses to physical activity have been implicated in literature (*Parker BA et al. 2010*).

7.3.2 Chapter II

Genetic variants of the *NOTCH3* gene in the elderly and magnetic resonance imaging correlates of age-related cerebral small vessel diseases

- We present a study to investigate the relationship between the *NOTCH3* gene and age-related morphological brain changes in a healthy elderly cohort, which was a novelty by the time of publication.
- We found no mutations affecting cysteine residues, which have been reported to be directly responsible for CADASIL (*Chabriat H et al 2009*).
- We reported a high variability of the gene and found four common SNPs that are in linkage disequilibrium, that enhance the risk of presence and presence of white matter lesions. This led us to hypothesize that common variations in *NOTCH3* substantially influence the risk of sporadic small vessel disease.
- The finding of the SNP with the strongest effect on white matter lesions in our cohort (rs10404382), could also be replicated in 4773 participants from 6 independent studies on stroke free elderly. When we stratified our results we noticed a confinement of the effect to hypertensives. A recent review suggests that risk factors such as smoking and hypertension may exacerbate white matter lesions in those with an innate genetic vulnerability (*Wardlaw JM et al. 2013*), which explains some of our findings.
- Contrary to our observations, a recent study could not replicate our findings for the SNPs associated with white matter lesion presence and progression, even adjusting for hypertension (*Rutten-Jacobs LC 2015*). However, this is likely due to the cohort being stroke patients compared to the healthy cohort in our investigation and differences in the assessment of WML.
- In addition we detected 9 rare non-synonymous SNPs in participants with extended white matter lesions, but they did not display clinical presentation specific for CADASIL. When we applied bioinformatics tools to these SNP we predicted 3 of them be functional.

7.3.3 Chapter 3

Heritability estimates identify a substantial genetic contribution to risk and outcome of intracerebral hemorrhage

- In the third chapter we first investigate the heritability of intra-cerebral hemorrhage to predict its risk and severity. We have used a novel approach to estimate heritability in a large case-control study setup of unrelated individuals and are the first to perform this type of investigation for ICH. The obtained results for the heritability confirm previous findings from twin studies (*Alberts MJ et al. 2002*).
- According to our findings, *APOE* variants differentially affect intracerebral hemorrhage depending on the location (lobar or deep), which has been reported (*Woo D et al 2002*).
- The differences in heritability are likely caused by different pathogenesis, lobar ICH being related to cerebral amyloid angioptathy opposed to the deep ICH being related to hypertension. As a consequence the deep ICH, which shows less heritability is more influenced by the environment. When genetic risk factors established from former investigations for ICH and hypertension were removed, the heritability estimates remained the same, which might be caused by missing heritability.
- The analysis of the heritability for the hemorrhage volume was estimated at 60%, but when the location was considered in the estimation the effect vanished, likely due to location being a key determinant for ICH volume (*Falcone GJ et al. 2013*).
- Lastly we estimated that 90-day mortality for ICH has a high heritability of 40% in our dataset. This suggests that unidentified genetic variants play a major role in the outcome of ICH.

Meta-analysis of genome-wide association studies identifies 1q22 as a susceptibility locus for intracerebral hemorrhage

- Consecutively to the estimation of the heritability of ICH we performed a meta-

analysis on 3223 cases and 3725 controls and identified and replicated a novel susceptibility locus for deep ICH at chromosome 1q22. Importantly, this region was implicated as a suggestive finding in a GWAS on white matter lesions in healthy adults (*Fornage M 2011*).

- This highlights the differences in pathology between deep and lobar ICH and suggests that deep ICH shares a pathogenesis similar to cerebral small vessel diseases, which has been proposed (*Rost NS et al. 2010*).
- The associated variants are located in a region harboring the *PMF1* and *SLC25A44* genes. Polyamine-modulating factor (PMF1) is a core protein, involved in chromosome alignment during mitosis and catabolic polyamine metabolism. Polyamines metabolites such as acrolein appear to be present in the plasma of stroke patients in higher concentrations, which is why they have been proposed as a predictive biomarker (*Tomitori H et al 2005*; *Igarashi K & Kashiwagi K 2011*).
- A second region 12q21.1 was suggestive for lobar ICH, but did not reach the required p-value threshold even though the effect was observed in all discovery cohorts, which makes the region potentially interesting for subsequent investigation.

<u>Meta-analysis of 74046 individuals identifies 11 new susceptibility loci for Alzheimer's disease</u>

- Here we present the findings from a two stage approach meta-analysis to illuminate the genetic underlying of late onset Alzheimer disease. We could successfully replicate previously discovered variants and discovered 11 new susceptibility loci.
- The most significant observation was in the *HLA-DRB5–DRB1* region, which code for major histo-compability complex II. Interestingly, the major histocompatibility complex has been assumed to be determinants in Alzheimer's disease ago (*Candore G et al 2004*). In addition, there has been a long standing debate over the involvement of microglial cells, which express the major histocompatibility complex II on their surfaces of microglia, in the formation of Aβ plaques (*Weitz TM & Town T 2012*).
- The next strongest signal was found in the *SORL1* gene, and plays a major role in the

etiology of Alzheimer's disease as SORL1 directly binds to the amyloid precursor protein (APP) and prohibits the accumulation of Aβ aggregates (*Rogaeva E et al 2007*). Unsurprisingly, it has been implicated to be a key player in both the familial (*Pottier C et al. 2012*) and late onset sporadic type of Alzheimer's disease (*Vadarajan BN et al 2015*).

- The third new locus lies in *PTK2B,* which encodes a protein kinase that plays a crucial role in the long-term potentiation and depression of the hippocampus and thus the creation memory (*Hsin H et al. 2010*) and has been found as a suggestive locus in a much smaller investigation on Alzheimer's disease (*Kamboh MI et al. 2012*). In addition, *PTK2B* is also frequently investigated due to its cancer signaling properties, along with *CASS4*, which was also significantly associated with late onset Alzheimer's disease in our meta-analysis (*Beck TN et al. 2014*).
- We also reported a finding in the locus of *SLC24A4,* a solute carrier, which is associated with eye and hair pigmentation (*Sulem P et al. 2007*), the development of amelogenesis imperfecta (*Parry DA et al. 2013*) and hypertension in African Americans (*Adeyemo A et al. 2009*).
- Importantly the obtained findings may point towards previously unidentified pathways involved in the etiology of Alzheimer's disease. These could revolve around hippocampus (*PTK2B and MEF2C*), cytoskeletal (*CASS4, CELF1*, and *NME8*), and microglial and myeloid cell function (*INPP5D*).

7.4 Future research

7.4.1 New phenotypes and biomarkers

Brain aging progresses with ongoing age. It is therefore important to detect diseases at the earliest time point, since prevalent dementia will likely remain incurable.

- Future phenotypes in brain aging will probably be subclinical, in order to reverse or ameliorate abnormal aging before clinical onset, earlier than MCI. One possibility to

achieve this is by longitudinal observation of healthy elderly patients over long time periods to identify potential high risk targets and key determinants.

- Advent of new MRI technology might also allow detection of more subtle age-related changes. Longitudinal studies have shown that white matter lesions are preceded by changes in normal appearing white matter, which are quantifiable (*de Groot M et al. 2013*). Diffusion tractography allows localization of tracts in white matter and can consequently predict progression of white matter lesions (*Maillard et. al. 2015*). The method of tractography imaging is still its infancy without widespread clinical application (*Dell'Aqua F & Catani M 2012*) and little correlation with other imaging methods such as Fluid attenuated inversion recovery (FLAIR) (*Zhan W et al. 2008)*.
- Relevant topics in future brain aging research will also likely include circulating markers, some of which have already been successfully identified. Blood inflammatory markers are associated with risk of vascular dementia (*Ravaglia G et al. 2007*). The brain derived neurotrophic (BDNF) factor and the vascular endothelial growth factor (VEGF) in serum are associated with stroke risk (*Pikula A et al. 2013*). In addition low blood red blood cell count and ω-3 fatty acid concentrations have been associated with accelerated brain aging and cognitive impairment (*Tan ZS et al. 2012*).
  Other biomarkers such as microRNAs might be predictive of brain aging as well, given their importance in neuron homeostasis (*Persengiev SP et al. 2012*). This is further highlighted as expression patterns of microRNAs display age-related differences in mouse brains (*Inukai S et. al. 2012*).

7.4.2 New techniques and approaches in genetic epidemiology

- Whole genome sequencing is becoming a more feasible option, since prices for an individual genome recently hit the 1000$ threshold in 2014. The major drawback is the creation of huge amounts of unwanted data (*Baker M 2010*), which represents a challenge for both analysis and interpretation (*Xuan J et al. 2013*).
  Given the necessity for high priced machinery for the conduct of NGS studies, there are several companies that provide whole genome sequencing services to customers.

The problem of data storage can be alleviated using cloud based solutions in which the generated data can be stored as well as analyzed online by providers such as Globus Genomics (*Madduri RK et al. 2014*) available at https://www.globus.org/genomics/ .

- Based on the success of previous GWAS on common variants, more effort will be undertaken to investigate rare variants with MAF ≤5% (*Lee S et al. 2014*). This requires investigations with large sample sizes, which can typically only be achieved through collaborative effort in consortia. A major benefit from these collaborations is that an analysis plan can be established for prospective studies, which in turn helps to minimize inter study heterogeneity. The results from the conducted GWAS can then be used for further research in a meta-analysis of study. In this type of study results from the primary investigation can be combined to achieve large enough statistical power to detect small effects as well as associations with rare polymorphisms (*Munafò MR & Flint J 2004*).
- A promising sector of upcoming genetic epidemiology is research in epigenetics. Research in the field is still in its early stages, as traditional approaches and study design in genetic epidemiology might not be applicable and need rethinking (*Mill J & Heijmans BT 2013*). Epigenome wide studies (EWA) s are currently underway and predominantly investigate DNA methylation patterns of common diseases (*Rakyan VK et al 2011*). The investigated phenotypes so far include the influence of high BMI (*Dick KJ et al 2014*) and multiple sclerosis on (*Huynh JL et al. 2014*).
- Findings from GWAS typically associate a phenotype and a genotype in the form of a SNP, which is both valuable and valid. The benefit, however is not necessarily apparent from a molecular or biochemical point of view. Therefore pathway analysis is performed subsequently to identify the biological pathways behind the genes housing polymorphisms (*Wang K et. al.2010*). On has to consider that SNPs can also fall into pathways by chance or through linkage disequilibrium (*Peng G et al. 2010*). However, many of the observed genotype-phenotype associations cannot be significantly associated with any known pathway (*Brodie A et al. 2014*). This is likely caused by our limited knowledge of existing pathways. Future genetic association studies may well help unravel new pathways, but not necessarily so, as the molecular underpinnings of

any disease are limited.

- Another example of likely future genetic research will be better prediction of diseases. More and more risk factors for common diseases are being discovered, which allow the estimation of risks, f.e. in the case of stroke (*Ibrahim-Verbaas CA et al. 2014*). Consecutive price drops in genotyping will easily allow detection of these risk factors or genotyping the broad majority. Knowledge of genetic risk factors will allow early interventions before the potential outbreak of the disease.

7.4.3 **Application of findings**

Findings from genetic association studies do not easily transition into clinical therapies. This is especially evident in the case of *APOE* ε4. The risk allele is associated with late onset Alzheimer disease (LOAD) and increases the risk of onset dramatically, depending on the amount of alleles. Homozygotes ε4 have a relative risk of 14.9 to develop LOAD (*Farrer LA et al. 1997*). Population attributable risk (PAR), which is the reduction in incidence that would hypothetically be observed without an exposure or risk factor is 50% in case of Alzheimer and *APOE* genotype (*Ashford JW et al. 2004*).

This discovery has been made well over two decades ago (*Corder EH et al. 1993*) and is indisputably accepted across the scientific community. Nevertheless no medical compound has been launched on the market for therapy of *APOE* ε4 carriers. This is especially bewildering as the *APOE* ε4 allele is very common and one in five individuals carrying the risk allele. Furthermore prevalence of dementia is estimated at ~36 million in 2010 (*Prince M et al. 2013*), which highlights the lucrative prospect for pharmaceutical companies.

It is therefore preferable to recommend a healthier life-style including regular exercise than to wait for development of therapeutic agents. Fitness obviously plays an important role to play in non-pharmacological therapeutic strategies in brain aging (*Philips C et al. 2014*). In contrast to genetic research benefits, these benefits are much more easily available for the individual and may also carry beneficial side effects, f.e. prevention of depression and disability,

strengthening of muscles and bones.

There is also indication that physical activity yields differential outcomes depending on genetic risk score (GRS) through gene x physical activity interactions (*Kilpeläinen TO et al. 2008*; *Ahmad S et al. 2013*). Medicine is becoming more personalized in general and life-style recommendations could also be tailored towards individuals depending on their general risk score (GRS) profile.

Age-related neurological diseases cannot be reversed in the near future, and changes in life-style are also becoming more difficult to handle with increasing age.
This raises the question at what age intervention should be considered in order to establish guidelines. According to recent observations there is no indication of neuroprotective effects of CRF in midlife and children with better cognition simply choose to lead healthier life's (*Belsky DW et al. 2015*). Several reports claim that moderate intensity physical exercise at later adulthoods is recommended for the prevention brain aging, especially for people who lead a sedentary life-style (*Middleton LE & Yaffe K 2010*).
The intensity, duration and type of exercise that is suitable for the individual to conserve cognitive function in the individual is uncertain and will remain a subject in studies to come.

8. Summary

This thesis provides novel insight into the complex topic of brain aging and brain aging associated diseases. Both environmental influences and genetic architecture is investigates. We collect novel evidence for a beneficial effect of maintaining cardiorespiratory fitness, which is protective of functional and structural brain aging. In addition we summarize genetic literature on vascular dementia and white matter lesions. Several state of the art approaches in genetic research are applied in the presented papers. We highlight the role of common polymorphisms of the NOTCH3 gene in the development of age-related white matter lesions in a candidate gene approach and direct sequencing. In a separate article we estimate heritability of intracerebral hemorrhage in unrelated individuals using a novel method.
In two separate meta-analyses of data we find novel genetic susceptibility loci for intracerebral hemorrhage and late onset Alzheimer's disease.
Additional detail is paid on how findings of the investigations can be applied and how genetic research transition does not necessarily transition to treatments. Much more importantly life-style recommendations, including an active life-style and high cardiorespiratory fitness are important in alleviating symptoms of brain aging. The findings of cardiorespiratory fitness on brain aging are robust, and the effect sizes are considerable. Participants perform a minimum of 5 years better in all cognitive domains than their peers.
Importantly, all the work presented in this paper represents a starting point for future investigations. Given the global aging of the populous and the rise in age-related diseases, brain aging is likely to stay a public and scientific interest.

9. Appendix

9.1 Course participation

- „Linux for scientists“ Erasmus MC, Rotterdam
- „Family based genetic studies“ Erasmus MC, Rotterdam

9.2 References

Adeyamo A, Gerry N, Chen G, et al. A genome-wide association study of hypertension and blood pressure in African Americans. *PLoS Genet* 2009;5(7):e1000564.

Adzhubei IA, Schmidt S, Peshking L et al. A method and server for predicting damaging missense mutations. *Nat Methods* 2010;7(4):248-9.

Ahmad S, Rukh G, Varga TV, et al. Gene x physical activity interactions in obesity: combined analysis of 111,421 individuals of European ancestry. *PloS Gent* 2013;9(7):e1003607.

Alberts MJ, McCarron MO, Hoffmann KL, Graffagnino C. Familial clustering of intracerebral hemorrhage: a prospective study in North Carolina. *Neuroepidemiology* 2002;21(1):18-21.

American Psychiatric Association (2000) Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition, Text Revision. Washington DC: American Psychiatric Association. 944 p.

Ashford JW. APOE genotype effects on Alzheimer's disease onset and epidemiology. *J Mol Neurosci* 2004;23(3):157-65.

Atwood LD, Wolf PA, Heard-Costa NL, et al. Genetic variation in white matter hyperintensity volume in the Framingham Study. *Stroke* 2004;35(7):1609-13.

Baker M. Next-generation sequencing: adjusting to data overload. *Nat Methods* 2010;7:495-99.

Barnes DE, Yaffe K, Satariano WA, Tager IB. A longitudinal study of cardiorespiratory fitness and cognitive function in healthy older adults. *J Am Geriatr Soc* 2003;51(4):459-65.

Barsh GS, Copenhaver GP, Gibson G, Williams SM, et al. Guidelines for genome-wide association studies. *PLoS Genet* 2012;8(7):e1002812.

Barzilai N, Huffman DM, Muzumdar RH, et al. The critical role of metabolic pathways in aging. *Diabetes* 2012;61(6):1315-22.

Batouli SA, Sachdev PS, Wen W, et a. Heritability of brain volumes in older adults: the Older Australian Twins Study. *Neurobiol Aging* 2014;35(4):937.e5-18.

Beck TN, Nicolas E, Kopp MC, Golemis EA. Adaptors for disorders of the brain? The cancer signaling proteins NEDD9, CASS4, and PTK2B in Alzheimer's disease. *Oncoscience* 2014;1(7):486-503.

Belsky DW, Caspi A, Israel S, et al. Cardiorespiratory fitness and cognitive function in midlife: Neuroprotection or Neuroselection? *Ann Neurol* 2015;77(4):607-17.

Benjamin P, Lawrence AJ, Lambert C. Strategic lacunes and their relationship to cognitive impairment in cerebral small vessel disease. *Neuroimage Clin* 2014;4:828-37.

Bertram L, Tanzi RE**.** Thirty years of Alzheimer's diseas genetics: the implications of systematic meta-analyses. *Nat Rev Neurosci* 2008;9(10):768-78.

Biffi A, Battey TW, Ayres AM, Warfarin-related intraventricular hemorrhage: imaging and outcome. *Neurology* 2011;77(20):1840-6.

Blackburn EH, Greider CW, Szostak JW. Telomeres and telomerase: the path from maize, Tetrahymena and yeast to human cancer and aging. *Nat Med* 2006;12(10):1133-8.

Botstein D, Risch N. Discovering genotypes underlying human phenotypes: past successes for mendelian disease, future approaches for complex disease. *Nat Genet* 2003;33 Suppl:228-37.

Bouaziz M, Ambroise C, Guedj M. Accounting for population stratification in practice: a comparison of the main strategies dedicated to genome-wide association studies. *PLoS One* 2011;6(12):e28845.

Bowles HR, FitzGerald SJ, Morrow JR Jr, et al. Construct validity of self-reported historical physical activity. *Am J Epidemiol* 2004;160(3):279-86.

Brockmole JR, Logie RH. Age-related change in visual working memory: a study of 55753 participants aged 8-75. *Front Psychol* 2013;4:12.

Brodie A, Tovia-Brodie O, Ofran Y. Large scale analysis of phenotype-pathway relationships based on GWAS results. *PloS One* 2014;9(7):e100887.

Campbell H, Rudan I. Interpretation of genetic association studies in complex disease. *Pharmacogenomics J* 2002;2(6):349-60.

Candore G, Balistreri CR, Colonna-Romano G, et al. Major histocompatibility complex and sporadic Alzheimer's disease: a critical reappraisal. *Exp Gerontol* 2004;39(4):645-52.

Cardon LR, Palmer LJ. Population stratification and spurious allelic association. *Lancet* 2003;361(9357):598-604.

Carmelli D, DeCarli C, Swan GE *et al.* Evidence for genetic variance in white matter hyperintensity volume in normal elderly male twins. *Stroke* 1998;29(6), 1177–1181**.**

Chabriat H, Joutel A, Dichgans M, et al. Cadasil. Lancet Neurol. 2009 Jul; 8(7):643-53.

Conrad DF, Pinto D, Redon R. Origins and functional impact of copy number variation in the human genome. *Nature* 2010;464(7289):704-12.

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):921-3.

Cordell HJ, Clayton DG. Genetic association studies. *Lancet* 2005 Sep 24-30;366(9491):1121-31.

Chuang YF, Hayden KM, Norton MC, et al. Association between APOE ε4 Allele and Vascular Dementia: The Cache County Study. *Dement Geriatr Cogn Disord* 2010;29(3): 248–253.

Colcombe S, Kramer AF. Fitness effects on the cognitive function of older adults: a meta-analytic study. *Psychol Sci* 2003;14(2):125-30.

Colhoun HM, McKeigue PM, Davy Smith G. Problems of reporting genetic associations with complex outcomes. *Lancet* 2003;361(9360):865-72.

Cullen B, O'Neill B, Evans JJ et al. A review of screening tests for cognitive impairment. *J Neurol Neurosurg Psychiatry* 2007;78(8):790-9.

Dawe B, Procter A, Philpot M. Concepts of mild memory impairment in the elderly and their relationship to dementia: a review. *Int J Geriatr Psychiatry* 1992;7473- 479.

Davie CA. A review of Parkinson's disease. *Br Med Bull* 2008;86:109-27.

Deary IJ, Penke L, Johnson W. The neuroscience of human intelligence differences. *Nat Rev Neurosci* 2010;11(3):201-11.

DeGroot M, Verhaaren BF, de Boer R, et al. Changes in normal-appearing white matter

precede development of white matetr lesions. *Stroke* 2013;44(4):1037-42.

Delgado-Rodriguez M, Llorca J. Bias. *J Epidemiol Community Health* 2004;58(8):635-41.

Dell'Aqua F, Catani M. Structural human brain networks: hot topics in diffusion tractography. *Curr Opin Neurol* 2012;25(4):375-83.

Del Ser T, Hachinski V, Merskey H, et al. An autopsy-verfied study of the effect of education on degenerative dementia. *Brain* 1999;122 (Pt 12):2309-19.

Dick KJ, Nelson CP, Tsaprouni L. DNA methylation and body-mass index: a genome-wide analysis. *Lancet* 2014;383(9933):1990-8.

Eichler EE, Flint J, Gibson G, et al. Missing heritability and strategies for finding the underlying causes of complex disease. Nat Rev Genet 2010;11(6):446-50.

Elias MF, Elias PK, Sullivan LM, et al. Obesity, diabetes and cognitive deficit: The Framingham Heart Study. *Neurobiol Aging* 2005;26 Suppl 1:11-6.

Falcone GJ; Biffi A, Brouwers HB, et al. Predictors of hematoma volume in deep and lobar supratentorial intracerebral hemmhorage. *JAMA Neurol* 2013;70(8):988-94.

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. APOE and Alzheimer Disease Meta Analysis Consortium. *JAMA* 1997;278:1349-1356.

Feigin VL, Lawes CMM, Bennett DA, et al. Worldwide stroke incidence and early case fatality reported in 56 population-based studies: a systematic review. *Lancet Neurol* 2009;8(4):355-69.

Feldstein CA. Association between chronic blood pressure changes and development of Alzheimer's disease. *J Alzheimers Dis* 2012;32(3):753-63.

Fewell Z, Davey Smith G, Sterne JA. The impact of residual and unmeasured confounding in epidemiologic studies: a simulation study. *Am J Epidemiol* 2007;166(6):646-55.

Fiebach JB, Schellinger PD, Gass A, et al. Stroke magnetic resonance imaging is accurate in hyperacute intracerebral hemorrhage: a multicenter study on the validity of stroke imaging. *Stroke* 2004;35(2):502-6.

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;12189- 198.

Fornage M, Debette S, Bis J, et al. Genome-wide association studies of cerebral white matter

lesion burden: the CHARGE consortium. *Ann Neurol* 2011;69(6):928-39.

Frazer KA, Murray SS, Shork NJ, Topol EJ. Human genetic variation and its contribution to complex traits. *Nat Rev Genet* 2009;10(4):241-51.

Freeman JL Perry GH, Feuk L, et al. Copy number variation: new insights in genome diversity. *Genome Res* 2006;16(8):949-61.

Freedman ML, Reich D, Penney KL, et al. Assessing the impact of population stratification on genetic association studies. *Nat Genet* 2004;36(4):388-93.

Freedson P, Bowles HR, Troiano R, Haskwell W. Assessment of physical activity using wearable monitors: recommendations for monitor calibration and use in the field. *Med Sci Sports Exerc* 2012;44(1 Suppl 1):S1-4

Geschwind DH, Miller BL, DeCarli C, et al. Heritability of lobar brain volumes in twins supports genetic models of cerebral laterality and handedness. *Proc Natl Acad Sci USA* 2002;99(5):3176-81.

Green DR, Galluzi L, Kroemer G. Mitochondria and the autophagy-inflammation-cell death axis in organismal aging. *Science* 2011;333(6046):1109-12.

Grundman M, Petersen RC, Ferris SH, et al. Mild cognitive impairment can be distinguished from Alzheimer disease and normal aging for clinical trials. *Arch Neurol* 2004;61(1):59-66.

Guralnik JM, Kritchevsky SB. Translating research to promote healthy aging: the complementary role of longitudinal studies and clinical trials. *J Am Geriatr Soc* 2010;58 Suppl 2:S337-42.

Haines JL, Hauser MA, Schmidt S, et al. Complement Factor H Variant Increases the Risk of Age-Related Macular Degeneration. *Science* 2005;308(5720):419-21.

Hayes SM, Forman DE, Verfaellie M. Cardiorespiratory Fitness Is Associated With Cognitive Performance in Older But Not Younger Adults. *J Gerontol B Psychol Sci Soc Sci* 2014 pii: gbu167.

He K, Song Y, Daviglus ML, et al. Fish consumption and incidence of stroke: a meta-analysi of cohrort studies. *Stroke* 2004;35(7):1538-42.

Hoeijmakers JH. DNA damage, aging, and cancer. *N Engl J Med* 2009;361(15):1475-85.

Hsin H, Kim MJ, Wang CF, Sheng M. Proline-rich tyrosine kinase 2 regulates hippocampal long-term depression. *J Neurosci* 2010;30(36):11983-93.

Huynh JL, Garg P, Thin TH, et al. Epigenome-wide differences in pathology-free regions of multiple sclerosis-affected brains. *Nat Neurosci* 2014 Jan;17(1):121-30.

Ibrahim-Verbaas CA, Fornage M, Bis JC, et al. Predicting stroke through genetic risk functions: the CHARGE Risk Score Project. *Stroke* 2014;45(2):403-12.

Igarashi K, Kashiwagi K. Use of polyamine metabolites as markers for stroke and renal failure. *Methods Mol Biol* 2011;720:395-408.

Ikram MA, Wieberdink RG, Koudstaal PJ. International epidemiology of intracerebral hemorrhage. *Curr Atheroscler Rep* 2012;14(4):300-6.

International Hapmap Consortium. A haplotype map of the human genome. *Nature* 2005;437(7063):1299-320.

Inukai S, de Lencastre A, Turner M, Slack F. Novel microRNAs differentially expressed during aging in the mouse brain. *PloS One* 2012;7(7):e40028.

Itsara A, Cooper GM, Baker C, et al. Population analysis of large copy number variants and hotspots of human genetic disease. *Am J Hum Genet* 2009;84(2):148-61.

Jackson AS, Sui X, Hébert JR et al. Role of life-style and aging on the longitudinal change in cardiorespiratory fitness. **Johnson RC**, Nelson GW, Troyer JL, et al. Accounting for multiple comparisons in a genome-wide association study (GWAS). *BMC Genomics* 2010;11:724.

Jones DL, Rando TA. Emerging models and paradigms for stem cell ageing. *Nat Cell Biol* 2011 May;13(5):506-12.

Jones RN, Gallo JJ. Education bias in the mini-mental state examination. *Int Psachogeriatr 2001*;13:299-310.

Joutel A, Corpechot C, Ducros A*,* et al. Notch3 mutations in CADASIL, a hereditary adult-onset condition causing stroke and dementia. *Nature* 1996;383(6602):707-10.

Kamboh MI, Demirci FY, Wang X, et al. Genome-wide association study of Alzheimer's disease. *Transl Psychiatry* 2012;2:e117.

Kanehisa M, Goto S. KEGG: Kyoto Encyclopedia of Genes and Genomes. *Nucleic Acids Res* 2000;28(1):27-30.

Kaprio J. Twins and the mystery of missing heritability: the contribution of gene-environment interactions. *J Intern Med* 2012;272(5):440-8.

Karran E, Mercken M, De Strooper B. The amyloid cascade hypothesis for Alzheimer's

disease: an appraisal for the development of therapeutics. *Nat Rev Drug Discov* 2011;10(9):698-712. doi: 10.1038/nrd3505.

Kilpeläinen TO, Lakka TA, Laaksonen DE, et al. SNPs in PPARG associate with type 2 diabeters and interact with physical activity. *Med Sci Sports Exerc* 2008;40(1):25-33.

Kim KW, MacFall JR, Payne ME. Classification of white matter lesions on magnetic resonance imaging in elderly persons. *Biol Psychiatry* 2008;64(4):273-80.

Kinugawa K, Schumm S, Pollina M, et al. Aging-related episodic memory decline: are emotions the key? *Front Behav Nurosci* 2013;7:2**.**

Kokmen E, Beard CM, O'Biren PC, et al. Is the incidence of dementing illness changing? A 25-year time trend study in Rochester, Minnesota (1960-1984). *Neurology* 1993;43(10):1887-92.

Kremen WS, Panizzon MS, Neale MC, et al. Heritability of brain ventricle volume: converging evidence from inconsistent results. *Neurobiol Aging* 2012;33(1):1-8.

Launer LJ, Andersen K, Deey ME, et al. Rates and risk factors for dementia and Alzheimer's disease: results from EURODEM pooled analyses. EURODEM Incidence Research Group and Work Groups. European Studies of Dementia. *Neurology* 1999;52(1):78-84.

Laurin D, Verreault R, Lindsay, et al. Physical activity and risk of cognitive impairment and dementia in elderly persons. *Arch Neurol* 2001;58, 498–504.

Lautenschlager NT, Cox KL, Flicker L, et al. Effect of physical activity on cognitive function in older adults at risk for Alzheimer disease: a randomized trial. *JAMA* 2008;300(9):1027–37.

Lee S, Abecasis GR, Boehnke M, Lin X. Rare-variant association analysis: study designs and statistical tests. *Am J Hum Genet* 2014;95(1):5-23.

Levine B, Svoboda E, Hay JF, et al. Aging and autobiographical memory: dissociating episodic from semantic retrieval. *Psychol Aging* 2002;17:677.

Levy S, Sutton G, Ng PC, et al. The diploid genome sequence of an individual human. PLoS Biol. 2007;5(10):e254.

Lobo A, Launer LJ, Fratiglioni L, et al. Prevalence of dementia and major subtypes in Europe: A collaborative study of population-based cohorts. Neurologic Diseasses in the Elderly Research Group. *Neurology* 2000;54(11 Suppl 5):S4-9.

López-Otin C, Blasco MA, Partridge L, et al. The hallmarks of aging *Cell* 2013;153(6):1194-217.

Lusis AJ. Atherosclerosis. *Nature* 2000;407(6801):233-41.

Madduri RK, Sulakhe D, Lacinski L, et al. Experiences Building Globus Genomics: A Next-Generation Sequencing Analysis Service using Galaxy, Globus and Amazon Web Services. *Concurr Comput* 2014;26(13):2266-2279.

Maillard P, Seshadri S, Beiser A, et al. Abstract 221: Regional Vulnerability within White Matter tracts to Whiter Matter Hyperintensities: a Diffusion Tensor Imaging Tractography Study. *Stroke* 2015;46:A221.

Manolio TA, Collins FS, Cox NJ, Finding the missing heritability of complex diseases. *Nature* 2009;461(7265):747-53.

Marchini J, Cardon LR, Philips MS, Donelly P. The effects of human population structure on large genetic association studies. *Nat Genet* 2004;36(5):512-7.

Matsukawa H, Shinoda M, Fujii M, et al. Factors associated with lobar vs. non-lobar intracerebral hemorrhage. *Acta Neurol Scand* 2012;126(2):116-21.

Mattson MP. Will caloric restriction and folate protect against AD and PD? *Neurology* 2003 Feb 25;60(4):690-5.

Mayeux R, Stern Y. Epidemiology of Alzheimer disease. *Cold Spring Harb Perspect Med* 2012;2(8).

Middleton LE, Yaffe K. Targets for the prevention of dementia. *J Alzheimers Dis* 2010;20(3):915-24.

Mill J, Heijmans BT. From promises to practical strategies in epigenetic epidemiology. *Nat Rev Genet* 2013 Aug;14(8):585-94. **Moskalev AA**, Shaposhnikov MV, Pluyusnina EN, et al. The role of DNA damage and repair in aging through the prism of Koch-like criteria. *Aging Res Rev* 2013;12(2):661-84.

Munafò MR, Flint J. Meta-analysis of genetic association studies. *Trends Genet* 2004(9):439-44.

Munoz DG. Small vessel diseas: neuropathology. *Int Psychogeriatr* 2003;15 Suppl 1:67-9.

NCI-NHGRI Working Group on Replication in Association Studies, Chandock SJ, Manolio T, et al. Replicating genotype-phenotype associations. *Nature* 2007;447(7145):655-

60.

Naqvi RM, Haider S, Tomlinson G, Alibhai S. Cognitive assessments in multicultural populations using the Rowland Universal Dementia Assessment Scale: a systematic review and meta-analysis. *CMAJ* 2015 Mar 17;187(5):E169-75.

Ng SB, Turner EH, Robertson PD, et al. Targeted capture and massively parallel sequencing of human exomes. *Nature* 2009;461(7261):272-6.

Ng PC, Henikoff S. Predicting deleterious amino acid substituations. *Genome Res* 2001 ;11(5):863-74.

Olsson E, Klasson, Berge K, et. al. White Matter Lesion Assessment in Patients with Cognitive Impairment and Healthy Controls: Reliability Comparisons between Visual Rating, a Manual and an Automatic Volumetrical MRI Method-The Gothenburg MCI Study. *J Aging Res* 2013;2013:198471.

Otsuka M, Yamaguchi K, Ueki A. Similarities and differences between Alzheimer's disease and vascular dementia from the viewpoint of nutrition. *Ann N Y Acad Sci* 2002;977:155-61.

Ott A, Stolk RP, van Harskamp, et al. Diabetes mellitus and the risk of dementia: The Rotterdam Study. *Neurology* 1999;53(9):1937-42.

Pantoni L, Garcia JH. Pathogenesis of leukoaraiosis: a review. *Stroke* 1997;28(3):652-9.

Parker BA, Kalasky MJ, Proctor DN. Evidence for sex differences in cardiovascular aging and adaptive responses to physical activity. *Eur J Appl Physiol* 2010;110(2):235-46.

Parry DA, Poulter JA, Logan CV, et al. Identification of mutations in SLC24A4, encoding a potassium-dependent sodium/calcium exchanger, as a cause of amelogenesis imperfecta. *Am J Hum Genet*. 2013;92(2):307-12.

Patnala R, Clements J, Batra J. Candidate gene association studies: a comprehensive guide to useful in silico tools. BMC Genet. 2013;14:39.

Pedraza O, Lucas JA, Smith GE, et al. Robust and expanded norms for the Dementia Rating Scale. *Arch Clin Neuropsychol* 2010;25(5):347-58.

Peloso GM, Auer PL, Bis JC, et al. Association of low frequency and rare coding sequence variants with blood lipids and coronary heart disease in 56,000 whites and blacks. *Am J Hum Genet* 2014;94(2):223-32

Peng G, Luo L, Siu H, et al. Gene and pathway-based second-wave analysis of genome-wide

association studies. *Eur J Hum Genet* 2010;18(1):111–117.

Persengiev SP. Kondova II, Bontrop RE. The Impact of MicroRNAs on Brain Aging and Neurodegeneration. *Curr Gerontol Geriatr Res* 2012;2012:359369.

Petersen RC, Smith GE, Waring SC, et al. Mild cognitive impairment: clinical characterization and outcome. *Arch Neurol* 1999 Mar;56(3):303-8.

Philips C, Baktir MA, Srivatsan M, Salehi A. Neurotrophic effects of physical activity on the brain: a closer look at trophic factor signaling. *Front Cell Neurosci* 2014;8:170.

Piguet O, Ridley LJ, Grayson DA, et al. Comparing white matter lesions on T2 and FLAIR MRI in the Sydney Older Persons Study. Eur J Neurol 2005;12(5):399-402.

Pikula A, Beiser AS, Chen TC, et al. Serum brain-derived neurotrophic factor and vascular endothelial grwoht factor levels are associated with risk of stroke and vascular brain injury: Framingham Study. *Stroke* 2013;44(10):2768-75.

Plassman BL, Langa KM, Fisher GG, Heeringa SG, Weir DR, Ofstedal MB, et al. Prevalence of dementia in the United States: the aging, demographics, and memory study. *Neuroepidemiology* 2007;29:125–32.

Potter GM, Doubal F, Jackson CA, et al. Long-term appearance of lacunar infarction on imaging: what proportion become CSF-containing cavities? *Cerebrovasc Dis* 2008; 25 (Suppl 2): 30.

Pottier C, Hannequin D, Coutant S, et al. High frequency of potentially pathogenic SORL1 mutations in autosomal dominant early-onset Alzheimer disease. *Mol. Psychiatry* 2012;17:875–879.

Powers ET, Morimoto RI, Dillin A, et al. Biological and chemical approaches to diseases of proteostasis deficiency. *Annu Rev Biochem* 2009;78:959-91.

Preacher KJ, Hayes AF. Asymptotic and resampling strategies for assessing and comparing indirect effects in multiple mediator models. *Behav Res Methods* 2008;40(3):879-91.

Price AL, Patterson NJ, Plenge RM, et al. Principal components analysis corrects for stratification in genome-wide association studies. *Nat Genet* 2006 Aug;38(8):904-9.

Prince M, Bryce R, Albanese E, Wimo A, Ribeiro W, Ferri CP. The global prevalence of dementia: a systematic review and metaanalysis. *Alzheimers Dement* 2013;9(1):63-75.

Prince SA, Adamo KB, Hamel ME, et al. A comparison of direct versus self-report measures

for assessing physical activity in adults: a systematic review. *Int J Behav Nutr Phys Act* 2008;5:56.

Prins ND, van Straaten EC, van Dijk Ej et al. Measuring progression of cerebral white matter lesions on MRI: visual rating and volumetrics. *Neurology* 2004. 62,1533–1539.

Pritchard JK, Cox NJ. The allelic architecture of human disease genes: common disease-common variant...or not? *Hum Mol Genet* 2002;11(20):2417-23.

Pruim RJ, Welch RP, Sanna S, et al. LocusZoom: regional visualization of genome-wide association scan results. *Bioinformatics* 2010;26(18):2336-7.

Purcell S, Neale B, Todd-Brown K, et al. PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am J Hum Genet* 2007;81(3):559-75.

Quigley C, Müller MM. Feature selective attention in healhty old age: a selective decline in selective attention? *J Neurosci* 2014;34(7):2471-6.

Rakyan VK, Down TA, Balding DJ, Beck S. Epigenome-wide association studies for common human diseases. *Nat Rev Genet* 2011;12(8):529-41.

Ravaglia G, Forti P, Maioli F, et al. Blood inflammatory markers and risk of dementia: The Conselice Study of Brain Aging. *Neurobiol Aging* 2007;28(12):1810-20.

Raz N, Lindenberger U, Rodrigue KM, et al. Regional brain changes in aging healthy adults: general trends, individual differences and modifiers. *Cereb Cortex* 2005 Nov;15(11):1676-89.

Razvi SS, Davidson R, Bone I, et al. The prevalence of cerebral autosomal dominant arteriopathy with subcortical infarcts and leucoencephalopathy (CADASIL) in the west of Scotland. *J Neurol Neurosurg Psychiatry* 2005;76(5), 739–741.

Reich DE, Cargill M, Bolk S. Linkage disequilibrium in the human genome. *Nature* 2001;411(6834):199-204.

Reilly D. Gender, culture, and sex-typed cognitive abilities. *PLoS One* 2012;7(7):e39904.

Rogaeva E, MengY, Lee JH, et al. The neuronal sortilin-related receptor SORL1 is genetically associated with Alzheimer disease. *Nat Genet* 2007;39:168–177.

Rolland Y, Abellan van Kan G, Vellas B. Physical activity and Alzheimer's disease: from prevention to therapeutic perspectives. *J Am Med Dir Assoc* 2008;9(6):390–405.

Rost NS, Fitzpatrick K, Biffi A, et al. White matter hyperintensity burden and susceptibility to cerebral ischemia. *Stroke* 2010;41(12):2807-11.

Rothman KJ, Greenland S, Lash TL. Modern Epidemiology. Third ed. Philadelphia, USA:Lipincott Williams and Wilkins; 2008.

Ruitenberg A, van Swieten JC, Witteman JC, et al. Alcohol consumption and risk of dementia: the Rotterdam Study. *Lancet* 2002;359(9303):281-6.

Rutten-Jacobs LC, Traylor M, Adib-Samii P, et al. Common NOTCH3 Variants and Cerebral Small- Vessel Disease. *Stroke* 2015;46(6):1482-7.

Salat DH, Buckner RL, Snyder AZ, et al. Thinning of the cerebral cortex in aging. *Cereb Cortex 2004*, 14:721-730.

Sahni R, Weinberger J. Management of intracerebral hemorrhage. *Vasc Health Risk Manag* 2007;3(5):701-9.

Scahill RI, Frost C, Jenkins R. et al. A longitudinal study of brain volume changes in normal ageing using serial registered magnetic resonance imaging. *Arch Neurol* 2003;60(7):989-94.

Schmidt R, Lechner H; Fazekas, F, et al. Assessment of cerebrovascular risk profiles in healthy persons: definition of research goals and the Austrian Stroke Prevention Study (ASPS). *Neuroepidemiology* 1994;13(6):308-13.

Schmidt R, Reinhart B, Schumacher M, et al. Prevalence of risk factors in the population of Graz (Austrian Stroke Prevention Study). *Wien Med Wochenschr* 1997;147(2):36-40.

Schmidt R, Schmidt H, Fazekas F, et al Apolipoprotein E polymorphism and silent microangiopathy-related cerebral damage. Results of the Austrian Stroke Prevention Study. *Stroke* 1997;28(5):951-6.

Schmidt R, Fazekas F, Kapeller P, et al. MRI white matter hyperintensities: three year follow-up of the Austrian Stroke Prevention Study. *Neurology* 1999;53(1):132-9.

Schmidt R, Schmidt H, Fazekas F, et al. Angiotensinogen polymorphism M235T, carotid atherosclerosis, and small-vessel disease-related cerebral abnormalities. *Hypertension* 2001;38(1):110-5.

Schmidt R, Enzinger C, Ropele S, et al Progression of cerebral white matter lestions: 6-year results of the Austrian Stroke Prevention Study. *Lancet* 2003.361, 2046–2048.

Schmidt R, Ropele S, Enzinger C, et al. White matter lesion progression, brain atrophy, and cognitive decline: the Austrian stroke prevention study. *Ann Neurol* 2005;58(4);610-6.

Seiler S, Schmidt H, Lechner A, et al. Driving cessation and dementia: results of the

prospective registry on dementia in Austria (PRODEM). *PLoS One* 2012;7(12):e52710.

Seshadri S, Wolf PA, Beiser A, et al. Lifetime risk of dementia and Alzheimer's disease. The impact of mortality on risk estimates in the Framingham Study. *Neurology* 1997;49(6):1498-504.

Seshadri S, Beiser A, Kelly-Hayes M, et al. The lifetime risk of stroke: estimates from the Framingham Study. *Stroke* 2006;37(2):345-50.

Sheet P, Stephens M. A fast and flexible statistical model for large-scale population genotype data: applications to inferring missing genotypes and haplotypic phase. *Am J Hum Genet* 2006;78(4):629-44.

Sherry ST, Ward MH, Kholodov M, et al. dbSNP: the NCBI database of genetic variation. *Nucleic Acids Res* 2001;29(1):308-11.

Shork NJ, Murray SS, Frazer KA, Topol EJ. Common vs. Rare allele hypotheses for complex diseases. *Curr Opin Genet Dev* 2009;19(3):212-9.

Skoog I, Lernfelt B, Landahl S, et al. 15-year longitudinal study of blood pressure and dementia. *Lancet* 1996;347(9009):1141-5.

Slatkin M. Epigenetic inheritance and the missing heritability problem. *Genetics* 2009;182(3):845-50.

Small SA, Duff K. Linking Abeta and tau in late-onset Alzheimer's disease: a dual pathway hypothesis. Neuron 2008 Nov 26;60(4):534-42.

Stankiewicz P, Lupski JR. Structural variation in the human genome and its role in disease. *Annu Rev Med* 2010;61:437-55.

Staff RT, Murray AD; Deary IJ, Whalley LJ. What provides cerebral reserve? *Brain* 2004;127(Pt 5):1191-9.

Stefanis L. α-Synuclein in Parkinson's disease. *Cold Spring Harb Perspect Med* 2012;2(2):a009399.

Stern Y, Gurland B, Tatemichi TK, et al. Influence of education and occupation on the incidence of Alzheimer's disease. *JAMA* 1994;271(13):1004–10.

Strong K, Mathers C, Bonita R. Preventing stroke: saving lives around the world. *Lancet Neurol* 2007 6, 182–187.

Sulem P, Gudbjartsson DF, Stacey SN, et al. Genetic determinants of hair, eye and skin

pigmentation in Europeans. *Nat Genet*. 2007;39(12):1443-52.

Sutherland GR, Auer RN. Primary intracerebral hemorrhage. *J Clin Neurosci* 2006;13(5):511-7.

Tabor HK, Risch NJ, Myers RM. Candidate-gene approaches for studying complex genetic traits: practical considerations. *Nat Rev Genet* 2002;3(5):391-7.

Tan ZS, Harris WS, Beiser AS, et al. Red blood cell ω-3 fatty acid levels and markers of accelerated brain aging. *Neurology* 2012;78(9):658-64.

Tager IB, Hollenberg M, Satariano WA. Association between self-reported leisure-time phyiscal activity and measures of cardiorespiratory fitness in an elderly population. *Am J Epidemiol* 1998;147(10):921-31.

Talens RP, Christensen K, Putter H, et al. Epigenetic variation during the adult lifespan: cross-sectional and longitudinal data on monozygotic twin pairs. *Aging Cell* 2012;11(4):694-703.

Teng EL, Hasegawa K, Homma K, et al. The cognitive abilities Screening Instrument (CASI): a practical test for cross cultural epidemiological studies of dementia. *Int Psychogeriatr* 1994;6(1):45-58;

Thomas D. Methods for investigating gene-environment interactions in candidate pathway and genome-wide association studies. *Ann Rev Public Health* 2010;31:21-36.

Tseng BY, Gundapuneedi T, Khan MA. et al. White matter integrity in physically fit older adults. *Neuroimage* 2013;82:510-6.

Tomitori H, Usui T, Seaki N, et al. Polyamin oxidase and acrolein as novel biochemical markers for diagnosis of cerebral stroke. *Stroke* 2005;36(12):2609-13

Tournier-Lasserve E., Joutel A., Melki J., et al. Cerebral autosomal dominant arteriopathy with subcortical infarcts and leukoencephalopathy maps to chromosome 19q12. *Nat Genet* 1993;3(3):256–259.

Tsang PS, Shaner TL. Age, attention, expertise, and time-sharing performance. *Psychol Aging* 1998;13:323.

Turner S, Armstrong LL, Bradford Y, et al. Quality control procedures for genome-wide association studies. *Curr Protoc Hum Genet* 2011;Chapter 1:Unit1.19.

Vadarajan BN, Zhang Y, Lee JH, et al. Coding mutations in SORL1 and Alzheimer disease.

Ann Neurol 2015;77(2):215-27.

Valdés Hernández Mdel C, Morris Z, Dickie DA, Close correlation between quantitative and qualitative assessments of white matter lesions. *Neuroepidemiology* 2013;40(1):13-22.

Verstynen TD, Lynch B, Miller DL, et al. Caudate Nucleus Volume Mediates the Link between Cardiorespiratory Fitness and Cognitive Flexibility in Older Adults. *J Aging Res* 2012;2012:939285.

Wallace DC. Mitochondrial DNA mutations in disease and aging. *Environ Mol Mutagen* 2010;51(5):440-50.

Wang K, Dickson SP, Stolle CA, et al. Interpretation of association signals and identification of causal variants from genome-wide association studies. *Am J Hum Genet* 2010;86(5):730-42.

Wang K, Li M, Hakonarson H. ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. *Nucleic Acids Res* 2010;38(16):e164.

Wang K, Li M, Hakonarson H. Analyzing biological pathways in genome-wide association studies. *Nat Rev Genet* 2010;11(12).843-54.

Wardlaw JM. What is a lacune? *Stroke* 2008;39(11):2921-2.

Wardlaw JM, Smith C, Dichgans M. Mechanisms of sporadic cerebral small vessel disease: insights from neuroimaging. *Lancet Neurol* 2013;12(5):483-97.

Weitz TM, Town T. Microglia in Alzheimer's disease: It's All About Context. *Int J Alzheimer's Dis* 2012;2012:314185.

Wenger HA; Bell GJ. The interactions of intensity, frequency and duration of exercise training in altering cardiorespiratory fitness. *Sports Med* 1986;3(5):346-56.

Wenk GL. Neuropathologic changes in Alzheimer's disease. *J Clin Psychiatry* 2003;64 Suppl 9:7-10.

Whitmer RA, Gustafson DR, Barrett-Connor E, et al. Central obesity and increased risk of dementia more than three decades later. *Neurology* 2008 Sep 30;71(14):1057-64.

Willer CJ, Li Y, Abecasis GR. METAL: fast and efficient meta-analysis of genome-wide association scans. *Bioinformatics* 2010;26(17):2190-1.

Winkler TW, Day FR, Croteau-Chonka DC, et al. Quality control and conduct of genome-wide association meta-analyses. *Nat Protoc* 2014;9(5):1192-212.

Wisløff U, Støylen A, Loennechen JP, et al. Superior cardiovascular effect of aerobic interval training versus moderate continuous training in heart failure patients: a randomized study. *Circulation* 2007;115(24):3086-94.

Wittke-Thompson JK, Pluzhnikov A, Cox NJ. Rational inferences about departures from Hardy-Weinberg equilibrium. *Am J Human Genet* 2005;76(6):967-86.

Woo D, Sauerbeck LR, Kissela BM, Khoury JC, Szaflarski JP, Gebel J, et al. Genetic and environmental risk factors for intracerebral hemorrhage: preliminary results of a population-based study. *Stroke* 2002;33:1190–1195.

Woodford HJ, George J. Cognitive assessment in the elderly: a review of clinical methods. *QJM* 2007;100(8):469-84.

Xuan J, Yu Y, Qing T, et al. Next-generation sequencing in the clinic: promises and challenges. *Cancer Lett* 2013;340(2):284-95.

Yang J, Lee SH, Goddard ME, Visscher PM. GCTA: a tool for genome-wide complex trait analysis. *Am J Hum Genet* 2011 Jan 7;88(1):76-82.

Yaffe K, Kanaya A, Lindquist K, et al. The metabolic syndrome, inflammation, and risk of cogntive decline. *JAMA* 2004;292(18):2237-42.

Zhan W. Zhang Y, Mueller SG, et al. Characterization of white matter degeneration in elderly subjects by magnetic resonance diffusion and FLAIR imaging correlation. *Neuroimage* 2009;47 Suppl 2:T58-65.

Zuk O, Hechter E, Sunyaev SR, Lander ES. The missing heritability: Genetic interactions create phantom heritability. *Proc Natl Acad Sci U S A* 2012;109(4):1193-8.

1000 Genomes Project Consortium. A map of human genome variation from population-scale sequencing. *Nature* 2010;467(7319):1061-73.