General introduction

The human brain is the most complex organ in the human body. The adult human brain comprises billions of neuronal and glial cells interconnected via trillions of synapses.[1,2] It is responsible for motor functions, processing sensory information, language, cognitive processes, and function of other organs. Pathology of the brain may occur prenatal, in early childhood or adolescence up to senescence. Disorders of the brain comprise a heterogeneous group of neurological and psychiatric disorders and are an important cause of disability and death worldwide.[3-5] These disorders are the result of a combination of genetic, environmental, and lifestyle factors. The focus of research presented in this thesis are most common neurological disorders from an epidemiological perspective. They include late-onset neurodegeneration and cerebrovascular pathology and the most common neurodevelopmental disorders including attention deficit hyperactivity disorder (ADHD) and autism spectrum disorder (ASD). I have also studied Duchenne muscular dystrophy, a recessive inherited disorder.

Expanding our knowledge on the molecular processes and pathways of these disorders and early pathology may facilitate development of new prevention and treatment strategies. The early changes manifested prior to the onset of clinical symptoms of the disease are usually approached as heritable quantitative measures and referred to as endophenotypes.[6-8] Endophenotypes can be measured accurately on a continuous scale, overcoming the problem of defining the arbitrary boundary between the presence and absence of subclinical disease in controls.[9] For long, cognitive ability has been studied as endophenotype of neurodegenerative and psychiatric disorders,[10-15] whereas more recently brain volumetric and vascular measures depicted by state-of-the-art imaging techniques have been studied as endophenotype of neurodegeneration and neurovascular pathology.[16-18]

LATE ONSET NEUROLOGICAL DISORDERS AND RELATED ENDOPHENOTYPES

The most common presentation of cerebrovascular pathology is stroke, a neurological disorder of sudden onset. Risk factors come in many varieties, including genetic factors and various modifiable risk factors. Beyond a large number of rare monogenic disorders underlying stroke,[9] 32 risk loci encompassing common and less-frequent variants have been associated with stroke in a study of 520,000 subjects.[19] These provide additional insights into stroke pathophysiology.[19] Several biological pathways including enlarged heart, decreased cardiac muscle contractility, and oxaloacetate metabolism emerged as relevant for any stroke, whereas various cardiac pathways, muscle-cell fate commitment, and nitric oxide metabolism are implicated in cardioembolic stroke.[19] A significant proportion of stroke risk also resides in modifiable risk factors including hypertension,

diabetes mellitus, cardiovascular disease, and smoking.[20,21] As management of these risk factors demonstrated reduction of stroke burden, additional research efforts to identify high-risk patients have been sought to improve the chances of success. Several studies performed to date searched for novel metabolic disturbances and identified various small circulating compounds to be associated with stroke.[22-27] The most comprehensive study to date is conducted in China Kadoorie Biobank, involving patients with both ischemic stroke (IS) (n = 1,146) and intracerebral hemorrhage (ICH) (n = 1,138).[26] The study reported association between lipoproteins and lipids with IS, but not with ICH. Additionally, the study reported association of glycoprotein acetyls and several non-lipid related metabolites with both IS and ICH.[26] To date, the studies in Europeans are based on relatively small samples.[25,27] A study involving 268 patients with incident stroke revealed no metabolites associated with stroke,[25] whereas another study reported association between lysophosphatidylcholine and stroke recurrence.[27] This asks for larger metabolomics studies of stroke in persons of European origin as presented in this thesis.

Cerebrovascular disease is also an important cause of dementia and cognitive decline.[28] A large number of genes have been implicated in dementia, predominately Alzheimer's disease (AD) but also frontotemporal dementia and Lewy body dementia.[29-32] The growing interest in early prevention of AD and cognitive decline, brought research of cognition in in the spotlight. Also there has been major progress in finding genes for cognitive function as endophenotype for various neurological and psychiatric disorders.[10-13,15] The major cognitive domains that have been studied in relation to these disorders include memory, language, executive function, and visuospatial ability.[10-13,15] Although the search for genes implicated in specific domains of cognition yielded some genes (e.g. *CADM2*, *HS3ST4*, *SPOCK3*),[33,34] the gene discovery improved its success when using general cognitive function, which captures all cognitive subdomains and shows a high correlation with intelligence and education.[35,36] General cognitive function is determined by environmental and genetic factors. Heritability estimates are reported to be more than 50% in adolescence and adulthood twin sample and 20-30% of variance is attributed to common variants.[35,37-39] Recent efforts identified more than 140 genomic regions encompassing common variants.[39] Furthermore, recent effort also reported evidence for a shared genetic origin with body mass index, waist to hip ratio, high-density lipoprotein levels, and cardiovascular diseases.[39] Even though these are drivers of the human metabolism, we have not linked yet genetic determinants of general cognitive function to circulating metabolites. Furthermore, most studies conducted to date included participants of European ancestry and a question to answer is whether the findings are generalizable to other ethnic groups. In this thesis I aim to find genetic determinants of general cognitive function, evaluate their generalizability to other ethnic groups and explore metabolic pathophysiology underlying established genetic variants. Despite all

efforts to date, common variants explain only a small proportion of cognitive test scores. Furthermore, diverse environmental factors have been implicated to influence cognitive function and the complex balance between genes and environment to cognitive function is poorly understood.[36,40]As studying epigenetic modifications may provide insights into molecular mechanisms underlying cognitive function, in this thesis we made an attempt to identify DNA methylation signatures of cognitive function.

At present, imaging is emerging as an endophenotype used in large-scale research of neurodegenerative disorders and stroke.[9] Finding genetic loci that influence this endophenotype may lead to identification of genes underlying related disorders. Studying brain structures using magnetic resonance imaging (MRI),[41,42] carotid intima-media thickness measured by carotid ultrasound and carotid artery calcification measured through computerized tomography (CT)[9,43] will expand our knowledge and provide novel insights into the pathophysiology of related disorders. In this thesis, I aim to explore genetic determinants of lateral ventricular volume, a measure of neurodegeneration, and intima-media thickness of carotid artery. Further, I aim to study metabolic determinants of carotid artery calcification, a measure of atherosclerosis.

NEURODEVELOPMENTAL DISORDERS

The most common neurodevelopmental disorders are ASD and ADHD.[44]

ASD is characterized by deficits in social communication and social interaction and restricted and repetitive patterns of activities and behavior.[45] The importance of genetic etiology is highlighted by heritability estimates ranging from 37% to 90%.[46-49] Progress in understanding genetic architecture of ASD has been made by identifying rare and de novo structural and sequence variation.[50,51] From a genetic perspective, ASD is an interesting disorder, as novel mutations have been implicated in patients that are not found back in either parent.[50] These variants have been identified in family-based studies.[52] Although most of the genetic risk for ASD is attributed to common variants, only a few genetic regions were successfully linked to ASD in family-based and population-based studies including unrelated patients and controls.[49,53-56] Despite a substantial increase in sample size, the most recent effort including over 16,000 individuals with ASD failed to identify common genetic variants associated with ASD asking for other approaches.[57] In this thesis, besides assessing the effect of single variants on ASD, I aim to evaluate the joint effect of multiple single genetic variants in a gene in a gene-based association analysis in patients with ASD.

ADHD is characterized by age-inappropriate inattentiveness, increased impulsivity and hyperactivity.[58] Heritability estimates in childhood are reported to be 70-80%, whereas estimates in adults show moderate heritability of 30-40%.[59] Several candidate genes have been associated with ADHD.[60] Although 10-28% of genetic risk is attributed to common variants,[61,62] the first risk loci with a high frequency have been reported recently.[63] Several of these loci are located near or in genes implicated in neurodevelopmental processes including *FOXP2* and *DUSP6*.[63] ADHD has been regarded as the extreme end of continuous distribution of inattentiveness and/or hyperactivity,[64-66] just like hypertension is the extreme of the continuous distribution of blood pressure in the population. As ADHD diagnosis is the extreme end of a continuous ADHD symptom scores[67] and genetic factors for ADHD diagnosis and ADHD symptoms showed an overlap,[67] novel more powerful approaches involving continuous measures in population-based setting could provide an opportunity to discover additional common variants and detect genes underlying ADHD. I aim to use this approach in order to evaluate contribution of common genetic variants in ADHD symptoms.

Furthermore, I have also studied Duchenne muscular dystrophy (DMD), the most common form of muscular dystrophy during childhood caused by mutations in dystrophin gene (*DMD*).[68] This fatal disease leads to progressive muscular weakness and less well described non-progressive central nervous system manifestations. As the risk of cognitive impairment is increased among the patients with DMD and higher occurrence of various neurodevelopmental disorders such as ASD and ADHD is also reported,[69-75] I address the question in this thesis whether *DMD* gene has an effect in general populations.

MOLECULAR APPROACHES USED IN THIS THESIS

To improve our understanding of the pathogenesis and heterogeneity in diseases and to facilitate development of personalized and more precise prevention and treatment, various omics approaches may be used to study changes underlying diseases at the molecular level. Omics approaches refer to large-scale high throughput technologies.[76,56] These technologies cover different molecular layers from the level of DNA (genomics) to DNA methylation/histone modification (epigenomics), RNA (transcriptomics), proteins (proteomics), and metabolites (metabolomics) as depicted in **Figure 1**. Furthermore, omics technologies also address microorganisms colonizing human body (microbiomics).

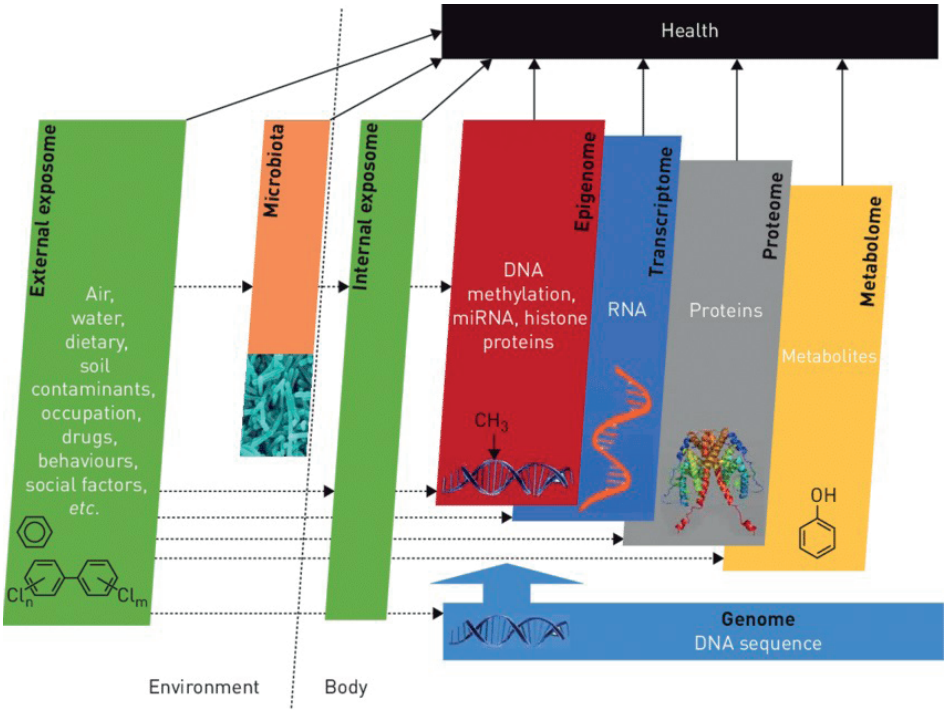

Figure 1. Diagram of omics layers and some of the interactions between them. Source: Siroux *et al.*[77]

In this thesis, I concentrated on several omics approaches including genomics, epigenomics, metabolomics, and microbiomics in relation to neurological and psychiatric disorders.

Genomics

The human genome captures all variations in our DNA, the blueprint of our proteins. Focusing on whole human genome, genomics provides important insights into genetic architecture of complex disorders, which involve effects of rare and common variants and variants conveying a small or large effect on pathology. Genetic determinants including single nucleotide polymorphisms (SNPs) or structural variation (SV) can be found in either protein-coding regions and may impact sequence of the protein or in non-coding regions more likely affecting gene expression and splicing processes.[78-80] Contribution of genetic variants commonly occurring in general population (minor allele frequency (MAF) > 5%) is often assessed by genome-wide association studies (GWAS).[81] The genetic variants often have a small effect on the trait. Although their individual effect is not informative, the joint effect is for a large part determining the risk of common diseases, as predicted by RA Fisher even before the structure of DNA was unraveled.[82] Thus, common variants provide important insights into the biology,

unravelling the pathological pathways, and jointly improve the proportion of variance explained by genetic factors, surpassing that of important epidemiological factors such as that of body mass index (BMI) on lipid levels.[83,84] Availability of relatively inexpensive SNP arrays and the possibility of imputing variants using large reference panels such as 1000 Genomes and Haplotype Reference Consortium (HRC), enabled the number of genetic variants for association testing to be increased and facilitated meta-analyses of studies using different arrays. [85-87] This resulted in mega GWAS of large sample size (currently up to a million).[88] A typical GWAS design involves hypothesis-free discovery study followed by replication of the associations in an independent sample.[89,90] Both

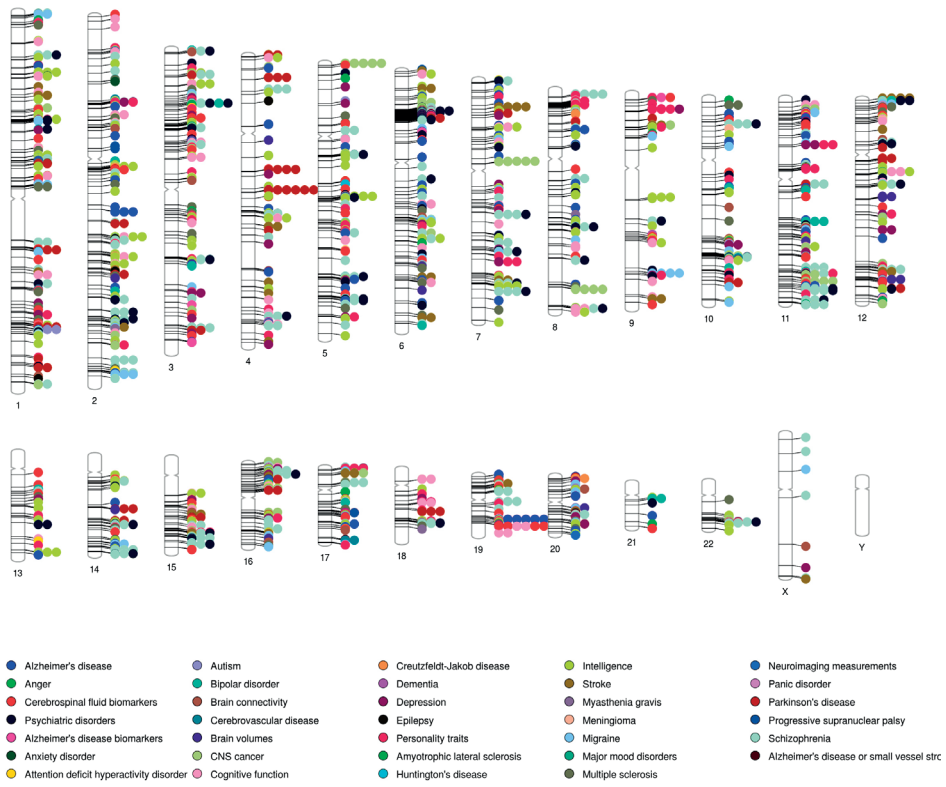

Figure 2. Associations of neurological and psychiatric disorders with SNPs accross the genome (GWAS catalog as of April 2018).[97] The genome is displayed divided into separate chromosomes. Color denotes disorders.

the discovery and replication are subjected to a stringent level of significance, adjusting for the large number of tests with the low *a priori* probability of association.[91] To date, more than 800 associations have been reported between the SNPs and neurological and psychiatric disorders (GWAS catalog as of April 2018) (**Figure 2**). Identified associations

not only confirmed previously identified genes (e.g. *APOE* locus was firstly identified in AD families followed by association analyses and later on replicated in GWAS)[92-94] but also identified novel genetic regions.[95] Additionally, GWAS provided opportunity to explore genetic architecture between the various complex disorders with methods such as LD score regression.[96]

The big data meta-analyses allowed to include more low-frequency and rare variants (MAF < 5%) in the GWAS.[98] However, there is a limit in that very rare variants are difficult to impute.[98] Thus, GWAS is unable to systematically explore the contribution of the rare variants which could also contribute to the genetic architecture and explain "missing heritability".[99,100] More importantly, these rare variants are key to personalized and precision prevention, e.g. as occurred in the prevention of breast cancer in *BRCA1/2* carriers through preventive mastectomy[101] and early mortality in carriers of *LDLR* mutations through treatment with statins starting in early adolescence.[102] Development of next-generation sequencing technologies including whole-genome sequencing (WGS) and whole-exome sequencing (WES) allowed detection of low-frequency or rare variants with large or moderate effects.[79,103] Applied to neurological and psychiatric disorders, some of the examples of success to date include discovery of rare coding variant in *TREM2* associated with AD,[104,105] rare variant in *VPS35* associated with Parkinson disease,[106,107] and several rare variants underlying the genetic etiology of ASD.[108] The development of dedicated rare variant arrays (e.g. the exome arrays), allowed the application of GWAS for rare variants in large datasets, i.e. as was successfully done for AD.[109] With increasing application to other disorders, more discoveries are underway, using both classical family-based methods as well as GWAS methodology.[110]

Epigenomics

Epigenomics focuses on genome-wide characterization of chemical modifications of DNA or DNA-associated proteins such as DNA methylation or histone modification.[111] Those modifications of DNA and histones play important role in the regulation of gene expression without changing the DNA sequence and are influenced by both genetic and environmental factors.[112] The most studied and best characterized epigenetic modification is DNA methylation -- addition of methyl group to the CpG sites of the DNA molecule. DNA methylation is essential for regulating X chromosome inactivation, genomic imprinting, and tissue-specific gene expression.[113,114] The pattern of DNA methylation established either during development[114] or late in life can have consequences within the brain. Abnormal methylation in *FMR1* gene causes mental retardation (Fragile X Syndrome),[115] whereas improper methylation of a single imprinted allele causes mental impairment (Prader-Willi Syndrome).[116,117] Late in life, environmental risk factors may have major impact, e.g. smoking and obesity-related pathologies are known to be

major determinants of expression.[118] With the development of epigenome-wide studies (EWAS), an opportunity to study DNA methylation pattern underlying complex neurological and psychiatric disorders has become available. Although methylation may be tissue-specific, there are many instances reported where there is a high correlation between the methylation in the brain and in blood.[119-121] Alteration of DNA methylation pattern has been observed in both psychiatric disorders such as schizophrenia and bipolar disorder and neurodegenerative disorders such as dementias.[122,123] Even though our understanding of the role of epigenetics in etiology of neurological and psychiatric disorders is still limited and may involve not only methylation but also acetylation in the brain,[124,125] epigenomics holds great potential for identifying useful biomarkers that could contribute to unraveling underlying mechanisms of these disorders. In addition to the etiological significance of methylation, one may speculate that methylation may possibly lead not only to timely diagnosis but also defining preclinical stages of disorders.

Metabolomics

The rapid development of new technologies enabled quantification of substrates and products of metabolism referred to as metabolites.[126] These low molecular weight compounds are influenced by genetic factors, lifestyle factors, pharmacological treatments, mechanisms of disease, and microbiota.[126,127] Last but not least, metabolites may reflect the disease process and may be a cause rather than a consequence of disease.

Identifying the metabolites and metabolic pathways has a potential to provide new insights into pathophysiology and for discovery of new diagnostic markers for disease risk that could facilitate the development of novel and precise diagnostic tools, and treatment and preventive strategies.[128,129] Metabolic profiling of biological fluids, including blood, urine, and cerebrospinal fluid, and tissues holds great potential for investigation of neurological and psychiatric disorders. Thousands of metabolites may be detected by targeted approaches, whereas this number increases if untargeted approaches are applied.[130] Although metabolite processes may be tissue specific, there is growing interest in vascular origin of neurodegeneration and cerebrovascular pathology. To date, metabolic profiling has been reported for various psychiatric disorders such as schizophrenia, bipolar disorder, and neurological conditions including AD and stroke.[24,131-137] However, not all studies performed to date were well powered, emphasizing need to explore metabolomics profiles in large epidemiological follow-up studies.

Microbiomics

Microbiomics focuses on microorganisms colonizing different parts of human body, such as skin (skin microbiota), the mouth (oral microbiota), the gut (gut microbiota) and

so on. The gut harbors thousands of microbial species which are considered to be a central signaling hub that integrates environmental inputs summarized as exposome (e.g. diet, life style, medication) with genetic and immune signals to affect the host's metabolism.[138] Gut microbiota is responsible for several functions including food digestion, vitamin and short chain acid (SCFA) production, amino acid synthesis, activation

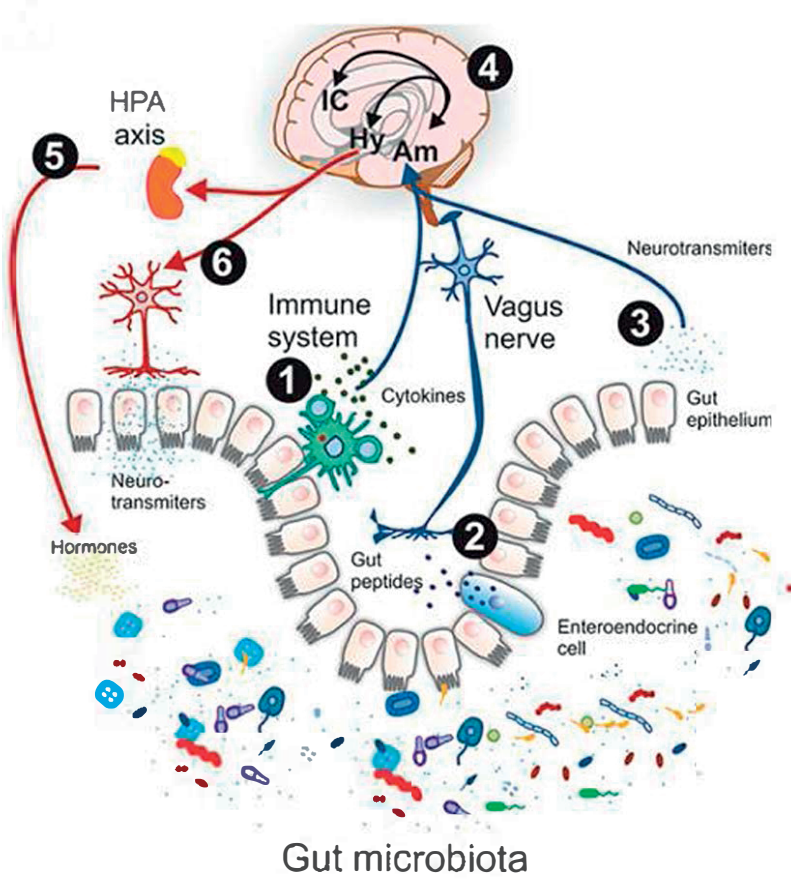

Figure 3. Bidirectional interaction between the gut microbiota and the central nervous system involving direct and indirect endocrine, immune and neural pathways. For instance: (1) cytokines released by lymphocytes which may sense the gut lumen can have endocrine or paracrine actions, (2) gut peptides released by enteroendocrine cells may activate sensory neuronal terminals, such as on the vagus nerve, (3) microbiota metabolites (neurotransmitters or its precursors) may reach the gut epithelium having endocrine or paracrine effects. (4) Centrally, after brainstem relays (e.g. nucleus *tractus solitarii*) a neural network involving the amygdala (Am) and the insular cortex (IC) integrates visceral inputs. Consistently hypothalamic (Hy) activation initiates: (5) corticosteroids release (results of the hypothalamic-pituitary-adrenal (HPA) axis activation) which modulates gut microbiota composition, (6) neuronal efferent activation ("anti-inflammatory cholinergic reflex" and/or sympathetic activation) liberating neurotransmitters that may affect the gut microbiota composition. Source: Montiel-Castro et al.[150]

of certain drugs, signaling molecules and anti-microbial compounds production, bile acid biotransformation and development of our immune system.[138-141] With advances in technology of microbial phenotyping methods, gut microbiota has been implicated in various neurological and psychiatric disorders and has been linked to cognitive ability, neurodevelopmental disorders (e.g. ASD), and neurodegenerative disorders (e.g. Parkinson disease, Alzheimer's disease).[142-148] The gut-brain axis has been long recognized. As depicted in **Figure 3** it involves metabolic and immune signals from the gut to the brain and vice versa from the brain to the gut and direct nerve innervation (nervus vagus).[148,149] Understanding mechanisms of complex nature of host-microbiome metabolism may help develop new strategies for preventing and treating diseases. In this thesis, we aim to explore link between gut microbiota and the metabolome.

AIM OF THIS THESIS

The aim of this thesis is to identify genomic and metabolomic determinants underlying neurological and psychiatric diseases and their related endophenotypes.

In **Chapter 2** omics studies of neurodegeneration are described. **Chapter 2.1** explores genetic determinants of brain structures determined by brain MRI. More specifically, I examine contribution of common genetic variants underlying lateral ventricular volume. Subsequently, other endophenotypes of neurological and psychiatric disorders are explored. Firstly, **Chapter 2.2** addresses common genetic determinants of general cognitive function and furthermore explores metabolic pathophysiology underlying established genetic variants implicated in cognitive ability. Then **Chapter 2.3** provides insights into complex DNA methylation signatures in relation to cognitive function. Finally, **Chapter 2.4** and **Chapter 2.5** apply candidate gene approach to study effect of rare variants mapped to a dystrophin gene on cognitive ability in general population and to determine whether the location of mutations in dystrophin gene and its impact on specific dystrophin isoforms has an effect on cognitive ability.

Chapter 3 addresses determinants of neurovascular pathology. In **Chapter 3.1**, contribution of rare genetic variants underlying carotid intima-media thickness is studied. Carotid intima-media thickness is a marker of subclinical atherosclerosis that predicts future cardiovascular events. **Chapter 3.2** addresses associations of metabolites measured by state-of-the-art metabolomics and carotid artery calcification, whereas **Chapter 3.3** focusses on metabolomic determinants of stroke in large prospective population-based studies including participants of European ancestry. **Chapter 3.4** provides insights into the relationship between gut microbiota and circulating metabolites.

Chapter 4 focusses on genetic determinants of neurodevelopmental disorders. **Chapter 4.1** explores genetic determinants in ASD, whereas the contribution of common genetic variants in ADHD symptoms is evaluated in **Chapter 4.2.**

Finally, **Chapter 5** summarizes the main findings of this thesis and provides suggestions for future research. **Chapter 5.1** describes major findings and in **Chapter 5.2** information derived from the genomic research of cardiovascular disorders is used to develop translational models aiming at effective prevention programs, earlier diagnosis and prognosis, and individualized treatments.

REFERENCES

1. Azevedo, F.A.C. *et al.* Equal Numbers of Neuronal and Nonneuronal Cells Make the Human Brain an Isometrically Scaled-Up Primate Brain. *Journal of Comparative Neurology* **513**, 532-541 (2009).
2. Qureshi, I.A. & Mehler, M.F. An evolving view of epigenetic complexity in the brain. *Philosophical Transactions of the Royal Society B-Biological Sciences* **369**(2014).
3. Whiteford, H.A., Ferrari, A.J., Degenhardt, L., Feigin, V. & Vos, T. The Global Burden of Mental, Neurological and Substance Use Disorders: An Analysis from the Global Burden of Disease Study 2010. *Plos One* **10**(2015).
4. Wittchen, H.U. *et al.* The size and burden of mental disorders and other disorders of the brain in Europe 2010. *Eur Neuropsychopharmacol* **21**, 655-79 (2011).
5. Group, G.B.D.N.D.C. Global, regional, and national burden of neurological disorders during 1990-2015: a systematic analysis for the Global Burden of Disease Study 2015. *Lancet Neurol* **16**, 877-897 (2017).
6. Burggren, A. & Brown, J. Imaging markers of structural and functional brain changes that precede cognitive symptoms in risk for Alzheimer's disease. *Brain Imaging and Behavior***.8**, pp (2014).
7. Vermeer, S.E. *et al.* Silent brain infarcts and the risk of dementia and cognitive decline. *N Engl J Med* **348**, 1215-22 (2003).
8. Bearden, C.E. & Freimer, N.B. Endophenotypes for psychiatric disorders: ready for primetime? *Trends Genet* **22**, 306-13 (2006).
9. Markus, H.S. Stroke genetics. *Hum Mol Genet* **20**, R124-31 (2011).
10. Bennett, D.A., De Jager, P.L., Leurgans, S.E. & Schneider, J.A. Neuropathologic intermediate phenotypes enhance association to Alzheimer susceptibility alleles. *Neurology* **72**, 1495-1503 (2009).
11. Kohler, S. *et al.* Progression to dementia in memory clinic patients without dementia: A latent profile analysis. *Neurology***.81**, pp (2013).
12. Gur, R.E. *et al.* Neurocognitive endophenotypes in a multiplex multigenerational family study of schizophrenia. *Am J Psychiatry* **164**, 813-9 (2007).
13. Hu, M.R. *et al.* Semantic fluency and executive functions as candidate endophenotypes for the early diagnosis of schizophrenia in Han Chinese. *Neuroscience Letters* **502**, 173-177 (2011).
14. Glahn, D.C. *et al.* Neurocognitive Endophenotypes for Bipolar Disorder Identified in Multiplex Multigenerational Families. *Archives of General Psychiatry* **67**, 168-177 (2010).
15. Shang, C.Y., Gau, S.S., Id, Shang, C.Y.O.h.o.o. & Gau, S.S.O.h.o.o. Visual memory as a potential cognitive endophenotype of attention deficit hyperactivity disorder. *Psychological Medicine***.41**, pp (2011).
16. Meschia, J.F., Worrall, B.B. & Rich, S.S. Genetic susceptibility to ischemic stroke. *Nature Reviews Neurology* **7**, 369-378 (2011).
17. Weinstein, G. *et al.* Brain Imaging and Cognitive Predictors of Stroke and Alzheimer Disease in the Framingham Heart Study. *Stroke* **44**, 2787-2794 (2013).
18. Burggren, A. & Brown, J. Imaging markers of structural and functional brain changes that precede cognitive symptoms in risk for Alzheimer's disease. *Brain imaging behav.* **8**, 251-61 (2014).
19. Malik, R. *et al.* Multiancestry genome-wide association study of 520,000 subjects identifies 32 loci associated with stroke and stroke subtypes. *Nat Genet* **50**, 524-537 (2018).
20. Andersen, K.K., Olsen, T.S., Dehlendorff, C. & Kammersgaard, L.P. Hemorrhagic and ischemic strokes compared: stroke severity, mortality, and risk factors. *Stroke* **40**, 2068-72 (2009).
21. Boehme, A.K., Esenwa, C. & Elkind, M.S. Stroke Risk Factors, Genetics, and Prevention. *Circ Res* **120**, 472-495 (2017).

22. Wang, D., Kong, J., Wu, J., Wang, X. & Lai, M. GC-MS-based metabolomics identifies an amino acid signature of acute ischemic stroke. *Neurosci Lett* **642**, 7-13 (2017).
23. Lee, Y., Khan, A., Hong, S., Jee, S.H. & Park, Y.H. A metabolomic study on high-risk stroke patients determines low levels of serum lysine metabolites: a retrospective cohort study. *Mol Biosyst* **13**, 1109-1120 (2017).
24. Jung, J.Y. *et al.* 1H-NMR-based metabolomics study of cerebral infarction. *Stroke* **42**, 1282-8 (2011).
25. Floegel, A. *et al.* Serum metabolites and risk of myocardial infarction and ischemic stroke: a targeted metabolomic approach in two German prospective cohorts. *Eur J Epidemiol* **33**, 55-66 (2018).
26. Holmes, M.V. *et al.* Lipids, Lipoproteins, and Metabolites and Risk of Myocardial Infarction and Stroke. *J Am Coll Cardiol* **71**, 620-632 (2018).
27. Jove, M. *et al.* Metabolomics predicts stroke recurrence after transient ischemic attack. *Neurology* **84**, 36-45 (2015).
28. Knopman, D.S. Dementia and cerebrovascular disease. *Mayo Clin Proc* **81**, 223-30 (2006).
29. Van Cauwenberghe, C., Van Broeckhoven, C. & Sleegers, K. The genetic landscape of Alzheimer disease: clinical implications and perspectives. *Genet Med* **18**, 421-30 (2016).
30. Olszewska, D.A., Lonergan, R., Fallon, E.M. & Lynch, T. Genetics of Frontotemporal Dementia. *Curr Neurol Neurosci Rep* **16**, 107 (2016).
31. Vergouw, L.J.M. *et al.* An update on the genetics of dementia with Lewy bodies. *Parkinsonism Relat Disord* **43**, 1-8 (2017).
32. Meeus, B., Theuns, J. & Van Broeckhoven, C. The genetics of dementia with Lewy bodies: what are we missing? *Arch Neurol* **69**, 1113-8 (2012).
33. Ibrahim-Verbaas, C.A. *et al.* GWAS for executive function and processing speed suggests involvement of the CADM2 gene. *Mol Psychiatry* **21**, 189-197 (2016).
34. Debette, S. *et al.* Genome-wide studies of verbal declarative memory in nondemented older people: the Cohorts for Heart and Aging Research in Genomic Epidemiology consortium. *Biol Psychiatry* **77**, 749-63 (2015).
35. Davies, G. *et al.* Genetic contributions to variation in general cognitive function: a meta-analysis of genome-wide association studies in the CHARGE consortium (N=53 949). *Molecular Psychiatry* **20**, 183-192 (2015).
36. Deary, I.J., Johnson, W. & Houlihan, L.M. Genetic foundations of human intelligence. *Hum Genet* **126**, 215-32 (2009).
37. Haworth, C.M. *et al.* The heritability of general cognitive ability increases linearly from childhood to young adulthood. *Mol Psychiatry* **15**, 1112-20 (2010).
38. Davies, G. *et al.* Genome-wide association study of cognitive functions and educational attainment in UK Biobank (N=112 151). *Mol Psychiatry* **21**, 758-67 (2016).
39. Davies, G. *et al.* Study of 300,486 individuals identifies 148 independent genetic loci influencing general cognitive function. *Nat Commun* **9**, 2098 (2018).
40. Anstey, K.J. Optimizing cognitive development over the life course and preventing cognitive decline: Introducing the Cognitive Health Environment Life Course Model (CHELM). *International Journal of Behavioral Development* **38**, 1-10 (2014).
41. Stein, J.L. *et al.* Identification of common variants associated with human hippocampal and intracranial volumes. *Nat Genet* **44**, 552-61 (2012).
42. Adams, H.H.H. *et al.* Novel genetic loci underlying human intracranial volume identified through genome-wide association. *Nature Neuroscience***.19**, pp (2016).

43. Wendell, C.R. *et al.* Carotid atherosclerosis and prospective risk of dementia. *Stroke* **43**, 3319-24 (2012).
44. Thapar, A., Cooper, M. & Rutter, M. Neurodevelopmental disorders. *Lancet Psychiatry* **4**, 339-246 (2017).
45. Huguet, G., Ey, E. & Bourgeron, T. The genetic landscapes of autism spectrum disorders. *Annu Rev Genomics Hum Genet* **14**, 191-213 (2013).
46. Hallmayer, J. *et al.* Genetic Heritability and Shared Environmental Factors Among Twin Pairs With Autism. *Archives of General Psychiatry* **68**, 1095-1102 (2011).
47. Bailey, A. *et al.* Autism as a Strongly Genetic Disorder - Evidence from a British Twin Study. *Psychological Medicine* **25**, 63-77 (1995).
48. Sandin, S. *et al.* The Familial Risk of Autism. *Jama-Journal of the American Medical Association* **311**, 1770-1777 (2014).
49. Gaugler, T. *et al.* Most genetic risk for autism resides with common variation. *Nature Genetics* **46**, 881-885 (2014).
50. Iossifov, I. *et al.* The contribution of de novo coding mutations to autism spectrum disorder. *Nature* **515**, 216-U136 (2014).
51. Pinto, D. *et al.* Functional impact of global rare copy number variation in autism spectrum disorders. *Nature* **466**, 368-372 (2010).
52. McClellan, J. & King, M.C. Genetic heterogeneity in human disease. *Cell* **141**, 210-7 (2010).
53. Wang, K. *et al.* Common genetic variants on 5p14.1 associate with autism spectrum disorders. *Nature* **459**, 528-533 (2009).
54. Ma, D.Q. *et al.* A Genome-wide Association Study of Autism Reveals a Common Novel Risk Locus at 5p14.1. *Annals of Human Genetics* **73**, 263-273 (2009).
55. Weiss, L.A., Arking, D.E. & Consortium, J.H.A. A genome-wide linkage and association scan reveals novel loci for autism. *Nature* **461**, 802-U62 (2009).
56. Anney, R. *et al.* Individual common variants exert weak effects on the risk for autism spectrum disorders. *Hum Mol Genet* **21**, 4781-92 (2012).
57. Anney, R.J.L. *et al.* Meta-analysis of GWAS of over 16,000 individuals with autism spectrum disorder highlights a novel locus at 10q24.32 and a significant overlap with schizophrenia. *Molecular Autism* **8**(2017).
58. American Psychiatric Association Diagnostic and Statistical Manual of Mental Disorders4th edn, revised. *American Psychiatric Press: Washington, DC* (2000).
59. Brikell, I., Kuja-Halkola, R. & Larsson, H. Heritability of attention-deficit hyperactivity disorder in adults. *Am J Med Genet B Neuropsychiatr Genet* **168**, 406-413 (2015).
60. Hawi, Z. *et al.* The molecular genetic architecture of attention deficit hyperactivity disorder. *Molecular Psychiatry* **20**, 289-297 (2015).
61. Cross-Disorder Group of the Psychiatric Genomics, C. *et al.* Genetic relationship between five psychiatric disorders estimated from genome-wide SNPs. *Nat Genet* **45**, 984-94 (2013).
62. Anttila, V. Analysis of shared heritability in common disorders of the brain. *Science* **360**.6395 (2018).
63. Demontis, D., Waters, R.K. & Martin, J. Discovery of the first genome-wide significant risk loci for ADHD. *bioRxiv* (2017).
64. Lubke, G.H., Hudziak, J.J., Derks, E.M., van Bijsterveldt, T.C. & Boomsma, D.I. Maternal ratings of attention problems in ADHD: evidence for the existence of a continuum. *J Am Acad Child Adolesc Psychiatry* **48**, 1085-93 (2009).

65. Levy, F., Hay, D.A., McStephen, M., Wood, C. & Waldman, I. Attention-deficit hyperactivity disorder: a category or a continuum? Genetic analysis of a large-scale twin study. *J Am Acad Child Adolesc Psychiatry* **36**, 737-44 (1997).
66. Larsson, H., Anckarsater, H., Rastam, M., Chang, Z. & Lichtenstein, P. Childhood attention-deficit hyperactivity disorder as an extreme of a continuous trait: a quantitative genetic study of 8,500 twin pairs. *Journal of Child Psychology and Psychiatry* **53**, 73-80 (2012).
67. Middeldorp, C.M. *et al.* A Genome-Wide Association Meta-Analysis of Attention-Deficit/Hyperactivity Disorder Symptoms in Population-Based Pediatric Cohorts. *J Am Acad Child Adolesc Psychiatry* **55**, 896-905 e6 (2016).
68. Hoffman, E.P., Brown, R.H. & Kunkel, L.M. Dystrophin - the Protein Product of the Duchenne Muscular-Dystrophy Locus. *Cell* **51**, 919-928 (1987).
69. Cyrulnik, S.E., Fee, R.J., De Vivo, D.C., Goldstein, E. & Hinton, V.J. Delayed developmental language milestones in children with Duchenne's muscular dystrophy. *Journal of Pediatrics* **150**, 474-478 (2007).
70. Wicksell, R.K., Kihlgren, M., Melin, L. & Eeg-Olofsson, O. Specific cognitive deficits are common in children with Duchenne muscular dystrophy. *Developmental Medicine and Child Neurology* **46**, 154-159 (2004).
71. Hinton, V.J., De Vivo, D.C., Nereo, N.E., Goldstein, E. & Stern, Y. Poor verbal working memory across intellectual level in boys with Duchenne dystrophy. *Neurology* **54**, 2127-2132 (2000).
72. Mento, G., Tarantino, V. & Bisiacchi, P.S. The Neuropsychological Profile of Infantile Duchenne Muscular Dystrophy. *Clinical Neuropsychologist* **25**, 1359-1377 (2011).
73. D'Angelo, M.G. *et al.* Neurocognitive Profiles in Duchenne Muscular Dystrophy and Gene Mutation Site. *Pediatric Neurology* **45**, 292-299 (2011).
74. Hendriksen, J.G.M. & Vles, J.S.H. Neuropsychiatric disorders in males with Duchenne muscular dystrophy: Frequency rate of attention-deficit hyperactivity disorder (ADHD), autism spectrum disorder, and obsessive-compulsive disorder. *Journal of Child Neurology* **23**, 477-481 (2008).
75. Wu, J.Y., Kuban, K.C., Allred, E., Shapiro, F. & Darras, B.T. Association of Duchenne muscular dystrophy with autism spectrum disorder. *J Child Neurol* **20**, 790-5 (2005).
76. Valdes, A.M., Glass, D. & Spector, T.D. Omics technologies and the study of human ageing. *Nat Rev Genet* **14**, 601-7 (2013).
77. Siroux, V., Agier, L. & Slama, R. The exposome concept: a challenge and a potential driver for environmental health research. *Eur Respir Rev* **25**, 124-9 (2016).
78. Gonzaga-Jauregui, C., Lupski, J.R. & Gibbs, R.A. Human genome sequencing in health and disease. *Annu Rev Med* **63**, 35-61 (2012).
79. Metzker, M.L. Sequencing technologies - the next generation. *Nat Rev Genet* **11**, 31-46 (2010).
80. Zhang, F. & Lupski, J.R. Non-coding genetic variants in human disease. *Hum Mol Genet* **24**, R102-10 (2015).
81. Visscher, P.M. *et al.* 10 Years of GWAS Discovery: Biology, Function, and Translation. *Am J Hum Genet* **101**, 5-22 (2017).
82. Bodmer, W. & Bonilla, C. Common and rare variants in multifactorial susceptibility to common diseases. *Nat Genet* **40**, 695-701 (2008).
83. Yang, J. *et al.* Common SNPs explain a large proportion of the heritability for human height. *Nat Genet* **42**, 565-9 (2010).
84. Aulchenko, Y.S. *et al.* Loci influencing lipid levels and coronary heart disease risk in 16 European population cohorts. *Nat Genet* **41**, 47-55 (2009).

85. Bush, W.S. & Moore, J.H. Chapter 11: Genome-wide association studies. *PLoS Comput Biol* **8**, e1002822 (2012).
86. Genomes Project, C. *et al.* An integrated map of genetic variation from 1,092 human genomes. *Nature* **491**, 56-65 (2012).
87. McCarthy, S. *et al.* A reference panel of 64,976 haplotypes for genotype imputation. *Nat Genet* **48**, 1279-83 (2016).
88. Evangelou, E. *et al.* Genetic analysis of over one million people identifies 535 novel loci for blood pressure. *bioRxiv* (2017).
89. Pe'er, I., Yelensk, R., Altshuler, D. & Daly, M.J. Estimation of the multiple testing burden for genome-wide association studies of nearly all common variants. *Genetic Epidemiology* **32**, 381-385 (2008).
90. Clayton, D.G. *et al.* Population structure, differential bias and genomic control in a large-scale, case-control association study. *Nature Genetics* **37**, 1243-1246 (2005).
91. Broer, L. *et al.* Distinguishing true from false positives in genomic studies: p values. *Eur J Epidemiol* **28**, 131-8 (2013).
92. Corder, E.H. *et al.* Gene dose of apolipoprotein E type 4 allele and the risk of Alzheimer's disease in late onset families. *Science* **261**, 921-3 (1993).
93. Pericak-Vance, M.A. *et al.* Linkage studies in familial Alzheimer disease: evidence for chromosome 19 linkage. *Am J Hum Genet* **48**, 1034-50 (1991).
94. Saunders, A.M. *et al.* Association of apolipoprotein E allele epsilon 4 with late-onset familial and sporadic Alzheimer's disease. *Neurology* **43**, 1467-72 (1993).
95. Lambert, J.C. *et al.* Meta-analysis of 74,046 individuals identifies 11 new susceptibility loci for Alzheimer's disease. *Nature Genetics* **45**, 1452-U206 (2013).
96. Bulik-Sullivan, B.K. *et al.* LD Score regression distinguishes confounding from polygenicity in genome-wide association studies. *Nat Genet* **47**, 291-5 (2015).
97. MacArthur, J. *et al.* The new NHGRI-EBI Catalog of published genome-wide association studies (GWAS Catalog). *Nucleic Acids Res* **45**, D896-D901 (2017).
98. Iglesias, A.I. *et al.* Haplotype reference consortium panel: Practical implications of imputations with large reference panels. *Human Mutation* **38**, 1025-1032 (2017).
99. Price, A.L., Spencer, C.C. & Donnelly, P. Progress and promise in understanding the genetic basis of common diseases. *Proc Biol Sci* **282**, 20151684 (2015).
100. Maher, B. Personal genomes: The case of the missing heritability. *Nature* **456**, 18-21 (2008).
101. Olopade, O.I., Grushko, T.A., Nanda, R. & Huo, D. Advances in breast cancer: pathways to personalized medicine. *Clin Cancer Res* **14**, 7988-99 (2008).
102. Neil, A. *et al.* Reductions in all-cause, cancer, and coronary mortality in statin-treated patients with heterozygous familial hypercholesterolaemia: a prospective registry study. *Eur Heart J* **29**, 2625-33 (2008).
103. Guerreiro, R., Bras, J., Hardy, J. & Singleton, A. Next generation sequencing techniques in neurological diseases: redefining clinical and molecular associations. *Human Molecular Genetics* **23**, R47-R53 (2014).
104. Guerreiro, R. *et al.* TREM2 variants in Alzheimer's disease. *N Engl J Med* **368**, 117-27 (2013).
105. Jonsson, T. *et al.* Variant of TREM2 associated with the risk of Alzheimer's disease. *N Engl J Med* **368**, 107-16 (2013).
106. Zimprich, A. *et al.* A mutation in VPS35, encoding a subunit of the retromer complex, causes late-onset Parkinson disease. *Am J Hum Genet* **89**, 168-75 (2011).
107. Vilarino-Guell, C. *et al.* VPS35 mutations in Parkinson disease. *Am J Hum Genet* **89**, 162-7 (2011).

108. Krumm, N., O'Roak, B.J., Shendure, J. & Eichler, E.E. A de novo convergence of autism genetics and molecular neuroscience. *Trends in Neurosciences* **37**, 95-105 (2014).
109. Sims, R. *et al.* Rare coding variants in PLCG2, ABI3, and TREM2 implicate microglial-mediated innate immunity in Alzheimer's disease. *Nature Genetics* **49**, 1373-+ (2017).
110. Sanders, S.J. *et al.* Whole Genome Sequencing in Psychiatric Disorders: the WGSPD Consortium. *bioRxiv* (2017).
111. Hasin, Y., Seldin, M. & Lusis, A. Multi-omics approaches to disease. *Genome Biology* **18** (2017).
112. Smith, Z.D. & Meissner, A. DNA methylation: roles in mammalian development. *Nature Reviews Genetics* **14**, 204-220 (2013).
113. Moore, L.D., Le, T. & Fan, G. DNA methylation and its basic function. *Neuropsychopharmacology* **38**, 23-38 (2013).
114. Smith, Z.D. & Meissner, A. DNA methylation: roles in mammalian development. *Nat Rev Genet* **14**, 204-20 (2013).
115. Verkerk, A.J.M.H. *et al.* Identification of a Gene (Fmr-1) Containing a Cgg Repeat Coincident with a Breakpoint Cluster Region Exhibiting Length Variation in Fragile-X Syndrome. *Cell* **65**, 905-914 (1991).
116. Buiting, K. Prader-Willi Syndrome and Angelman Syndrome. *American Journal of Medical Genetics Part C-Seminars in Medical Genetics* **154c**, 365-376 (2010).
117. Moore, L.D., Le, T. & Fan, G.P. DNA Methylation and Its Basic Function. *Neuropsychopharmacology* **38**, 23-38 (2013).
118. Zeilinger, S. *et al.* Tobacco Smoking Leads to Extensive Genome-Wide Changes in DNA Methylation. *Plos One* **8** (2013).
119. Ewald, E.R. *et al.* Alterations in DNA methylation of Fkbp5 as a determinant of blood-brain correlation of glucocorticoid exposure. *Psychoneuroendocrinology* **44**, 112-22 (2014).
120. Horvath, S. *et al.* Aging effects on DNA methylation modules in human brain and blood tissue. *Genome Biol* **13**, R97 (2012).
121. Masliah, E., Dumaop, W., Galasko, D. & Desplats, P. Distinctive patterns of DNA methylation associated with Parkinson disease: identification of concordant epigenetic changes in brain and peripheral blood leukocytes. *Epigenetics* **8**, 1030-8 (2013).
122. Mill, J. *et al.* Epigenomic profiling reveals DNA-Methylation changes associated with major psychosis. *American Journal of Human Genetics* **82**, 696-711 (2008).
123. Fransquet, P.D. *et al.* Blood DNA methylation as a potential biomarker of dementia: A systematic review. *Alzheimers & Dementia* **14**, 81-103 (2018).
124. Konsoula, Z. & Barile, F.A. Epigenetic histone acetylation and deacetylation mechanisms in experimental models of neurodegenerative disorders. *J Pharmacol Toxicol Methods* **66**, 215-20 (2012).
125. Lu, X., Wang, L., Yu, C.J., Yu, D.H. & Yu, G. Histone acetylation modifiers in the pathogenesis of Alzheimer's disease. *Frontiers in Cellular Neuroscience* **9**(2015).
126. Johnson, C.H., Ivanisevic, J. & Siuzdak, G. Metabolomics: beyond biomarkers and towards mechanisms. *Nat Rev Mol Cell Biol* **17**, 451-9 (2016).
127. Menni, C., Zierer, J., Valdes, A.M. & Spector, T.D. Mixing omics: combining genetics and metabolomics to study rheumatic diseases. *Nat Rev Rheumatol* **13**, 174-181 (2017).
128. Kaddurah-Daouk, R., Kristal, B.S. & Weinshilboum, R.M. Metabolomics: a global biochemical approach to drug response and disease. *Annu Rev Pharmacol Toxicol* **48**, 653-83 (2008).
129. Guest, P.C., Guest, F.L. & Martins-de Souza, D. Making Sense of Blood-Based Proteomics and Metabolomics in Psychiatric Research. *Int J Neuropsychopharmacol* **19**(2016).
130. Markley, J.L. *et al.* The future of NMR-based metabolomics. *Curr Opin Biotechnol* **43**, 34-40 (2017).

131. McIntyre, R.S. *et al.* Advancing biomarker research: utilizing 'Big Data' approaches for the characterization and prevention of bipolar disorder. *Bipolar Disorders* **16**, 531-547 (2014).
132. Paredes, R.M. *et al.* Metabolomic profiling of schizophrenia patients at risk for metabolic syndrome. *International Journal of Neuropsychopharmacology* **17**, 1139-1148 (2014).
133. Pickard, B.S. Schizophrenia biomarkers: Translating the descriptive into the diagnostic. *Journal of Psychopharmacology* **29**, 138-143 (2015).
134. Guest, P.C., Guest, F.L. & Martins-de Souza, D. Making Sense of Blood-Based Proteomics and Metabolomics in Psychiatric Research. *International Journal of Neuropsychopharmacology* **19** (2016).
135. Tynkkynen, J. *et al.* Association of branched-chain amino acids and other circulating metabolites with risk of incident dementia and Alzheimer's disease: A prospective study in eight cohorts. *Alzheimers Dement* (2018).
136. Sethi, S. & Brietzke, E. Omics-Based Biomarkers: Application of Metabolomics in Neuropsychiatric Disorders. *Int J Neuropsychopharmacol* **19**, pyv096 (2015).
137. van der Lee, S.J. *et al.* Circulating metabolites and general cognitive ability and dementia: Evidence from 11 cohort studies. *Alzheimers Dement* (2018).
138. Thaiss, C.A., Zmora, N., Levy, M. & Elinav, E. The microbiome and innate immunity. *Nature* **535**, 65-74 (2016).
139. Vernocchi, P., Del Chierico, F. & Putignani, L. Gut Microbiota Profiling: Metabolomics Based Approach to Unravel Compounds Affecting Human Health. *Front Microbiol* **7**, 1144 (2016).
140. Conlon, M.A. & Bird, A.R. The impact of diet and lifestyle on gut microbiota and human health. *Nutrients* **7**, 17-44 (2014).
141. Holmes, E., Li, J.V., Athanasiou, T., Ashrafian, H. & Nicholson, J.K. Understanding the role of gut microbiome-host metabolic signal disruption in health and disease. *Trends Microbiol* **19**, 349-59 (2011).
142. Hsiao, E.Y. *et al.* Microbiota modulate behavioral and physiological abnormalities associated with neurodevelopmental disorders. *Cell* **155**, 1451-63 (2013).
143. Sampson, T.R. *et al.* Gut Microbiota Regulate Motor Deficits and Neuroinflammation in a Model of Parkinson's Disease. *Cell* **167**, 1469-1480 e12 (2016).
144. Rogers, G.B. *et al.* From gut dysbiosis to altered brain function and mental illness: mechanisms and pathways. *Mol Psychiatry* **21**, 738-48 (2016).
145. Vogt, N.M. *et al.* Gut microbiome alterations in Alzheimer's disease. *Scientific Reports* **7**(2017).
146. Fernandez-Real, J.M. *et al.* Gut Microbiota Interacts With Brain Microstructure and Function. *Journal of Clinical Endocrinology & Metabolism* **100**, 4505-4513 (2015).
147. Cattaneo, A. *et al.* Association of brain amyloidosis with pro-inflammatory gut bacterial taxa and peripheral inflammation markers in cognitively impaired elderly. *Neurobiology of Aging* **49**, 60-68 (2017).
148. Fung, T.C., Olson, C.A. & Hsiao, E.Y. Interactions between the microbiota, immune and nervous systems in health and disease. *Nature Neuroscience* **20**, 145-155 (2017).
149. Powell, N., Walker, M.M. & Talley, N.J. The mucosal immune system: master regulator of bidirectional gut-brain communications. *Nat Rev Gastroenterol Hepatol* **14**, 143-159 (2017).
150. Montiel-Castro, A.J., Gonzalez-Cervantes, R.M., Bravo-Ruiseco, G. & Pacheco-Lopez, G. The microbiota-gut-brain axis: neurobehavioral correlates, health and sociality. *Front Integr Neurosci* **7**, 70 (2013).