Aetiology of Depression:

Insights from epidemiological and genetic research

Olivera Story-Jovanova

Acknowledgments: Financial support for the publication of this thesis by the Department of Epidemiology of the Erasmus MC, is gratefully acknowledged.

ISBN: 978-94-6233-909-5

Cover: Concept by Olivera Story-Jovanova, design by Chris van Wolferen-Ketel, photography by Michal Macku.

Layout: Gildeprint, Enschede.

Printing: Gildeprint, Enschede.

Aetiology of Depression:
Insights from epidemiological and genetic research

Etiologie van depressie:
Inzichten vanuit de epidemiologisch en de genetisch onderzoek

Proefschrift

ter verkrijging van de graad van doctor aan de
Erasmus Universiteit Rotterdam
op gezag van de
rector magnificus

Prof. dr. H.A.P. Pols
en volgens besluit van het College voor Promoties.

De openbare verdediging zal plaatsvinden op
4 April 2018 om 15:30 uur

door
Olivera Story-Jovanova
geboren te Skopje, Macedonia

Erasmus University Rotterdam

Erasmus

PROMOTIECOMMISSIE

Promotor	Prof. dr. H. Tiemeier
Overige leden	Prof. D. Boomsma Prof. K. Berger Prof. C. van Duijn
Copromotor	Dr. N. Amin

Paranimfen	Ayesha Sajjad Dina Atlas

For my husband, children, sister, parents, grandparents, my parents in law and all you who believed in me. You are a gift of unconditional love, acceptance, joy and wisdom.

I am thankful for that!

ACKNOWLEDGEMENTS

The research described in this thesis was performed within the frame work of the Rotterdam Study.

The contribution of participants, the staff from the Rotterdam Study and the participating general practitioners, and pharmacists is gratefully acknowledged. The Rotterdam Study is supported by the Erasmus Medical Center and the Erasmus University, Rotterdam, the Netherlands; the Organization for the Health Research and Development (ZonMw); the Research Institute for Diseases in the Elderly (RIDE); the Ministry of Education, Culture and Science, the Ministry for Health, Welfare, and Sports; the European Commission (DG XII); and the Municipality Rotterdam. The research was supported by a Netherlands Organization for Scientific Research grant (NOW-ZonMw VIDI grant no. 017.106.370) awarded to prof. H. Tiemeier and by the European Commission – ERAWEB Grant (grant no. 2011-2586/001-001-EMA2) awarded to O. Jovanova, MD.

CONTENTS

MANUSCRIPTS UPON WHICH THIS THESIS IS BASED

Chapter 2.1: **Jovanova O**, Aarts N, Noordam R, Carola-Zillekens M, Hofman A, Tiemeier H. Vitamin D serum levels are cross-sectionally but not prospectively associated with late-life depression. *ACTA Psychiatrica Scandinavica*. 2017 Mar; 135(3):185-194.

Chapter 2.2: Zalli A, **Jovanova O,** Hoogendijk WJ, Tiemeier H, Carvalho LA. Low-grade inflammation predicts persistence of depressive symptoms. *Psychopharmacology*. 2016 May; 233(9):1669-78.

Chapter 3.1: Amin N, **Jovanova O,** Adams HH, Dehghan A, Kavousi M, Vernooji MW, Peeters RP, de Vrij FM, van der Lee SJ, van Rooij JG, van Leeuwen EM, Chaker L, Demirkan A, Hofman A, Brouwer RW, Kraaij R, Willems van Dijk K, Hankemeier T, van Ijcken WF, Uitterlinden AG, Niessen WJ, Franco OH, Kushner SA, Ikram MA, Tiemeier H, van Duijn CM. Exome-sequencing in a large population-based study reveals a rare Asn396Ser variant in the LIPG gene associated with depressive symptoms. *Molecular Psychiatry*. 2017 Apr; 22(4):537-543.

Chapter 3.2: Amin N, Belonogova NM, **Jovanova O**, Brouwer RW, van Rooij JG, van den Hout MC, Svishcheva GR, Kraaij R, Zorkoltseva IV, Kirichenko AV, Hofman A, Uitterlinden AG, van IJcken WF, Tiemeier H, Axenovich TI, van Duijn CM. Nonsynonymous Variation in NKPD1 Increases Depressive Symptoms in the European Populations. *Biological Psychiatry*. 2017 Apr; 81(8):702-707.

Chapter 4.1: **Jovanova O**, Nedeljkovic I, Lemaitre R, Brody J, Swenson B, Liu Ch, Hongsheng, Lahti J, Kunze S, Kuhnel B, Luciano M, Deary IJ, Marioni R, Walker RM, Evans KL, Wang Z, Gondalia R, Levy D, Seshadri S, Karl-Heintz L, Waldenberg M, McRae AF, Starr JM, Wray N, Bressler J, Mosley TH, Guo X, Sotoodehnia N, Mendelson MM, Peters A, Russ TC, McIntosh AM, Porteous DJ, Fornage M, Whitsel AE, Tiemeier H, Amin N. DNA methylation signatures of depressive symptoms identified in a large multi-ethnic meta-analysis of epigenome-wide studies. *Under review.*

Chapter 5.1: **Jovanova O**, Wolters FJ, Ikram MA, Tiemeier H, Schmitz N. DNA methylation signatures of depressive symptoms identified in a large multi-ethnic meta-analysis of epigenome-wide studies. *Under review.*

Chapter 5.2: **Jovanova O**, Luik AI, Leening MJG, Noordam R, Aarts N, Hofman A, Franco OH, Dehghan A, Tiemeier H. The long-term risk of recognized and unrecognized myocardial infarction for depression in older men. *Psychological Medicine*. 2016 Mar; 46(9):1951-60.

CHAPTER 1

General Introduction

INTRODUCTION

Have you ever felt depressed? The answer to this question is usually "YES" and this is not unreasonable since most of us have felt depressed once or more in our life-time. Typically, "feeling depressed" or "feeling BLUE" is not more than the well-accepted low mood or just being temporarily unhappy. On the contrary, clinical depression is a serious neuropsychiatric mood disorder that has a leading role in the global burden of diseases causing much of the disability world-wide.[1,2] Individuals suffering from depression have feelings of guilt, helplessness, hopelessness, are occupationally impaired, often have desire for social withdrawal, suffer from sleep and concentration disturbances, have a loss of interest in life-pleasures and almost all other activities, and suffer from a loss of appetite and sex drive.[3] This extensive list of depressive symptoms can also include daily experience of suicidal thoughts that can result in a suicidal act.[3] As a matter of fact, depression is a complete contrast to the beauty of the colour BLUE; a colour that awakens feelings of tranquillity, stability and inspiration (Yves Klein dedicated a life-time of work to the colour blue: "Blue Revolution").

®"Blue Monochrome (anthropometry)" - Yves Klein 1928-1962

In the last 15 years a number of successful campaigns took place around the world to raise the awareness for depression.[4,5] One of the main goals of such campaigns was to reduce the stigma of depression and to clarify the differences between being sad and suffering from depressive disorders. These goals were mostly achieved and the knowledge over depressive illness among the public has increased. More persons have some understanding of the physical and mental exhaustion caused by this medical illness. However, the reality is: "We still do not completely understand all aspects of depression"! Even though depression may not be an enigma anymore, there are many core questions asked by lay people, medical doctors, and academic experts that remain unclear. One such question is "What actually causes depression?"! Briefly, this thesis is an attempt to unravel some aspects of the aetiology of depression.

IS DEPRESSION the DIABETES OF the BRAIN or a DEMONIC POSSESSION? A brief historical aspect

I would first like to quote the famous Chinese philosopher Confucius who said "Study the past if you would define the future". Indeed, if we try to obtain new insights into the aetiology of depression and draw conclusions on this illness; and finally define some future perspectives for a better understanding of depression, we should certainly understand the history of this disorder. Throughout history, depression transformed from the old concept of melancholia to the current concept of depression which is mainly viewed as a multifactorial behavioural mental disorder. This transition did not occur all of a sudden, but throughout the centuries, many philosophers and scientists struggled explaining depression while facing a lot of controversies. This explains why studying depression and its aetiology is a great challenge even today.[6]

It was in the ancient times that the first theories were postulated to define and explain depression: Melancholia, a condition characterized by fear, loss-of appetite, and sleeplessness. The humoral theory proposed by Hippocrates (370-460 B.C), the father of medicine; explained melancholia.[7] According to the humoral concept melancholia was a condition caused by disequilibrium between the four humors and more specifically by an increase in the black bile.[7] The core of this theory was slightly modified and edited during the centuries to come. It took almost 2,500 years to move from Hippocrates melancholia to Emil Kraepelin's concept of depression (1856 –1926). He was the first scientist to propose the use of the term depression, but what made it so difficult to progress to defining depression as a disorder?[8]

In order to define this medical condition as a disorder, depression must be associated with specific symptoms and signs that are caused by external factors and internal dysfunctions. Scientists, doctors and people described and classified the symptoms of depression to a moderate extent. However, they faced difficulties in explaining and understanding what causes depression.

Depression was mainly viewed as a mental insanity characterized by personality changes and altered/disturbing emotions that did not have a clear cause. This limited understanding in what causes depression restrained the process of defining depression as a mental disorder.

In the golden scientific ages of the 19th and the 20th century, mental diseases for the first time, were proposed as diseases of the brain by Wilhelm Griesinger (1817-1869).[10] Krapelin's "unitary concept of depression" was the first theory to define depression as a unitary endogenous disorder.[11] Based on clinical observations, this concept emphasized depression as a specific psychiatric illness characterized by a combination of several symptoms with a specific organic aetiology and pathology. Krapelin's categorization of mental diseases established the basic tool for the widely-used classifications of mental disorders by the American Psychiatric Association (DSM-Classification)[3] and the World Health Organization (International classification of Disease)[12]. However, these classifications do not account for the aetiology of the disease. A group of scientists, e.g. Meyer, refused to regard depression as a biological disease only and proposed to view depression as a psychobiological reaction of the human body to stress.[13] Whether depression is a reactive answer to stress, a systemic disease, or both remains the object of discussion until today. Solving the aetiological puzzles would help to disentangle the true causes of depression and improve diagnosis.

Over the past 60 years, a large number of scientists spent literally hundreds of thousands of research hours, producing and publishing a substantial number of scientific articles on depression. This enormous work helped in better understanding the disease. One of the most remarkable scientific discovery of the last century (1952) was detecting the presence of serotonin in the brain and this discovery started the so-called antidepressant revolution.[14] Betty Twarog was the first scientist to confirm that indeed depression is the diabetes of the brain and supported the biological theory defining depression as a mental disorder caused by an imbalance in brain neurotransmitters.[15] This start of the era of antidepressant drugs, inspired many researchers to reconcile different aetiologic orientations of depression, from social-psychodynamic to biological aspects. Due to this work, many questions related to the causes of depression have been answered. However, the knowledge over the aetiology of depression remains patchy. Important elements in the aetiology of depression are still missing.

What do we really know about the AETIOLOGY OF DEPRESSION? – Epidemiological aspects

Depression is generally seen as a bio-psychosocial disease. Thus, there is no single explanation of what causes depression, and no minimalist aetiology could capture the complexity of this disorder. The academic world widely accepts depression as a complex multifactorial disease that

develops as a result of interaction and accumulation of various different psychosocial, biological and environmental risk factors. Psychosocial factors such as traumatic early-childhood events[16] such as abuse, socio-economic status[17], marital status[18], and loss of a partner[19] are all established aetiologic risk factors involved in the pathogenesis of depression. These factors were observed by clinicians and are also related to depression in population-based studies. But these associations do not necessarily imply causality between two entities and the typically cross-sectional study design may only answer questions such as "Are older persons more likely to be depressed?" On the contrary, longitudinal study designs, such as those included in this thesis are more informative when we try to infer causality. Specifically, these designs help answer temporal questions: "Do events such as abuse precede depression?". For a factor to be considered causal to depression, the time between that factor and the consequence is important and that factor has to precede the depressive event. Indeed, this conceptualizes the well-known criterion for temporality by Bradford Hill that perhaps is the only criterion which epidemiologists universally agree on and is essential to infer causality.[20] Longitudinal studies are still relatively rare when studying aetiologic factors of depression and therefore most studies presented in this thesis are focused on investigations performed within a longitudinal framework.

Various biological factors such as neuroendocrine, neuro-immunological, and genetic, have been related to depression and have an important role in the complex aetiology of this disease.[21-23] However, whether these factors are in the causal path to development of depression or they appear as a consequence of depression it is still unclear. In this thesis we carefully study few blood extracted neuro-inflammatory markers such as interleukin (IL)-6, alpha-1-antichymotrypsin (ACT) and C reactive protein (CRP) as well as serum vitamin D levels (a neuroendocrine factor) and their impact on the development of depression. These potential biomarkers could reflect a disease cause, biological signals of a pathophysiological processes related to extraneous factors, or response to a therapeutic intervention specific to depression, or a related condition.[21] Determining such biomarkers for depression could one day have an influential role in clinical practice and may increase the possibility for early detection, treatment and successful management of depression.[24]

Moreover, psychiatric epidemiology traditionally showed more interest in studying risk factors for depression than studying the consequences of depression for health in the general population. Taking into account that depression threatens to become the leading global cause of disability[2], the interest to study the long-term consequences of depression rises. Therefore, one of the studies described in this thesis studies the cognitive decline that appears as a long-term consequence in persons suffering from depression.

In order to completely understand the biology and the pathophysiology of depression, we need to identify genetic loci susceptible of depression, determine the genetic risk, and understand the

involvement of those genes in the pathology of depression. The technical development in the last years has allowed easy and cheap DNA sequencing, thus studying the genetics of depression on a large epidemiologic scale became easily accessible. Many studies aimed to identify common genetic variation involved in depression using large scale genome-wide association studies (GWAS) were performed.[25] Only recently, one study showed that 44 different common genetic variants are associated with depression (study under review). Compared to other outcomes studied with GWAS medical studies, those studying depression are challenging and with moderate success only. Genetic epidemiological research using genotyping arrays approach has detected common variants with small effect sizes which may not be appropriate when studying heterogenetic traits such as depressive disorder.[26] In addition to the common genetic variants, rare variants with substantial effect sizes may also be involved in the development of depression. Therefore, exome-sequencing and exome-chip genotyping methods were proposed as a good solution in order to identify rare genetic variants associated with depression.[27] This thesis presents two studies that employ such methods in discovering rare variants with possible large effects on depression in the general population.

Depression is a mental disorder with an estimated heritability from 30 to 40%, and most of the risk for depression is explained by environmental factors.[22] A complex interplay between environment and genetics is conceptualized to increase neurodevelopmental processes involved in depression.[28] The basis of this hypothesis, as well as the moderate success in determining the genetic risk for depression, increased the interest in studying the epigenetics of depression. A few investigations in post-mortem depression patients have performed epigenome wide association studies (EWAS)[29], however a population-based study has not yet been conducted. Such an approach is essential in detecting epigenetic associations for psychiatric conditions given the lack of prior knowledge.[22] Therefore, in Chapter 4 we present the first and the largest population-based EWAS of depression.

AIMS

This thesis includes several population-based studies that explore the aetiology of depression, with a specific interest on biological factors, genetics and epigenetics, and physical health factors for depression.

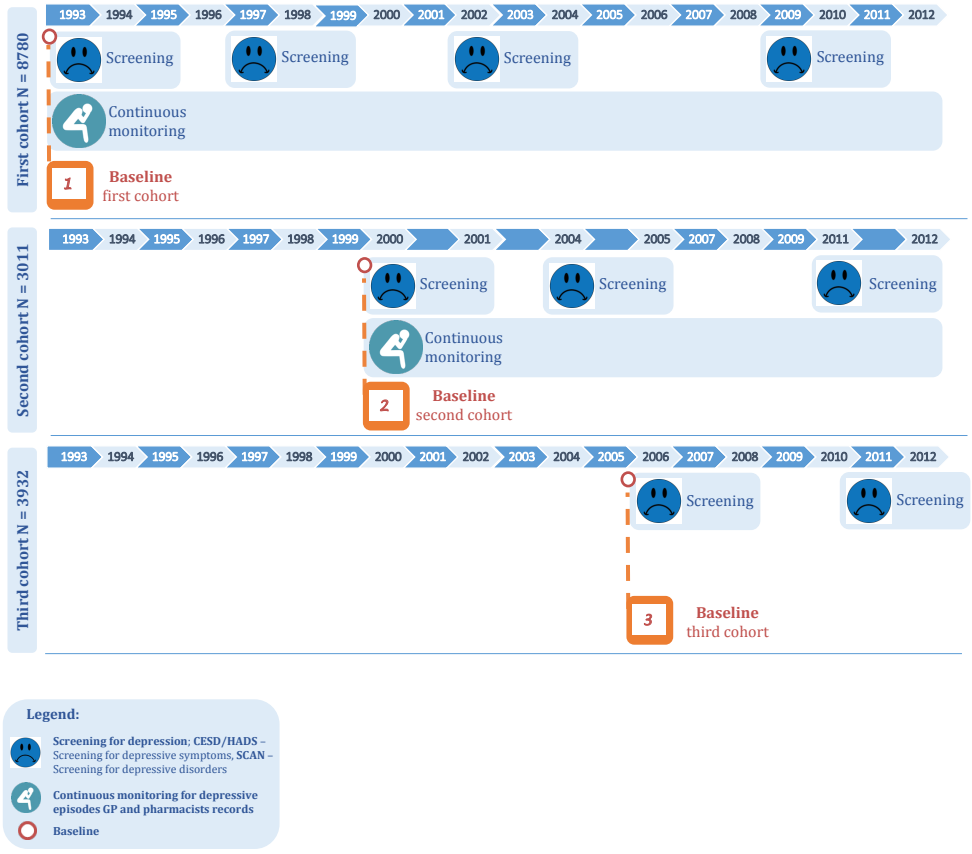
Unravelling the aetiology of depression could potentially answer some remaining questions about depression, and finally may explain why we consistently fail to develop effective management and treatment tools for depression. Therefore, this thesis aims to apply advanced epidemiological studies and to extend the existing knowledge on the aetiology of depression. Using population-

based data, the studies described in this thesis examine several risk factors and predictors that may enlighten the pathophysiological mechanism that underlie the development of depression. Specifically, **Chapter 2** of this thesis presents longitudinal studies that examine the impact of potential biomarkers, such as vitamin D (**Chapter 2.1**), and inflammatory markers (**Chapter 2.2**) on the occurrence of depression. **Chapter 3** of this thesis presents two studies which apply advanced genetic epidemiological methods to study the genetics of depression. **Chapter 4** focuses on the epigenetics of depression and presents the largest epigenome wide association population-based study so far. Moreover, we dedicated a chapter to the impact of physical health conditions such as myocardial infarct on depression (**Chapter 5.1**) as well as one to the physical consequences of depression such as cognitive decline (**Chapter 5.2**). Finally, **Chapter 6** provides a more general discussion of the main findings in this thesis and addresses several methodological considerations of the studies. Clinical implications of the results in this thesis, and future directions are also presented.

SETTING

The studies described in this thesis are large population-based studies of older adults screened for depressive symptoms and continuously followed for the occurrence of depressive disorders. Data presented in **Chapter 4** includes eleven large, population-based cohorts that contributed to an epigenome-wide meta-analysis and a replication analysis within the Cohorts for Heart and Aging Research in Genomic Epidemiology (CHARGE) Consortium. The remaining chapters present studies embedded in the Rotterdam Study, a population-based cohort that enrolled 14 926 adults aged 45 and older in Rotterdam, the Netherlands.[30] Participants underwent cycles of extensive home interviews and research examinations every 3 to 4 years. During the home interview participants were asked to self-report on the presence of depressive symptoms (using the Center for Epidemiologic Studies Depression Scale - CES-D). Those with clinically relevant depressive symptoms underwent a semi-structured psychiatric interview (Schedules for Clinical assessment in Neuropsychiatry - SCAN) during the research centre visit to diagnose depressive disorders. Moreover, participants were continuously monitored for various disorders, among which depression, via computerized linkage of data retrieved from pharmacists and general practitioner's reports. Detailed depression assessment was conducted since 1993 onwards.[31] A diagram of the depression assessment within the Rotterdam Study is presented in **Figure 1**.

Figure 1. Diagram of measurements of depression in the Rotterdam Study used in the current thesis.



REFERENCES

R1 1. Friedrich MJ. Depression Is the Leading Cause of Disability Around the World. *JAMA.* 2017;317(15):1517.

R2 2. Murray CJ, Lopez AD. Global mortality, disability, and the contribution of risk factors: Global Burden of Disease Study. *Lancet.* 1997;349(9063):1436-1442.

R3 3. Uher R, Payne JL, Pavlova B, Perlis RH. Major depressive disorder in DSM-5: implications for clinical practice and research of changes from DSM-IV. *Depress Anxiety.* 2014;31(6):459-471.

R4 4. Kohls E, Coppens E, Hug J, et al. Public attitudes toward depression and help-seeking: Impact of the OSPI-Europe depression awareness campaign in four European regions. *J Affect Disord.* 2017;217:252-259.

R5 5. Gronholm PC, Henderson C, Deb T, Thornicroft G. Interventions to reduce discrimination and stigma: the state of the art. *Soc Psychiatry Psychiatr Epidemiol.* 2017;52(3):249-258.

R6 6. Barchas JD, Brody BD. Perspectives on depression--past, present, future(a). *Ann N Y Acad Sci.* 2015;1345:1-15.

R7 7. Jackson SW. *Melancholia and depression: From hippocratic times to modern times.* Yale University Press; 1986.

R8 8. Kraepelin E. Manic depressive insanity and paranoia. *The Journal of Nervous and Mental Disease.* 1921;53(4):350.

R9 9. Lolas F. Bioethics and psychiatry: a challenging future. *World Psychiatry.* 2002;1(2):123-124.

R10 10. Mondimore FM. Kraepelin and manic-depressive insanity: an historical perspective. *Int Rev Psychiatry.* 2005;17(1):49-52.

R11 11. Ebert A, Bar KJ. Emil Kraepelin: A pioneer of scientific understanding of psychiatry and psychopharmacology. *Indian J Psychiatry.* 2010;52(2):191-192.

R12 12. Organization WH. *The ICD-10 classification of mental and behavioural disorders: diagnostic criteria for research.* Vol 2: World Health Organization; 1993.

R13 13. Paykel ES. Basic concepts of depression. *Dialogues Clin Neurosci.* 2008;10(3):279-289.

R14 14. Greenberg G. *Manufacturing Depression: The Secret History of a Modern Disease.* Simon & Schuster; 2010.

R15 15. Whitaker-Azmitia PM. The discovery of serotonin and its role in neuroscience. *Neuropsychopharmacology.* 1999;21(2 Suppl):2S-8S.

R16 16. Fryers T, Brugha T. Childhood determinants of adult psychiatric disorder. *Clin Pract Epidemiol Ment Health.* 2013;9:1-50.

R17 17. Everson SA, Maty SC, Lynch JW, Kaplan GA. Epidemiologic evidence for the relation between socioeconomic status and depression, obesity, and diabetes. *J Psychosom Res.* 2002;53(4):891-895.

R18 18. Kiecolt-Glaser JK, Newton TL. Marriage and health: his and hers. *Psychol Bull.* 2001;127(4):472-503.

R19 19. Bruce ML. Psychosocial risk factors for depressive disorders in late life. *Biological Psychiatry.* 2002;52(3):175-184.

R20 20. Fedak KM, Bernal A, Capshaw ZA, Gross S. Applying the Bradford Hill criteria in the 21st century: how data integration has changed causal inference in molecular epidemiology. *Emerg Themes Epidemiol.* 2015;12:14.

R21 21. Gururajan A, Clarke G, Dinan TG, Cryan JF. Molecular biomarkers of depression. *Neurosci Biobehav Rev.* 2016;64:101-133.

R22 22. Abbasi J. 23andMe, Big Data, and the Genetics of Depression. *JAMA.* 2017;317(1):14-16.

R23 23. Mechawar N, Savitz J. Neuropathology of mood disorders: do we see the stigmata of inflammation? *Transl Psychiatry.* 2016;6(11):e946.

R24 24. Gadad BS, Jha MK, Czysz A, et al. Peripheral biomarkers of major depression and antidepressant treatment response: Current knowledge and future outlooks. *J Affect Disord.* 2017.

R25 R26 R27 R28 R29 R30 R31 R32 R33 R34 R35 R36 R37 R38 R39

25. Mullins N, Lewis CM. Genetics of Depression: Progress at Last. *Curr Psychiatry Rep.* 2017;19(8):43.
26. Boyle EA, Li YI, Pritchard JK. An Expanded View of Complex Traits: From Polygenic to Omnigenic. *Cell.* 2017;169(7):1177-1186.
27. Pirooznia M, Wang T, Avramopoulos D, Potash JB, Zandi PP, Goes FS. High-throughput sequencing of the synaptome in major depressive disorder. *Mol Psychiatry.* 2016;21(5):650-655.
28. Nestler EJ. Epigenetic mechanisms of depression. *JAMA Psychiatry.* 2014;71(4):454-456.
29. Menke A, Binder EB. Epigenetic alterations in depression and antidepressant treatment. *Dialogues Clin Neurosci.* 2014;16(3):395-404.
30. Hofman A, Brusselle GG, Darwish Murad S, et al. The Rotterdam Study: 2016 objectives and design update. *Eur J Epidemiol.* 2015;30(8):661-708.
31. Luijendijk HJ, van den Berg JF, Dekker MJ, et al. Incidence and recurrence of late-life depression. *Arch Gen Psychiatry.* 2008;65(12):1394-1401.

R1 R2 R3 R4 R5 R6 R7 R8 R9 R10 R11 R12 R13 R14 R15 R16 R17 R18 R19 R20 R21 R22 R23 R24 R25 R26 R27 R28 R29 R30 R31 R32 R33 R34 R35 R36 R37 R38 R39

CHAPTER 2

Biomarkers for depression

CHAPTER 2.1

Vitamin D serum levels and depression in the elderly

Jovanova Olivera, Aarts Nikkie, Noordam Raymond, Carola Zillikens,
Hofman Albert, Tiemeier Henning

ACTA Psychiatrica Scandinavica. 2017 Mar; 135(3):185-194

ABSTRACT

Objective: The evidence for a prospective association of vitamin D deficiency with the occurrence of late-life depression is limited. We aimed to study the long-term association between vitamin D serum levels and depression in a large population-based study of older adults.

Method: We included 3 251 participants from the Rotterdam Study, aged 55 and older with 32 400 person-years follow-up for depression. Baseline 25-Hydroxivitamin D (25(OH)D) serum levels were analyzed continuously and categorically. Repeated depressive symptoms questionnaire assessments were used to assess the change of depressive symptoms. Semi-structured psychiatric interviews, and GP-records were used to assess incident major depressive disorder according to DSM-IV criteria.

Results: Low serum vitamin D levels were cross-sectionally associated with more depressive symptoms. However, low 25(OH)D serum levels were not prospectively associated with change of depressive symptoms (unstandardized beta $\beta = 0.02$, 95%CI = -0.23; $p = 0.26$) or incident MDD (hazard ratio *HR* = 0.95, 95%CI = 0.86; $p = 1.05$).

Conclusion: We observed a cross-sectional but no prospective association between serum vitamin D levels and depression. A cross-sectional association in the absence of the longitudinal association can mostly be attributed to reverse causality or residual confounding. Probably, vitamin D deficiency is not an independent risk factor for depression but co-occurs with late-life depression.

Significant outcomes: 1. Serum vitamin D deficiency and late-life depression are not prospectively associated. 2. Only in the cross-sectional analysis we observed that older persons with lower vitamin D serum levels were more likely to have more depressive symptoms. 3. This study provides further evidence that vitamin D deficiency accompanies late-life depression, but does not suggest a causal role of vitamin D deficiency in the development of depression.

Limitations: 1. This study reports analysis of single intra-individual assessment of vitamin D serum levels thus variability of vitamin D serum levels over time cannot be observed. 2. This observational study design limits the possibility to address reverse causality and draw conclusion about the directionality of the association between vitamin D serum levels and depression.

INTRODUCTION

Chronic/long-standing vitamin D deficiency is a common health risk factor that affects as many as one third to half of all elderly people world-wide.[1-3] The main sources of this neuro-steroid hormone are the synthesis in the skin in response to sunlight exposure and dietary intake. In the elderly, the skin capacity to synthesize vitamin D is reduced up to 25%.[4] Often the frequency of outdoor physical activity and the sunlight exposure are also low, making older people prone to vitamin D deficiency.[5]

Vitamin D deficiency and depression can coexist during late-life. It has been reported that one in every five geriatric patients with vitamin D deficiency suffer from depression.[6] A growing body of epidemiological evidence reports a relationship between low 25(OH)D serum levels and depression.[7-9] Several studies showed a consistent cross-sectional relationship between low serum levels of 25(OH)D and more depressive symptoms in older adults.[6,8,10] Yet, few epidemiological studies addressed the longitudinal relationship between vitamin D deficiency and late-life depression, and yielded inconclusive results, or failed to confirm the possible relationship.[7,8,10-13] Two recently published reviews concluded that the available prospective data is scarce, the analysis are not comprehensive, and the reported estimates were not precise.[8,14] In order to clarify the uncertainty about a prospective association between vitamin D and depression well-designed longitudinal studies are needed.[8,14]

Further, the few available longitudinal studies had limited precision for defining vitamin D deficiency and used different cut-offs. Second, their results were mostly based on self-reported depressive symptoms, thus the association with well-defined cases of major depression remains understudied in prospective research.[8,11] Third, three of the previous reports were focused on men and one of these was conducted in cardiovascular patients. These studies cannot necessarily be considered representative of the general population.[7,13,15]

Forth, potential confounding bias is a main challenge of all vitamin D studies.[8] Serum 25(OH)D levels depend on UVB-induced synthesis; thus sunlight exposure, season, and outside activity, are all relevant determinants of a person's serum 25(OH)D levels. Sunlight exposure is correlated with both vitamin D serum deficiency and depression and its confounding effect on this relationship was previously acknowledged.[8] Still, whether sunlight exposure and health-related problems influence the prospective association between 25(OH)D serum levels and depression is not well understood.[8]

Importantly, it is uncertain whether vitamin D deficiency is an independent risk factor for depression or a marker of poor health status that occurs as a consequence of a prior depression

or other chronic diseases.[8,10] Earlier studies indicate that vitamin D deficiency may contribute to the development of many systemic diseases like cardiovascular diseases, diabetes, and cancer; all highly prevalent in elderly population.[2,3,16,17]

Aims of the study

In conclusion, the cross-sectional association between vitamin D and depression is observed consistently. In contrast, the more recently published prospective studies reported inconsistent results and failed to establish a longitudinal association. Given the inconclusive evidence of a longitudinal association between vitamin D deficiency and depression, additional prospective studies are needed. In the current study we sought to clarify whether vitamin D serum levels are prospectively associated with depression, using data with a well-defined change of depressive symptoms as well as major depression during follow-up. The first aim was to replicate the cross-sectional relation between vitamin D and depressive symptoms. The second aim was to explore whether there is any long-term relationship of the vitamin D serum levels with the change of depressive symptoms or with incident major depression. We hypothesized that 25(OH)D serum levels are not only cross-sectionally associated with depression but also affect the risk of incident depression after carefully controlling for confounders.

METHODS

Study population

The present study was embedded in the Rotterdam Study, a population-based cohort designed to investigate diseases and their determinants among people aged 55 years and older. All residents of a district in Rotterdam, The Netherlands, were invited to participate.[18] Trained research assistants collected data on health, medication use, medical and family history, and lifestyle factors in extensive home interviews. From March 1997 till December 1999 (baseline of this study), 4 214 participants visited the research center for clinical examination and blood sampling. Blood serum sample of 25(OH)D was available for 3 828 participants. Out of these, 562 participants with an Mini-Mental State Examination (MMSE) score < 26 or missing, were excluded.[19] Another 15 participants, with no assessment of baseline depressive symptoms (CES-D - the Dutch version of the Centre for Epidemiologic Studies Depression scale) were excluded; leaving 3 251 participants in the study population. Women (57% versus 61%, $p = 0.03$), younger persons (mean age 72 versus 75 years, $p < 0.05$); and those with less depressive symptoms (mean CES-D score 4 versus 5.8, $p < 0.05$) were more likely to take part in our study.

The medical ethics committee approved the study (Wet Bevolkingsonderzoek-Population Study Act executed by the Ministry of Health, the Netherlands).[18] Written informed consent was obtained for all participants.

Vitamin D

Vitamin D was measured at baseline only, by assessing serum levels of 25(OH)D. This is the product of cutaneous synthesis from sun exposure and dietary sources, and after 25 hydroxylation in the liver; it represents the major storage form of vitamin D in the human body.[20] In 1997 - 1999, fasting blood samples were collected and centrifuged for 20 min. The serum was separated, dispensed and frozen within 3 hours at -80° C. Serum total 25(OH)D was quantitatively determined using Elecsys vitamin D total assay (COBAS, Roche Diagnostics GmbH, Germany). The electrochemiluminescence immunoassay is intended for use on Elecsys and cobas-e immunoassay analyzers. The test functional sensitivity was determined to be 10 nmol/L. Limit of quantitation was 22.5 nmol/L and intra-and inter-assay coefficients of variation were < 8% and < 11% for concentration between 7.5 and 175 nmol/L.[21] We analyzed different cut-off points of vitamin D, rather than using a single definition of deficiency to reduce the likelihood of chance findings.[8]

Depressive symptoms

Depressive symptoms were assessed with the Dutch version of the Centre for Epidemiologic Studies Depression scale (CES-D).[22] The CES-D scale was designed to assess presence and severity of self-reported depressive symptoms.[23] We asked participants 20 questions that correspond with criterion based-symptoms associated with depression, and participants could score from 0 up-to 60. The screening for depressive symptoms was performed at baseline 1997 - 1999. To assess for change of depressive symptoms the screening was reassessed twice; first during examination round 2002 until 2004 and second during examination round 2009 until 2011.[24]

Major depressive disorder

We continuously followed participants for the occurrence of incident major depression from baseline 1997 - 1999 until 1 January 2012. The participants were followed for on average 10.0 years (± 3.5 *SD*, between 1997 - 2012; for 32 400 person-years) for the occurrence of depression. If no depression was identified, participants were censored at the end of follow-up, if they moved or at death. Incident events of major depressive disorder (MDD) were identified from two sources of information, as reported in detail previously.[25] At first, all screen-positive participants identified by a CES-D score of 16 or above in each follow-up examination, were interviewed by a clinician (psychiatrist, psycho-geriatrician or clinical psychologist) with a semi-structured clinical interview

(Dutch version of the Schedules for Clinical Assessment in Neuropsychiatry - SCAN)[26] to diagnose depressive disorders. Major depressive disorders were classified according to the Statistical Manual of Mental Disorders, 4th revised edition (DSM-IV) criteria.

Second, information on occurrence of episodes of MDD were continuously collected from general practitioners medical records. The CES-D screening and SCAN interview provided information on depressive episodes that are present during a follow-up examination. The medical records data identified depressive episodes that occurred and remitted in the intervals between follow-up examinations. The Netherlands has a primary physician health-care system, thus all medical records (hospital discharge letters, specialists reports, and GP notes) could be extracted and copied by a research-assistants looking for potential depressive symptoms. These data were rated/categorized by two medical doctors.[25] Consensus decisions were made for disagreeing categorizations. Finally, incident episodes of MDD were defined as the first event that chronologically occurred in one of the two data sources described above.[25] In order to control for prevalent baseline depression (both clinical and subclinical) we adjusted the longitudinal analysis for baseline depressive symptoms.

Covariates

The following socio-demographic variables, were assessed during the baseline home interview and included in the analysis: age, gender, partner status, living independently or in a nursing home, and level of education. Partner status was classified as never married or divorced, married, and widowed. Education was classified as low, intermediate, or high. Alcohol consumption and smoking habits were both assessed at baseline and classified as never, past, or current smoker/consumer. Self-reported vitamin dietary supplementation was assessed at baseline.[27] Everyday functional competence was assessed by using the Activities of daily living scale ADL.[28] In order to assess cognitive performance the MMSE was measured.[19] Body mass index BMI (kg/m^2), systolic blood pressure (mmHg), creatinine and estimated glomerular filtration rate were assessed with standard medical and laboratory procedures. Presence of chronic conditions/diseases (stroke, diabetes mellitus, cardiovascular disease, liver conditions (based to elevated liver enzymes), and Parkinson's disease) was based on self-report, examination, medical record information, and drug utilization. A past diagnosis of cancer was based on self-report only. Latitude, season and sunlight exposure are widely accepted as determinants of serum vitamin D.[7,29] Therefore, season was categorized as blood intake in winter, spring, summer and autumn. Moreover, we used the data of the Dutch Royal Meteorological Center KNMI to calculate the individual hours of sunlight in the 10 weeks preceding the blood intake.[30] Our participants live in one area only thus we did not account for latitude.

DATA ANALYSIS

Serum 25(OH)D level (nmol/L) was analyzed continuously as a Z-transformed square root variable, in order to normalize the distribution and for better interpretation of results. Additionally, we analyzed 25(OH)D using two pre-defined cut-offs; a ≤ 37.5 nmol/L, a ≤ 50 nmol/L, and ≤ 75 nmol/L to define vitamin D serum level deficiency. Also vitamin D was categorized into quartiles, within our study sample to explore effects across the range of the continuum. These alternative categorizations used in previous research[8], facilitate comparison and tests consistency of results across cut-off.[8]

We studied the association between serum 25(OH)D levels and depression both cross-sectionally and longitudinally. First, we addressed the cross-sectional association between serum 25(OH)D and depressive symptoms (CES-D score) with linear regression. Second, we studied the longitudinal association between serum 25(OH)D and change of depressive symptoms (CES-D score) with linear regression and generalized estimating equation analysis (GEE). Moreover, we studied the longitudinal association between serum 25(OH)D and incident MDD by Cox proportional hazard survival analysis. The proportional hazards assumption was assessed by visual inspection of log-survival curves and by performing an interaction test with time. For all approaches we built two models: first, age and sex adjusted (in the longitudinal approach we additionally adjusted for baseline depressive symptoms) and second, a fully adjusted model. We included a covariate in the model if it changed the estimate of the main determinant by more than 10%, predicted depression ($p < 0.05$) or was an important priory confounder.[31] Sex differences in serum 25(OH) D levels and depression have been described.[6,32] Gender interaction-term and possibility of non-linear relationship were tested.

Multiple imputations were used in order to account for missing data on potential confounding variables (missing values: systolic blood pressure 0.4%, BMI 0.9%, ADL 0.1%, EGRF 0.2%, creatinine 0.2% and calcium serum levels 0.1%, education 1.4%, sunlight exposure 0.1%, chronic conditions 0.1%, and vitamin dietary supplements 6.7%). All analyses were rerun in the complete case and five imputed data sets. In this manuscript we present results from the imputed data. All statistical tests were two-sided and $p < 0.05$ was considered statistically significant. Analysis were performed using SPSS Statistics (version 21).

RESULTS

Baseline characteristics of the cohort are presented in Table 1. The mean age was 71.6 (± 6.6 *SD*) years and 57.4% of participants were women. The 3 251 participants had a mean 25(OH)D serum level of 49.68 nmol/l. In total, 1 843 (56.7%) individuals had a deficient 25(OH)D serum levels (≤ 50 nmol/L), and 1 408 (43.3%) individuals had a sufficient 25(OH)D serum levels (> 50 nmol/L).

Cross-sectional analysis

The cross-sectional analysis of the relationship between 25(OH)D level serum level and depressive symptoms are presented in Table 2. In both continuous and categorical analyses of vitamin D we found strong inverse relationship between 25(OH)D serum levels and depressive symptoms. After adjustment for age, sex, BMI, baseline CES-D score, ADL score, chronic conditions, sunlight, and other lifestyle factors, participants with a lower serum 25(OH)D levels had more depressive symptoms unstandardized beta (β) = -0.27, 95% CI = -0.51;-0.04, p = 0.023).

Additionally, we analyzed low 25(OH)D serum levels as a categorical exposure using two different cut-off points and quartiles. A serum 25(OH)D level ≤ 37.5 nmol/l was associated with depressive symptoms (β = 0.48, 95% CI = -0.01;0.95, p = 0.046) compared to serum 25(OH)D levels > 37.5 nmol/l. Similarly, when we analyzed the exposure using the cut-off ≤ 75 nmol/L serum 25(OH)D levels, these were associated with depressive symptoms (β = 0.61, 95% CI = -0.02;1.20, p = 0.045) compared to levels > 75 nmol/l. When analyzing 25(OH)D serum level quartiles, persons in each of the lower quartile groups had more depressive symptoms than those in the reference group (> 63.21 nmol/L)(see Table 2.).

Table 1. Baseline characteristics of the study population

	Total sample *N*=3251	Vitamin D deficiency *(<50 nmol/L)* *N*=1843	Vitamin D sufficiency *(>50 nmol/L)* *N*=1408
	Mean (*SD*)[a]		
Age (years)	71.6 (6.6)	73 (6.8)	69.7 (5.8)
Female gender, *N*(%)	1866 (57.4)	1206 (64.6)	660 (35.4)
Systolic blood pressure (mmHg)	143.2 (21)	144.7 (21.4)	141.1 (20.3)
Body mass index (kg/m2)	26.8 (4)	27.3 (4.3)	26.3 (3.4)
Mini mental state examination (score)	28.2 (1.3)	28.1 (1.3)	28.2 (1.2)
Smoking status			
Non smoker, ***N*(%)**	1032 (31.7)	653 (63.3)	379 (36.7)
Past smoker, ***N*(%)**	1627 (50.0)	830 (51.0)	797 (49.0)
Current smoker, ***N*(%)**	592 (18.2)	360 (60.8)	232 (39.2)
Alcohol consumption			
Non consumer, ***N*(%)**	319 (9.8)	233 (73.0)	86 (27.0)
Past consumer, ***N*(%)**	215 (6.6)	138 (64.2)	77 (35.8)
Current consumer, ***N*(%)**	2717 (83.6)	1472 (54.2)	1245 (45.8)
Activities of Daily Living (ADL score)	1.39 (0.5)	1.5 (0.6)	1.3 (0.4)
Creatinine serum levels (nmol/L)	80.5 (27.2)	78.4 (21.2)	83.2 (33.2)
Calcium serum levels (mmol/L)	2.4 (0.1)	2.4 (0.1)	2.4 (0.1)
Estimated glomerular filtration rate	74.6 (15.5)	74.4 (15.8)	74.8 (15.1)
Education			
Low education, ***N*(%)**	1788 (55.0)	1079 (60.3)	709 (39.7)
Middle education, ***N*(%)**	1055 (32.5)	550 (52.1)	505 (47.9)
High education, ***N*(%)**	363 (11.2)	179 (49.3)	184 (50.7)
Marital status			
Never married or divorced, ***N*(%)**	357 (11.0)	222 (62.2)	135 (37.8)
Married or living with partner, ***N*(%)**	2182 (67.1)	1120 (51.3)	1062 (48.7)
Widowed, ***N*(%)**	712 (21.9)	501 (70.4)	211 (29.6)
Season of blood intake			
Winter, ***N*(%)**	1089 (33.5)	548 (50.3)	541 (49.7)
Spring, ***N*(%)**	668 (20.5)	420 (62.9)	248 (37.1)
Summer, ***N*(%)**	1049 (32.3)	669 (63.8)	380 (36.2)
Autumn, ***N*(%)**	442 (13.6)	203 (45.9)	239 (54.1)
Sunlight exposure, (hours)[b]	305.6 (120.5)	292.5 (114.6)	322.8 (125.8)
Chronic conditions, *N*(%)[c]	1049 (32.3)	671 (64.0)	378 (36.0)
Cancer status, *N*(%)	235 (7.2)	136 (57.9)	99 (42.1)
Vitamin dietary supplements, *N*(%)	303 (9.3)	135 (44.6)	168 (55.4)
Baseline CES-D score[d]			
CES-D < 16, ***N*(%)**	3047 (93.7)	1714 (56.3)	1333 (43.7)
CES-D ≥ 16, ***N*(%)**	204 (6.3)	129 (63.2)	75 (36.8)

Abbreviations: SD, standard deviation; CES-D, Dutch version of the Centre for Epidemiologic Studies Depression scale.

[a] Unless otherwise is indicated;

[b] Sunlight hours during 10 week period preceding the blood drawing;

[d] Chronic conditions: History of stroke, history of diabetes mellitus, history of Parkinson disease, history of liver conditions and history of cardiovascular disease.[d] Those who score on CES-D ≥ 16 have clinically relevant depressive symptoms

Table 2. Cross-sectional association between serum vitamin D levels and depressive symptoms (CES-D) with linear regression (*N* = 3 251)

Vitamin D serum levels		Depressive symptoms (CES-D score)			
	N	**Model 1**		**Model 2**	
		β (95 % CI)	*p*	β (95 % CI)	*p*
Continuously					
Vitamin D √*SD* nmol/L[a]	3251	-0.54 (-0.77;-0.30)	<0.001	-0.27 (-0.51;-0.04)	0.023
Cut-off					
<37.5 nmol/L	1226	0.95 (0.48;1.43)	<0.001	0.48 (0.01;0.95)	0.046
>37.5 nmol/L	2025	**Reference**		**Reference**	
<50 nmol/L	1843	0.62 (0.17;1.08)	0.008	0.28 (-0.18;0.73)	0.23
>50 nmol/L	1408	**Reference**		**Reference**	
<75 nmol/L	2748	0.85 (0.24;1.46)	0.006	0.61 (0.02;1.20)	0.045
>75 nmol/L	503	**Reference**		**Reference**	
Quartiles					
<28.57 nmol/L	732	1.48 (0.82;2.15)	<0.001	0.74 (0.07;1.41)	0.030
28.58-43.81 nmol/L	819	1.11 (0.49;1.72)	<0.001	0.83 (0.22;1.43)	0.008
43.82-63.21 nmol/L	848	0.83 (0.23;1.43)	0.007	0.75 (0.19;1.31)	0.012
>63.21 nmol/L	852	**Reference**		**Reference**	

Abbreviations: *SD*, standard deviation; CES-D, Dutch version of the Centre for Epidemiologic Studies Depression scale; β, unstandardized beta; CI, confidence interval.
Interaction term vitamin D serum levels*gender was tested and showed no statistical significance ($p = 0.936$).
[a] Standard Deviation of the square root of the Vitamin D serum levels (nmol/L)
Model 1. Adjusted for gender and age
Model 2. Additionally adjusted for: body mass index, systolic blood pressure, chronic conditions, cancer status, smoking habits, alcohol consumption, marital status, level of education, 10 week sunlight exposure prior to the blood intake, calcium serum levels, and activity of daily leaving score.

Longitudinal analysis

Next, we studied the longitudinal relationship between 25(OH)D serum levels and change of depressive symptoms as well as incident major depressive disorder.

First, (see Table 3.) in contrast to the cross-sectional analysis, low levels of 25(OH)D did not predict higher depressive symptoms at the first follow-up (β = 0.01, 95%CI = -0.28;0.29, $p = 0.95$) or second follow-up assessment (β = 0.05, 95%CI = -0.31;0.40, $p = 0.80$). Moreover, we did not found an association between 25(OH)D serum levels and change of depressive symptoms in the combined analysis of the two assessment waves (β = 0.02, 95%CI = -0.23;0.26, $p = 0.89$).

Table 3. Longitudinal association between serum vitamin D levels and depressive symptoms assessed with linear regression and generalized estimated equations (*N* = 3 251)

Vitamin D serum levels	Depressive symptoms (CES-D score)								
	N	First assessment		*N*	Second assessment		*N*	Combined assessments	
		β (95%CI)	*p*		β (95%CI)	*p*		β (95%CI)	*p*
Continuously									
Vitamin D√*SD* nmol/L[a]	2595	0.01 (-0.28;0.29)	0.95	1702	0.05 (-0.31;0.40)	0.80	4297	0.02 (-0.23;0.26)	0.89
Cut-off									
<37.5 nmol/L	904	-0.39 (-0.96;0.18)	0.18	508	-0.56 (-1.28;0.17)	0.13	1412	-0.44 (-0.97;0.07)	0.10
>37.5 nmol/L	1691	**Reference**		1194	**Reference**		2885	**Reference**	
<50 nmol/L	1408	-0.16 (-0.73;0.41)	0.57	843	-0.08 (-0.76;0.59)	0.81	2251	-0.14 (-0.62;0.33)	0.57
>50 nmol/L	1187	**Reference**		859	**Reference**		2046	**Reference**	
<75 nmol/L	2156	0.36 (-0.00;0.72)	0.32	1366	0.11 (-0.32;0.53)	0.80	3522	0.25 (-0.29;0.79)	0.37
>75 nmol/L	439	**Reference**		336	**Reference**		775	**Reference**	
Quartiles									
<28,57 nmol/L	517	-0.25 (-1.06;0.57)	0.55	272	0.08 (-0.96;1.12)	0.88	789	-0.09 (-0.81;0.62)	0.80
28,58-43,81 nmol/L	646	-0.06 (-0.43;0.31)	0.86	412	-0.54 (-0.99;-0.08)	0.24	1058	-0.25 (-0.89;0.40)	0.45
43,82-63,21 nmol/L	693	0.25 (-0.11;0.61)	0.48	470	-0.02 (-0.44;0.41)	0.97	1163	0.22 (-0.38;0.81)	0.47
>63.21 nmol/L	739	**Reference**		548	**Reference**		1287	**Reference**	

Abbreviations: *SD*, standard deviation; CES-D, Dutch version of the Centre for Epidemiologic Studies Depression scale; β, unstandardized beta; CI, confidence interval.

[a] Standard Deviation of the square root of the Vitamin D serum levels (nmol/L)

The results present the fully adjusted model (adjusted or sex, age, body mass index, chronic conditions, smoking status, alcohol consumption, and activity of daily living score).

Second, we assessed the long-term relationship of 25(OH)D serum level and incident MDD. (see Table 4.) 32 400 person-years (mean 10.0 ± 3.5 *SD*, interquartile range 4.1 - 13.8), during which 150 incident MDD occurred. The serum level of 25(OH)D did not predict long-term risk of incident MDD (hazard ratio (*HR*) = 0.95, 95%CI = 0.86;1.05, *p* = 0.61).

Table 4. Longitudinal analysis of the association between serum vitamin D levels and incident major depressive disorder with Cox regression analysis (*N* = 2 466)

Vitamin D serum levels	Major depressieve disorder				
	events	**Model 1**		**Model 2**	
		HR (95%CI)	*p*	*HR* (95%CI)	*p*
Continuously					
Vitamin D √SD nmol/L[a]	150	0.98 (0.81 to 1.18)	0.79	0.95 (0.86 to 1.05)	0.61
Cut-off					
<37.5 nmol/L	70	0.96 (0.68 to 1.36)	0.84	1.04 (0.87 to 1.25)	0.81
>37.5 nmol/L	80	**Reference**		**Reference**	
<50 nmol/L	92	0.82 (0.58 to 1.17)	0.27	0.84 (0.70 to 1.01)	0.34
>50 nmol/L	58	**Reference**		**Reference**	
<75 nmol/L	137	1.32 (0.74 to 2.36)	0.35	1.28 (0.95 to 1.73)	0.41
>75 nmol/L	13	**Reference**		**Reference**	
Quartiles					
<28,57 nmol/L	38	0.77 (0.52 to 1.14)	0.19	0.82 (0.67 to 1.02)	0.36
28,58-43,81 nmol/L	45	1.13 (0.79 to 1.61)	0.49	1.13 (0.79 to 1.61)	0.51
43,82-63,21 nmol/L	43	1.36 (0.95 to 1.95)	0.10	1.24 (0.87 to 1.78)	0.24
>63.21 nmol/L	24	**Reference**		**Reference**	

Abbreviations: *SD*, standard deviation; *HR*, hazard ratio; CI, confidence interval.
Interaction term Vitamin D serum levels*gender was tested and showed no statistical significance (p = 0.160)
[a]Standard Deviation of the square root of the Vitamin D serum levels (nmol/L)
Model 1. Adjusted for gender, age and baseline depressive symptoms
Model 2. Additionally adjusted for: body mass index, alcohol consumption, smoking status, marital status and activity of daily living score.

Sensitivity analysis

Several sensitivity analyses were performed. First, we repeated the analysis of the association of 25(OH)D serum levels and incident MDD restricting the analysis to 2 and 5 years follow-up (Supplementary material 1). Again we found no indication of a prospective association between vitamin D serum levels and incident depression.

Second, sex interaction-term with 25(OH)D serum levels was tested and showed no statistical significance (*p* = 0.94). A quadratic terms of serum 25(OH)D was examined to test for non-linear relationship, no evidence for non-linearity was observed.

DISCUSSION

In this large population-based cohort of older people, we observed a cross-sectional association of both continuously modeled vitamin D serum levels and vitamin D deficiency with depressive symptoms, consistent with previous studies.[7,8] However low levels of serum vitamin D were not prospectively associated with either change of depressive symptoms or incident MDD. Further, analyzing vitamin D as a continuous measure did not reveal any prospective association between vitamin D serum levels and depression.

Several epidemiological studies explored the cross-sectional relation between vitamin D serum levels and depression.[6,8] In line with previous investigations[6,7,12,33], we found a consistent cross-sectional relation regardless of whether we analyzed vitamin D serum levels categorically or continuously. Older adults with vitamin D deficiency are clearly more likely to have depressive symptoms.[8] Importantly, the presence of a cross-sectional relationship between vitamin D serum levels and depression is not discredited by the absence of long-term risk between them. Yet, our findings suggest that if there is a true relationship between low vitamin D serum levels and depression this would reflect on a short-term rather than a long-term effect on the development of depression.

Our study extends the results of the few earlier studies exploring the prospective relationship with conflicting results.[10,12,34,35] The three studies showing a prospective association between vitamin D and depression[13,34,36] were based on selected study populations, i.e. cardiovascular patients or men only. In contrast to these findings, Toffanello, et al. and Chan, et al. reported no prospective association in men and during relatively short period of 4 years follow-up, respectively.[10,12] Our results were consistent, none of the analysis provided any statistical evidence for longitudinal relationship between vitamin D serum levels and depression, regardless of vitamin D serum level cut-offs or severity of depression (depressive symptoms and incident depression). Unlike previous studies, we assessed depression using different information sources, and did not rely on self-report only. Depression was continuously monitored in GP-records over a mean time period of 10 years. Combining multiple sources of depression reduces ascertainment bias often seen in other studies.[8]

The observed effect estimates present depression rates over 10 years follow-up with respect to the vitamin D serum levels measured at baseline. Vitamin D serum levels are highly variable, depend on sunlight exposure and diet that changes over time.[3] Our report is limited by the single intra-individual assessment of vitamin D serum levels. The time-dependent variation of vitamin D serum levels may explain why we did not capture an association with depression during follow-up. Given the within-person variability of vitamin D serum levels, the effect of lower vitamin D

serum level is probably diluted when the follow-up period is longer than 5 years.[37] Thus, we tested different follow-up times of shorter duration. Yet, the findings remained negative. Also, studies with other outcomes showed that the variability of vitamin D does not preclude the observation of long-term effects. An independent association of vitamin D deficiency with both the development of cardiovascular disease and mortality has been reported.[16,38]

Many authors have speculated that vitamin D activates vitamin D receptors in brain regions related to depression, stimulates neurotrophic release, and protects the brain by buffering antioxidant and anti-inflammatory defense against vascular injury.[3,29] Current pathophysiological theories postulate an etiological link between vitamin D and depression. However, whether the cross-sectional findings reflect such a causal mechanism is questionable. Cherniack, et al. suggested that cross-sectional studies may miss sensitive period underlying any plausible association because mental illness typically take many months or years to develop.[3]

We speculate that, if there is any relationship, then vitamin D deficiency rather has a short-term than a long-term effect on the development of depression. However, it is very probable that there is no causal association between vitamin D deficiency and depression. There are several alternative explanations why we observed a cross-sectional association between vitamin D deficiency and depression, in the absence of a longitudinal association, in particular residual confounding and reverse causality.

The association between vitamin D and depression is complex and likely to be subject of residual confounding. This study was embedded in a population-based study, allowing us to use prospective measures of outcome and assess a large number of confounding factors. Yet, vitamin D deficiency and depression coexist during late-life[33] and it is possible that the cross-sectional association is due to the higher prevalence of both conditions in the elderly. Chronic diseases are common among elderly and convey a high risk of depression. These largely affect the mental health by lowering the quality of life, restricting activities, social isolation, and interfering with the quality of sleep but also via biological mechanisms such as immune and vascular factors.[39] Hence, we cannot reject the possibility of residual confounding that might bias our results, e.g. by physical activity or unmeasured health problems. Indeed Anglin, et al. pointed out that the risk of bias due to unmeasured confounding remains high.[8] Arguably, vitamin D deficiency is partly a marker of chronic nonspecific disease rather than a determinant or a direct contributor to the pathogenesis of depression. Sunlight exposure was already acknowledged as an important factor that can confound the relationship between vitamin D and depression.[8] Assessment of sunlight exposure prior to the vitamin D serum level measurement remains a challenge in previous studies.[7,8] We have tested for confounding by both sunlight exposure 10 week preceding the blood intake and season at the time of blood intake, but both measures accounted for a very

small part of cross-sectional association. Further, sunlight exposure did not affect the longitudinal association. We did not account for parathyroid hormone (PTH) which has been acknowledged as a covariate and an intermediate factor with a role in the association between vitamin D and depression.[34] However, May et al. showed that PTH cannot explain the relation between vitamin D and depression.[34]

Other limitation is the possibility of reverse causality. Depression may precede the vitamin D deficiency and reverse causality may explain the observed cross-sectional association between depression and vitamin D deficiency. In particular, depressed individuals don not easily go outside thus are typically less exposed to sunlight. In this longitudinal study we did not repeatedly assess vitamin D and we cannot demonstrate the directionality of the association or rule out that depressive symptoms occur prior to any vitamin D deficiency.

Also, we should point out that even though many elderly persons had low levels of vitamin D, the analysis of more severe vitamin D insufficiency (≤ 37.5 nmol/L) may not have had optimal power to detect an increased risk of incident depression (in contrast to that of continuous depressive symptoms) as new onset MDD is not very common in elderly populations.

This study provides further evidence for the cross-sectional association between serum vitamin D levels and depression. Clearly, older persons with lower vitamin D serum levels are more likely to have more depressive symptoms. However, we observed that vitamin D deficiency and late-life depression are not prospectively associated. In conclusion, vitamin D deficiency co-occurs with late-life depression, but has no causal role in the development of depression. Although, this is an observational study and defining causal mechanisms is not possible, the co-occurrence of vitamin D deficiency and late-life depression is a fact that may have clinical implication. Screening for vitamin D deficiency in elderly with chronic conditions such as depression remains important even if the conditions are not causally related. Vitamin D deficiency is associated with bone health, osteoporosis, and increased risk of falls and subsequent fractures.[40] Detection of vitamin D serum deficiency is easy and treatable. Even if treatment of vitamin D deficiency has no effect on depression[41], it may prevent other chronic conditions.

REFERENCES

1. Gronli O, Kvamme JM, Jorde R, Wynn R. Vitamin D deficiency is common in psychogeriatric patients, independent of diagnosis. *BMC Psychiatry.* 2014;14:134.
2. Holick MF. High prevalence of vitamin D inadequacy and implications for health. *Mayo Clin Proc.* 2006;81(3):353-373.
3. Cherniack EP, Troen BR, Florez HJ, Roos BA, Levis S. Some New Food for Thought: The Role of Vitamin D in the Mental Health of Older Adults. *Curr Psychiat Rep.* 2009;11(1):12-19.
4. Need AG, Morris HA, Horowitz M, Nordin C. Effects of skin thickness, age, body fat, and sunlight on serum 25-hydroxyvitamin D. *Am J Clin Nutr.* 1993;58(6):882-885.
5. Scragg R, Camargo CA, Jr. Frequency of leisure-time physical activity and serum 25-hydroxyvitamin D levels in the US population: results from the Third National Health and Nutrition Examination Survey. *Am J Epidemiol.* 2008;168(6):577-586; discussion 587-591.
6. Lapid MI, Cha SS, Takahashi PY. Vitamin D and depression in geriatric primary care patients. *Clin Interv Aging.* 2013;8:509-514.
7. Ju SY, Lee YJ, Jeong SN. Serum 25-hydroxyvitamin D levels and the risk of depression: a systematic review and meta-analysis. *J Nutr Health Aging.* 2013;17(5):447-455.
8. Anglin RE, Samaan Z, Walter SD, McDonald SD. Vitamin D deficiency and depression in adults: systematic review and meta-analysis. *Br J Psychiatry.* 2013;202:100-107.
9. Okereke OI, Singh A. The role of vitamin D in the prevention of late-life depression. *J Affect Disord.* 2016;198:1-14.
10. Chan R, Chan D, Woo J, et al. Association between serum 25-hydroxyvitamin D and psychological health in older Chinese men in a cohort study. *J Affect Disord.* 2011;130(1-2):251-259.
11. Williams JA, Sink KM, Tooze JA, et al. Low 25-hydroxyvitamin D concentrations predict incident depression in well-functioning older adults: the health, aging, and body composition study. *J Gerontol A Biol Sci Med Sci.* 2015;70(6):757-763.
12. Toffanello ED, Sergi G, Veronese N, et al. Serum 25-hydroxyvitamin d and the onset of late-life depressive mood in older men and women: the Pro.V.A. study. *J Gerontol A Biol Sci Med Sci.* 2014;69(12):1554-1561.
13. Almeida OP, Hankey GJ, Yeap BB, Golledge J, Flicker L. Vitamin D concentration and its association with past, current and future depression in older men: The Health In Men Study. *Maturitas.* 2015;81(1):36-41.
14. Li G, Mbuagbaw L, Samaan Z, et al. Efficacy of vitamin D supplementation in depression in adults: a systematic review. *J Clin Endocrinol Metab.* 2014;99(3):757-767.
15. Chan R, Chan D, Woo J, et al. Association between serum 25-hydroxyvitamin D and psychological health in older Chinese men in a cohort study. *Journal of Affective Disorders.* 2011;130(1–2):251-259.
16. Wang TJ, Pencina MJ, Booth SL, et al. Vitamin D deficiency and risk of cardiovascular disease. *Circulation.* 2008;117(4):503-511.
17. Palomer X, Gonzalez-Clemente JM, Blanco-Vaca F, Mauricio D. Role of vitamin D in the pathogenesis of type 2 diabetes mellitus. *Diabetes Obes Metab.* 2008;10(3):185-197.
18. Hofman A, Darwish Murad S, van Duijn CM, et al. The Rotterdam Study: 2014 objectives and design update. *Eur J Epidemiol.* 2013;28(11):889-926.
19. Folstein MF, Folstein SE, McHugh PR. "Mini-mental state". A practical method for grading the cognitive state of patients for the clinician. *J Psychiatr Res.* 1975;12(3):189-198.
20. Wallace AM, Gibson S, de la Hunty A, Lamberg-Allardt C, Ashwell M. Measurement of 25-hydroxyvitamin D in the clinical laboratory: current procedures, performance characteristics and limitations. *Steroids.* 2010;75(7):477-488.

21. Wagner D, Hanwell HE, Vieth R. An evaluation of automated methods for measurement of serum 25-hydroxyvitamin D. *Clin Biochem.* 2009;42(15):1549-1556.
22. Breslau N. Depressive symptoms, major depression, and generalized anxiety: a comparison of self-reports on CES-D and results from diagnostic interviews. *Psychiatry Res.* 1985;15(3):219-229.
23. Hales DP, Dishman RK, Motl RW, Addy CL, Pfeiffer KA, Pate RR. Factorial validity and invariance of the center for epidemiologic studies depression (CES-D) scale in a sample of black and white adolescent girls. *Ethn Dis.* 2006;16(1):1-8.
24. Beekman AT, Deeg DJ, Van Limbeek J, Braam AW, De Vries MZ, Van Tilburg W. Criterion validity of the Center for Epidemiologic Studies Depression scale (CES-D): results from a community-based sample of older subjects in The Netherlands. *Psychol Med.* 1997;27(1):231-235.
25. Luijendijk HJ, van den Berg JF, Dekker MJ, et al. Incidence and recurrence of late-life depression. *Arch Gen Psychiatry.* 2008;65(12):1394-1401.
26. Bjelland I, Dahl AA, Haug TT, Neckelmann D. The validity of the Hospital Anxiety and Depression Scale. An updated literature review. *J Psychosom Res.* 2002;52(2):69-77.
27. Klipstein-Grobusch K, Geleijnse JM, den Breeijen JH, et al. Dietary antioxidants and risk of myocardial infarction in the elderly: the Rotterdam Study. *Am J Clin Nutr.* 1999;69(2):261-266.
28. Lawton MP, Brody EM. Assessment of older people: self-maintaining and instrumental activities of daily living. *Gerontologist.* 1969;9(3):179-186.
29. Kulie T, Groff A, Redmer J, Hounshell J, Schrager S. Vitamin D: an evidence-based review. *J Am Board Fam Med.* 2009;22(6):698-706.
30. Institut KNM. Koninklijk Nederlands Meteorologisch Institut Ministerie van Infrastructuur en Milieu. [Klimaatdesk]. 2015.
31. Hernan MA, Hernandez-Diaz S, Werler MM, Mitchell AA. Causal knowledge as a prerequisite for confounding evaluation: An application to birth defects epidemiology. *Am J Epidemiol.* 2002;155(2):176-184.
32. Piccinelli M, Wilkinson G. Gender differences in depression. Critical review. *Br J Psychiatry.* 2000;177:486-492.
33. Bertone-Johnson ER. Vitamin D and the occurrence of depression: causal association or circumstantial evidence? *Nutr Rev.* 2009;67(8):481-492.
34. May HT, Bair TL, Lappe DL, et al. Association of vitamin D levels with incident depression among a general cardiovascular population. *Am Heart J.* 2010;159(6):1037-1043.
35. Milaneschi Y, Shardell M, Corsi AM, et al. Serum 25-hydroxyvitamin D and depressive symptoms in older women and men. *J Clin Endocrinol Metab.* 2010;95(7):3225-3233.
36. Milaneschi Y, Hoogendijk W, Lips P, et al. The association between low vitamin D and depressive disorders. *Mol Psychiatry.* 2014;19(4):444-451.
37. Clarke R, Shipley M, Lewington S, et al. Underestimation of risk associations due to regression dilution in long-term follow-up of prospective studies. *Am J Epidemiol.* 1999;150(4):341-353.
38. Zittermann A, Iodice S, Pilz S, Grant WB, Bagnardi V, Gandini S. Vitamin D deficiency and mortality risk in the general population: a meta-analysis of prospective cohort studies. *Am J Clin Nutr.* 2012;95(1):91-100.
39. Katon W, Lin EH, Kroenke K. The association of depression and anxiety with medical symptom burden in patients with chronic medical illness. *Gen Hosp Psychiatry.* 2007;29(2):147-155.
40. Holick MF. Vitamin D deficiency. *N Engl J Med.* 2007;357(3):266-281.

CHAPTER 2.2

Inflammatory markers and depression in the elderly

Zalli Argita, Jovanova Olivera, Hoogendijk Witte, Tiemeier Henning, Carvalho Livia

Psychopharmacology. 2016 May; 233(9):1669-78.

ABSTRACT

Rationale: Evidence suggests that depression is cross-sectionally and longitudinally associated with activation of inflammatory response system. Few studies, however, have investigated the longitudinal relationship between raised inflammatory biomarkers and persistence of depressive symptoms. We examined the temporal relationship between serum levels of inflammatory biomarkers and persistence of depressive symptoms among older participants.

Methods: Center for Epidemiologic Studies Depression Scale (CES-D) was used to assess depressive symptoms at baseline and at 5 year follow up in 656 participants (233 men, 423 women) aged > 60 years of the Rotterdam Study. Markers of inflammation interleukin (IL)-6, alpha-1-antichymotrypsin (ACT) and C reactive protein (CRP) were assessed at baseline, and all participants taking antidepressant medications were excluded from the analysis.

Results: No cross-sectional association was found between IL-6, ACT and CRP with depressive symptoms at baseline. However, higher levels of IL-6 and CRP predicted depressive symptoms at 5 year follow-up. Adjustment for confounding variables had no impact on the observed associations. Similarly, a positive association was found between baseline levels of IL-6 (odds ratio [*OR*], 2.44; $p = 0.030$) and CRP (*OR*, 1.81; $p = 0.052$) and persistence of depressive symptoms over 5 years.

Conclusion: Our data suggest that dysregulation of the inflammatory response system is associated with a more severe form of depression more likely to re-occur.

INTRODUCTION

There is evidence implicating inflammation as a potential etiologic factor for mood disorders. Meta-analysis studies have reported higher levels of inflammatory cytokines (such as Interleukin-6, (IL)-6) and acute phase proteins (such as C-reactive protein; (CRP)) in the peripheral blood and cerebrospinal fluid of patients with major depression.[1,2] In addition, inflammation and depressive symptoms seem to be associated in large epidemiological cross-sectional studies.[3-5]

Further evidence stems from prospective studies showing that acute or chronic administration of cytokines leads to development of depressive symptoms. Chronic administration of the pro-inflammatory cytokine IFN-α for treatment of hepatitis C induced clinically significant depression in 30 - 50% of persons with no psychiatric disorders previous to interferon-alpha (IFN-α) treatment.[6,7] Additional evidence to support an association between elevated immune-inflammatory cytokines and depressive-like behavioural systems is provided in animal models. Elevated immune-inflammatory cytokines induce and exacerbate depressive-like symptoms, whereas tumour necrosis factor-alpha (TNF)-α and IL-6 receptor knockout mice show reduced behavioural indices of depression.[8-13] In humans, fewer large epidemiological studies points towards a possible association between activation of inflammatory system and future depressive symptoms[5,14,15], although controversy still exists.[16]

A possible explanation for this controversy is the hypothesis that activation of the inflammatory system distinguishes a particular subset of patients i.e. those who have a more severe form of depression. In agreement, an activation of the inflammatory system was particularly observed in major depressive patients who are older[17], have recurrent episodes[17,18]; have comorbid depression with other mental and physical illnesses[19-21]; present earlier onset of the disorder[17] and are resistant to antidepressant treatment.[22,23] Consistent with the hypothesis that inflammation is present in a particular subgroup of depressed patients, the anti-inflammatory drug infliximab showed antidepressant properties only in treatment resistant depressed patients who have high levels of the inflammatory marker C-reactive protein.[24]

In this study, we take a longitudinal approach to examine the temporal relationship between the inflammatory biomarkers IL-6, alpha-1-antichymotrypsin (ACT), CRP at baseline and incident depressive symptoms 5 years later or persistent depressive symptoms over 5 years. Participants taking antidepressant medications were excluded from the analysis due to the small population size.

METHODS

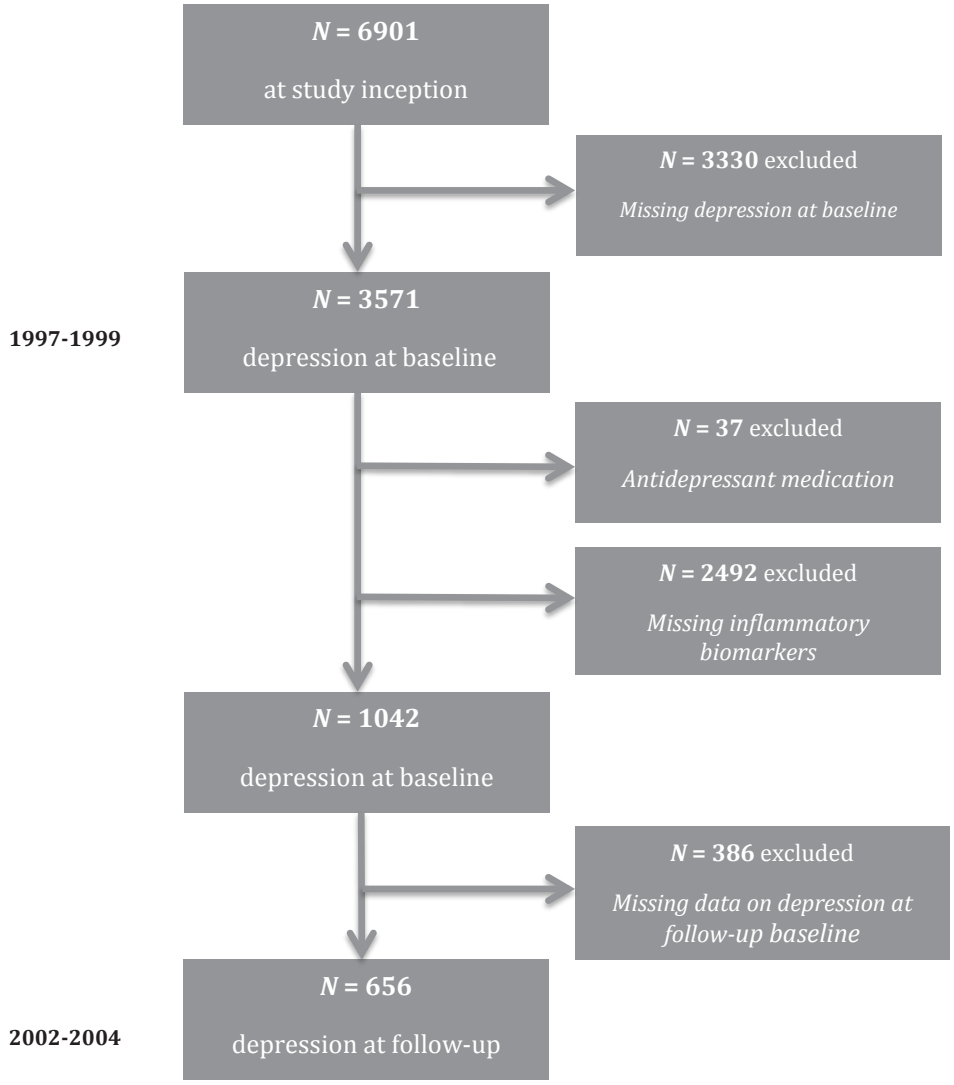
Study population

The present study is embedded within the Rotterdam Study, a population-based cohort study in which all inhabitants age 55 and over living in a defined geographic area of Rotterdam have been invited to participate.[25] The Medical Ethics Committee approved the study according to the Wet Bevolkingsonderzoek: ERGO (Population Study Act: Rotterdam Study) executed by the Ministry of Health, Welfare and Sports of the Netherlands. Written informed consent was obtained from all participants. During the third survey (1997 – 1999), participants were assessed for depressive symptoms. For the present study, of the 3 571 persons who were screened for depressive symptoms, only 1 211 had available information on serum level of inflammatory biomarkers. Moreover, thirty-seven participants were excluded for being on antidepressants, leaving 1 175 participants eligible for analysis. In the present analysis, we studied the association of inflammatory proteins and depressive symptoms in 656 participants free of antidepressant medications. Out of these participants, 102 were on anti-inflammatory medications. In this study, we do not have repeated measures of inflammatory cytokines (IL-6, ACT, and CRP), and therefore could not analyse whether long-term exposure to high levels of inflammatory markers are associated with depression. Sequential exclusions occurred according to the flow chart represented in Figure 1.

Depression assessment

Depressive symptoms were assessed through participant's completion of the Dutch version of the original Center for Epidemiological Studies Depression scale (CES-D) during a home interview.[26,27] The CES-D is a self-reporting 20-item measure of depressive symptoms scored on a scale from 0 to 3. For the analysis of persistent depressive symptoms, we used a score of 16 as a cut-off, to indicate clinically significant depressive symptoms in each wave.[26,28] Persistent depressive symptoms were defined by having CES-D scores $\geqq$ 16 in 1997 - 1999 (baseline) and when re-assessed in 2002 - 2004 (follow-up).

Figure 1. Flow chart for sample selection in the study



Blood specimens

At baseline, a venepuncture was performed by application of minimal stasis with a 21-gauge butterfly needle with tube (Surflo winged infusion set, Terumo). Fasting blood was collected in the morning and all tubes were put on ice directly and centrifuged at 2000×g for 10 minutes. Plasma was separated and dispensed into two 1.5-ml aliquots and then frozen within 3 hours at - 80°C. Both ACT and CRP were assessed by means of a nephelometric method (BN 100, Dade Behring,

Marburg, Germany). The IL-6 concentrations were determined with quantitative enzyme-linked immunosorbent assay with a test kit from R&D systems (Minneapolis, MN). The intra-assay and interassay coefficients for all measurements were < 5% and < 8%, respectively. High-sensitivity CRP was measured in a serum, which was stored at - 20°C until performance of the CRP measurements, using a Rate Near Infrared Particle Immunoassay (IMMAGE, Immunochemistry System, Beckman Coulter, San Diego, CA, detection limit 0.2 mg/L, coefficient of variation (CV) 3.2%). In this matter a fully automated Hitachi 747 system (Hitachi, Tokyo, Japan, detection limit 1 mg/L, CV<5%) was used. IL-6 plasma levels were determined using a quantitative enzyme-linked immunosorbent assay (ELISA) technique (Quantikine HS IL-6 kit, R&D Systems, Oxon, UK, detection limit 0.094 pg/mL, CV 8.7%) and ACT plasma levels using kinetic nephelometry (Behring Nephelometer BN200, Marburg, Germany, detection limit 1.5 mg/dL, CV 2.8%).

Other measurements

The following variables were considered as possible confounding variables: age, sex, education (low, middle, high), physical illness (including prevalent stroke, cardiovascular disease and diabetes), cognitive function (as measured by the mini mental-state examination), smoking and body mass index (BMI). A history of stroke was obtained through direct questioning and computerized linkage with general practitioner medical records.[29] Smoking was coded as number of cigarettes currently smoked per day and in categories of current, former and never smoker. To exclude for obvious sings of inflammation when we analysed IL-6 and ACT, we adjusted for acute inflammation as defined by C-reactive protein level > 10mg/L. BMI was defined by > 18.50 kg/m^2 underweight, 18.50-24.99 kg/m^2 normal, 25-29.99 kg/m^2 overweight, > 30 kg/m^2 obese.

DATA ANALYSIS

All data analysis was performed with IBM SPSS Statistics 22. We normalized the distribution of IL-6, ACT and CRP by natural logarithmic transformation and used binary logistic regression to estimate odds ratios (OR) and 95% confidence intervals (CIs) of inflammatory markers for depressive symptoms at baseline (1997 - 1999). We treated inflammatory markers, IL-6, ACT, and CRP as continuous variables. To assess the association between the log of the mean IL-6, ACT, CRP and depressive symptoms at 5 year follow-up, we performed linear regression analysis of the CES-D scores at 5 year follow-up (2002-2004) with adjustment for depressive symptoms at baseline (1997-1999). To further explore the association between inflammatory proteins and persistent depressive symptoms over the 5 years (1997/1999 - 2002/2004), multinomial binary logistic analysis were performed using the CES-D scores (≥ 16) at baseline and at 5 year follow-up.

Age (continuous) and sex were controlled for in all analyses. To further analyse the effect of confounding factors, we added potential confounders to the basic model. If this changed the effect estimate meaningfully, the contribution of each variable was individually explored. In subgroup analyses, we assessed the age- and sex-adjusted association following exclusion of participants with acute inflammation or those with low mini mental state examination (MMSE) scores.

RESULTS

Demographic characteristics

Table 1 represents information on socio-demographic and clinical baseline characteristics of the participants. The average age of the study participants was 73 years (range 61.1 - 105.8); 59.9% of whom were women. Among the study participants, the majority was either classified as being overweight or obese (65.3%). Most participants were past smokers 48.4%. Only the minority of participants reported history of physical illness (34%) or were screened positive for cognitive impairment (13.7%). Mean scores of levels of pro-inflammatory biomarkers in study participants are reported.

No association of IL-6, ACT or CRP and depressive symptoms at baseline

To determine if there was a cross-sectional relationship between serum levels of inflammatory biomarkers IL-6, ACT, or CRP and current depressive symptoms at baseline, we performed logistic regression analysis with CES-D scores at baseline. Our data indicate no association between inflammatory proteins IL-6 (*OR*, 1.08; $p = 0.731$), ACT (*OR*, 1.069; $p = 0.670$) and CRP (*OR*, 0.851; $p = 0.819$) and current depressive symptoms at baseline following the adjustments for age and gender.

Table 1. Demographic characteristics of participants in the Rotterdam Study

Demographics		Characteristics
Age (years) (mean, range)		73 (61-106)
Gender (%)		
	Male	40.1
	Female	59.9
Education (%)		
	Low	61.2
	Middle	29.1
	High	9.7
BMI (kg/m^2)		
	>18.50	0.9
	18.50-24.99	33.8
	25-29.99	46.4
	30-40	18.9
Smoking status (%)		
	Non-smoker	35.2
	Past smoker	48.4
	Current smoker	16.4
History of physical illness (%)		34
MMSE score (%)		
	>26	13.7
	≤26	86.3
Interleukin-6 (pg/mL)		4.07 (0.53-80)
α1-Antichymotrypsin (mg/dL)		40.3 (19.5-128.5)
C-Reactive protein (mg/L)		3.39 (0.06-88.8)

BMI body mass index, *physical illness* stroke, history of cardiovascular disease and diabetes, *MMSE* mini-mental state examination.

IL-6 and CRP predict depressive symptoms at 5 year follow-up in older people

To investigate if inflammatory markers are predictive of depressive symptoms at 5-year follow-up, stepwise linear regression were performed with the CES-D scores at 5-year follow-up (Table 2). In our basic model we corrected for age, gender and depressive symptoms at baseline. Subsequently, corrections included acute inflammation, socio-demographic and health characteristics. Our data indicate that IL-6 (*B*, 0.084; $p = 0.016$) and CRP (*B*, 0.086; $p = 0.013$) were significant predictors of depressive symptoms at 5-year follow-up and remained so after correction for all socio-demographic and health characteristics. ACT (*B*, 0.057; $p = 0.083$) showed a trend association with depressive symptoms at 5-year follow-up, following adjustment for age, gender and depressive symptoms at baseline. The trend association for ACT disappeared after correcting for additional socio-demographic and health characteristics including BMI, smoking, physical illness, low MMSE scores and acute inflammation (Table 2).

Table 2. The association between inflammatory proteins and depressive symptoms after 5 years

Inflammatory proteins	No. of cases	*B* (95 % CI)	*p* value
Log IL-6 (per 1 SD increment)α			
Model 1	656	0.107 (1.32, 5.38)	0.001
Model 2	650	0.105 (1.16, 5.33)	0.002
Model 3	650	0.084 (0.48, 4.73)	0.016
Log ACT (per 1 SD increment)α			
Model 1	655	0.057 (-0.67, 10.76)	0.083
Model 2	650	0.048 (-1.57, 10.04)	0.153
Model 3	650	0.020 (-4.20, 7.75)	0.561
Log CRP (per 1 SD increment)α			
Model 1	656	0.090 (0.55, 3.19)	0.006
Model 2	650	0.086 (0.38, 3.16)	0.013

The category "depressive symptoms" includes all subjects who were screened positive. α=levels of inflammatory biomarkers from measurements of 1997-1999: new onset of depressive symptoms at 5-years follow-up in participants with no depressive symptoms at baseline. Subjects on antidepressant medications were excluded from the analysis. *Model* 1 linear regression analysis adjusted for sex and age. *Model 2* as model 1 and additionally adjusted for body mass index (BMI), smoking, physical illness (including stroke, history of cardiovascular disease and diabetes) and mini-mental state examination (MMSE). *Model 3* as model 2 and additionally adjusted for acute inflammation. No model 3 was created for CRP as acute inflammation was calculated as CRP >10 mg/mL; therefore, no adjustment for acute inflammation could be done when CRP was used as predictor. *IL-6* interleukin-6, *ACT* α1-antichymotrypsin, *CRP* C-reactive protein, *SD* standard deviation, *B* standardized beta, *CI* cumulative interval.

The contribution of individual covariates in the association between inflammatory markers and depressive symptoms at 5 year follow-up

Association with IL-6 and CRP

In the secondary analysis, we included adjustments for individual covariates to further assess the effect of each variable on the association between inflammatory markers at baseline and depressive symptoms at 5 year follow-up (Table 3 and 4). Our results indicate that the association between the levels of IL-6 or CRP and depressive symptoms remained strong and was not substantially affected by age, gender, BMI, smoking status, physical illness, low MMSE scores, or acute inflammation (specific to IL-6 analysis only), when adjusted individually or in combination. As acute inflammation was calculated as CRP >10 mg/mL, no adjustment for acute inflammation could be done when CRP was used as a predictor.

In the IL-6 subgroup, exclusion of participants with acute inflammation (*B*, 0.089; $p = 0.010$) or low MMSE scores (*B*, 0.170*;* $p = 0.001$) did not alter the significance of the association between levels of IL-6 and depressive symptoms at 5 year follow-up (Table 3).

Similarly, in the fully adjusted model, IL-6 predicted incident depressive symptoms after 5 years following exclusion of CRP values > 10 mg/L (full Model: *B*, 0.089; $p = 0.013$; $N = 622$).

In contrast to IL-6, associations between CRP and incident depressive symptoms) were no longer observed (Full Model: *B*, 0.044; *p* = 0.223; *N* = 622) following exclusion of subjects with CRP values > 10 mg/l and adjustments of all covariates.

Table3. **The association between levels of IL-6 and depressive symptoms after 5 years**

Inflammatory proteins	No. of cases	*B* (95 % CI)	*p* values
Log IL-6 (per 1 SD increment)α			
Age, sex	656	0.107 (1.32, 5.38)	0.001
Age, sex, BMI (kg/m2)	652	0.109 (1.31, 5.46)	0.001
Age, sex, smoking status	656	0.105 (1.24, 5.32)	0.002
Age, sex, physical illness	656	0.106 (1.26, 5.34)	0.002
Age, sex, MMSE	654	0.106 (1.27, 5.33)	0.002
Age, sex, acute inflammation	656	0.088 (0.68, 4.83)	0.009
Fully adjusted*	650	0.084 (0.48, 4.73)	0.016
Subgroups (age and sex adjusted)			
Excluding those with acute inflammation	627	0.089 (0.68, 4.91)	0.010
Excluding those with MMSE	416	0.170 (2.99, 8.12)	0.001

Individual contribution of each covariate. The category "depressive symptoms" includes all subjects who were screened positive. α = levels of Il-6 from measurements of 1997-1999: new onset of depressive symptoms at 5-year follow-up in participants with no depressive symptoms at baseline. Subjects on antidepressant medications were excluded from the analysis. *IL-6* interleukin-6, *SD* standard deviation, *B* standardized beta, *CI* cumulative interval, *BMI* body mass index, *MMSE* mini-mental state examination

Similarly to IL-6, in the CRP subgroup, exclusion of participants with low MMSE scores (*B*, 0.108; *p* = 0.008), did not affect the association between CRP levels and depressive symptoms (Table 4). Our data reported no significant gender and IL-6 or CRP-interaction in the prediction of depressive symptoms at 5 year follow-up (following adjustment for age); thus, no further analysis stratified for gender were performed.

Table 4. The association between levels of CRP and depressive symptoms after 5 years

Inflammatory proteins	No. of cases	*B* (95 % CI)	*p* values
Log CRP (per 1 SD increment)α			
Age, sex	656	0.090 (0.55, 3.19)	0.006
Age, sex, BMI (kg/m2)	652	0.090 (0.46, 3.23)	0.009
Age, sex, smoking status	656	0.088 (0.50, 3.15)	0.007
Age, sex, physical illness	656	0.089 (0.51, 3.16)	0.007
Age, sex, MMSE	654	0.089 (0.53, 3.16)	0.006
Fully adjusted*	650	0.086 (0.68, 3.16)	0.013
Subgroups (age and sex adjusted)			
Excluding those with MMSE	416	0.108 (0.63, 4.06)	0.008

Individual contribution of each covariate. The category "depressive symptoms" includes all subjects who were screened positive. α = levels of CRP from measurements of 1997-1999: new onset of depressive symptoms at 5-year follow-up in participants with no depressive symptoms at baseline. Subjects on antidepressant medications were excluded from the analysis. *CRP* C-reactive protein, *SD* standard deviation, *B* standardized beta, *CI* cumulative interval, *BMI* body mass index, *MMSE* mini-mental state examination.

Association with ACT

As no association was found between ACT and depressive symptoms at 5 year follow-up, no further analysis was conducted.

IL-6 and CRP are associated with persistent depressive symptoms over 5 years

To investigate whether there was any association between inflammatory proteins and persistent depressive symptoms over 5 years, we generated cut-off scores CES-D $\geq$ 16 at baseline and at 5 year follow-up and subsequently performed logistic analysis. In the basic model 1, IL-6 (*OR*, 2.32; $p = 0.035$) and CRP (*OR*, 1.79; $p = 0.043$) were positively associated with persistent depressive symptoms over 5 years (Table 5). In model 2, this association remained relatively unchanged after adjusting for age, gender, BMI, smoking status, physical illness and low MMSE scores for IL-6 (*OR*, 2.44; $p = 0.03$) and CRP (*OR*, 1.81; $p = 0.052$) respectively. In contrast to the previous observations on the depressive symptoms at 5-year follow-up (incident depression-Table 2.), acute inflammation largely explained the association between IL-6 and persistent depressive symptoms in the fully adjusted model 3 (*OR*, 2.02; $p = 0.107$), (Table 5). Levels of ACT (*OR*, 1.38; $p = 0.80$) were not positively associated with persistent depressive symptoms following adjustment for age and gender. No gender interaction was found for IL-6, ACT or CRP in the prediction of persistence depressive symptoms over 5 years following adjustment for age, thus no further analysis stratified for gender was performed.

Table 5. **The association between inflammatory proteins and persistent depressive symptoms over 5 years**

Inflammatory proteins	No. of cases	Odds ratio (95 % CI)	*p* value
Log IL-6			
Model 1	656	2.32 (1.06, 5.06)	0.035
Model 2	650	2.44 (1.09, 5.45)	0.030
Model 3	649	2.02 (0.86, 4.77)	0.107
Log ACT			
Model 1	654	1.38 (0.12, 16.06)	0.80
Model 2	649	1.23 (0.09, 15.27)	0.871
Model 3	649	0.63 (0.05, 8.81)	0.729
Log CRP			
Model 1	655	1.79 (1.02, 3.13)	0.043
Model 2	649	1.81 (0.99,3.29)	0.052

The category "depressive symptoms" includes all subjects who were screened positive: CES-D ≥ 16 at baseline and at 5-year follow-up. Subjects on antidepressant medications were excluded from the analysis. *Model* 1 multinomial binary logistic regression analysis adjusted for sex and age. *Model 2* as model 1 and additionally adjusted for body mass index (BMI), smoking, physical illness (including stroke, history of cardiovascular disease and diabetes) and mini-mental state examination (MMSE). *Model 3* as model 2 and additionally adjusted for acute inflammation. No model 3 was created for CRP as acute inflammation was calculated as CRP >10 mg/mL; therefore, no adjustment for acute inflammation could be done when CRP was used as predictor. *IL-6* interleukin-6, *ACT* α1-antichymotrypsin, *CRP* C-reactive protein, *SD* standard deviation, *B* standardized beta, *CI* cumulative interval.

The contribution of individual covariates on the association between levels of IL-6 or CRP and persistent depressive symptoms over 5 years

Associations with IL-6 and CRP

We further investigated whether the individual covariates could explain the correlation between IL-6 or CRP and persistent depressive symptoms during the 5 years (Table 6 and 7). Our results indicate that the association between the levels of IL-6 and persistent depressive symptoms was little explained by age, gender, BMI, smoking status, physical illness or MMSE, when adjusted individually. In contrast, acute inflammation, largely explained the association between the levels of IL-6 and persistent depressive symptoms. Indeed correcting for (*OR*, 1.96; $p = 0.113$), or excluding (*OR*, 2.19; $p = 0.068$) participants with acute inflammation reduced this association to insignificant levels (Table 6).

However, IL-6 predicted persistent depressive symptoms following exclusion of CRP values > 10 mg/L and adjustments of all covariates (full model: *OR*, 2.37; $p = 0.049$; N = 622). In contrast, associations between CRP and persistent depressive symptoms were no longer observed following exclusion of subjects with CRP values > 10 mg/l (full model: *OR*, 1.45; $p = 0.310$; N = 621). The data suggest that the association between depression and inflammation seems to be stronger for IL-6 than CRP.

In the subgroup analysis, exclusion of participants with low MMSE scores (*OR*, 2.64; $p = 0.046$) did not affect the positive relationship between levels of IL-6 and persistent depressive symptoms. In contrast to IL-6, association between levels of CRP and persistent depressive symptoms over 5 years were largely explained by adjustment for physical illness (*OR*, 1.70; $p = 0.067$) and by exclusion (*OR*, 1.68; $p = 0.138$), but not adjustment for low MMSE scores (*OR*, 1.80; $p = 0.042$). Associations between CRP and depressive symptoms were little explained by age, gender, BMI, or smoking status.

Associations with ACT

As no associations was found between ACT and depressive symptoms at 5 years, no further analysis were conducted.

Table 6. The association between levels of IL-6 and persistent depressive symptoms over 5 years

Inflammatory proteins	No. of cases	*Odds ratio* (95 % CI)	*p* values
Log IL-6			
Age, sex	656	2.32 (1.06, 5.06)	0.035
Age, sex, BMI (kg/m2)	652	2.54 (1.15, 5.60)	0.021
Age, sex, smoking status	656	2.28 (1.04, 5.00)	0.040
Age, sex, physical illness	656	2.23 (1.01, 4.93)	0.048
Age, sex, MMSE	654	2.38 (1.09, 5.18)	0.030
Age, sex, acute inflammation	655	1.96 (0.85, 4.50)	0.113
Fully adjusted*	649	2.02 (0.86, 4.77)	0.107
Subgroups (age and sex adjusted)			
Excluding those with acute inflammation	626	2.19 (0.94, 5.01)	0.068
Excluding those with MMSE	416	2.64 (1.02, 6.88)	0.046

Individual contribution of each covariate. The category "depressive symptoms" includes all subjects who were screened positive: CES-D ≥ 16 at baseline and at follow-up. Subjects on antidepressant medications were excluded from the analysis. *IL-6* interleukin-6, *CI* cumulative interval, *BMI* body mass index, *MMSE* mini-mental state examination.

Table 7. **The association between levels of CRP and persistent depressive symptoms over 5 years**

Inflammatory proteins	No. of cases	*Odds ratio* (95 % CI)	*p* values
Log CRP			
Age, sex	656	1.79 (1.02, 3.13)	0.043
Age, sex, BMI (kg/m2)	651	1.93 (1.07, 3.49)	0.03
Age, sex, smoking status	655	1.76 (1.01, 3.09)	0.048
Age, sex, physical illness	655	1.70 (0.97, 2.98)	0.067
Age, sex, MMSE	653	1.80 (1.02, 3.16)	0.042
Fully adjusted*	649	1.81 (0.99, 3.29)	0.052
Subgroups (age and sex adjusted)			
Excluding those with MMSE	415	1.68 (0.85, 3.35)	0.138

Individual contribution of each covariate. The category "depressive symptoms" includes all subjects who were screened positive: CES-D ≥ 16 at baseline and at follow-up. Subjects on antidepressant medications were excluded from the analysis. *CRP* C-reactive protein, *CI* cumulative interval, *BMI* body mass index, *MMSE* mini-mental state examination.

DISCUSSION

In this large population based study of elderly persons, we found that the inflammatory biomarkers IL-6 and CRP were longitudinally associated with depressive symptoms. Both IL-6 and CRP levels predicted persistent depressive symptoms over 5 years independently of age, gender, BMI, smoking status or MMSE. Furthermore, we found that IL-6 and CRP predicted incident depressive symptoms at 5 year follow-up after adjustment for all socio-demographic and health characteristics. In contrast, no longitudinal association was observed between the levels of ACT and depressive symptoms. None of the inflammatory biomarkers were associated with depressive symptoms at baseline in the cross-sectional analysis.

To the best of our knowledge, this is the first population-based study to investigate whether inflammation predicts persistent depressive symptoms over 5 years. Persistent depressive symptoms and inflammation had been previously associated in a small study in individuals who were at high risk for coronary heart disease.[30] Our results are consistent with the literature suggesting that activation of inflammatory response system contributes significantly to the maintenance of symptoms of depression, and might be more relevant in people who are more severely ill. Inflammation is particularly observed in major depressive patients who are older[17], have comorbid depression with physical [19,21] and mental illnesses[20], present recurrent[18] or earlier age of onset of the disorder[17]; and in those who are resistant to antidepressant treatment.[22,23] Consistent with this hypothesis, we found that the association between inflammation and persistent depression is lost when adjusted for physical illnesses. It is possible that inflammation distinguishes a subgroup of depressed patients. In line with this idea, Raison et al. showed that the anti-inflammatory drug infliximab showed antidepressant properties only in treatment resistant depressed patients who have high levels of the inflammatory marker C-reactive protein.[24]

Secondly, in this study we report a positive longitudinal relationship between IL-6 or CRP levels and incident depressive symptoms at 5 year follow-up. Our data is consistent with findings from previous literature, which demonstrated a positive longitudinal association between inflammation and incident mental health. Similar to our study, Gimeno et al. found a positive association between CRP, IL-6 and incident depressive symptoms in healthy participants.[31] However, in their research only cognitive depressive symptoms were reported; and mental health was conducted by using the general health questionnaire (GHQ) which is less specific for depressive symptoms. In our study, depressive symptoms were measured using the CES-D scale which is reported to have an excellent sensitivity (100%) and specificity (88%) for major depression in a community based sample of older subjects.[26] Also, in agreement with the present results, Kivimaki et al. showed that persistent elevation of IL-6 levels increases risk of common mental disorder.[32]. Likewise, van den Biggelaar and colleagues showed that baseline levels of CRP significantly predicted incident depression at five year follow-up in elderly participants of > 85 years old.[5]

Further investigations on the link between inflammation and incident depression have also been conducted in different population groups compared to ours. For example, it has been reported that inflammation also predicts depression in females after an accident[33] and in middle-aged adults[4]; but not in younger adults[16,34] This study does not report a cross-sectional association between inflammatory biomarkers and depressive symptoms at baseline. Our data is consistent with findings from Matsushima et al. who reported no cross sectional association between baseline values of hsCRP and IL-6 and depressive symptoms in a community dwelling older participants.[35] Our data is also consistent with that of Krogh et al. 2014 who showed no cross-sectional association between hsCRP and IL-6 levels and depressive symptoms in people with

depression.[36] The fact that no cross-sectional association was found suggest that the response to peripheral infection on producing inflammatory cytokines that act on the brain to cause depression might occur only when this inflammatory stimuli is persistent; and might not occur in a normal acute and transitional response of the inflammatory system.[10]

Several mechanisms have been speculated on how inflammation and depression may be associated. Both clinical and experimental data strongly point toward the involvement of the enzyme indoleamine 2,3 dioxygenase in the development of inflammation-associated major depressive disorders.[10] Reduced function of the hypothalamus-pituitary-adrenal (HPA) axis has been observed in the presence of chronic inflammation in severely ill patients with major depression.[22,23] Inflammation also induces a reduction in the brain derived-neurotrophic factor (BDNF), which negatively influences neurogenesis and neuroplasticity.[37-39] Over time, a decrease in neurogenesis could contribute to the reduction in hippocampal volume seen in major depression[40] and cognitive dysfunction[41] which is positively correlated with a longer duration of depressive symptoms in the community.[42] Furthermore, it has been suggested that pro-inflammatory cytokines can modulate the tryptophan/kynurenine pathway and decrease tryptophan availability for serotonin synthesis.[10,38,43]

We did not find a longitudinal association between levels of ACT and depressive symptoms. Our findings seem surprising because ACT is an acute phase serum glycoprotein, which is positively associated with CRP.[44] A possible explanation for this could be that ACT is a less sensitive marker for inflammation than CRP, which signals early inflammation when other clinical parameters are yet unchanged and therefore, the latter reveals early inflammation when other clinical parameters are equivocal.[45-47]

This study has several strengths. First, the assessment of depressive symptoms at baseline and at 5 year follow-up, allowed us to examine the direction of the association between inflammatory biomarkers and depressive symptoms. Furthermore, we controlled for several confounding factors, which can influence the association between inflammatory markers and depression. Several limitations of our study need also to be addressed. First, our study was restricted to older population of men and women and thus our findings cannot be generalised for the younger population. A second limitation was the fact that we excluded participants who were on antidepressants, and thus the findings might be not generalizable for patients more severely ill and with clinical depression. Finally, we do not have repeated measures of inflammatory cytokines (IL-6, ACT, and CRP), and therefore could not analyse whether long term exposure to high levels of inflammatory markers are associated with depression.

To conclude, our findings indicate that inflammatory biomarkers IL-6 and CRP are longitudinally associated with persistence and incident depressive symptoms in older men and women, after adjustment for socio-demographic and health characteristics. Inflammation seems to distinguish a particular group of people who may benefit from preventive therapies.

REFERENCES

R1 1. Dowlati Y, Herrmann N, Swardfager W, et al. A meta-analysis of cytokines in major depression. *Biological psychiatry.* 2010;67(5):446-457.

R2 2. Howren MB, Lamkin DM, Suls J. Associations of depression with C-reactive protein, IL-1, and IL-6: a meta-analysis. *Psychosomatic medicine.* 2009;71(2):171-186.

R3 3. Liu Y, Al-Sayegh H, Jabrah R, Wang W, Yan F, Zhang J. Association between C-reactive protein and depression: modulated by gender and mediated by body weight. *Psychiatry research.* 2014;219(1):103-108.

R4 4. Matthews KA, Schott LL, Bromberger JT, Cyranowski JM, Everson-Rose SA, Sowers M. Are there bi-directional associations between depressive symptoms and C-reactive protein in mid-life women? *Brain, behavior, and immunity.* 2010;24(1):96-101.

R5 5. van den Biggelaar AH, Gussekloo J, de Craen AJ, et al. Inflammation and interleukin-1 signaling network contribute to depressive symptoms but not cognitive decline in old age. *Experimental gerontology.* 2007;42(7):693-701.

R6 6. Alavi M, Grebely J, Matthews GV, et al. Effect of pegylated interferon-alpha-2a treatment on mental health during recent hepatitis C virus infection. *Journal of gastroenterology and hepatology.* 2012;27(5):957-965.

R7 7. Birerdinc A, Afendy A, Stepanova M, Younossi I, Baranova A, Younossi ZM. Gene expression profiles associated with depression in patients with chronic hepatitis C (CH-C). *Brain and behavior.* 2012;2(5):525-531.

R8 8. Chourbaji S, Urani A, Inta I, et al. IL-6 knockout mice exhibit resistance to stress-induced development of depression-like behaviors. *Neurobiology of disease.* 2006;23(3):587-594.

R9 9. Dantzer R. Cytokine-induced sickness behavior: mechanisms and implications. *Annals of the New York Academy of Sciences.* 2001;933:222-234.

R10 10. Dantzer R, O'Connor JC, Freund GG, Johnson RW, Kelley KW. From inflammation to sickness and depression: when the immune system subjugates the brain. *Nature reviews. Neuroscience.* 2008;9(1):46-56.

R11 11. Felger JC, Alagbe O, Hu F, et al. Effects of interferon-alpha on rhesus monkeys: a nonhuman primate model of cytokine-induced depression. *Biological psychiatry.* 2007;62(11):1324-1333.

R12 12. McNamara RK, Lotrich FE. Elevated immune-inflammatory signaling in mood disorders: a new therapeutic target? *Expert review of neurotherapeutics.* 2012;12(9):1143-1161.

R13 13. Simen BB, Duman CH, Simen AA, Duman RS. TNFalpha signaling in depression and anxiety: behavioral consequences of individual receptor targeting. *Biological psychiatry.* 2006;59(9):775-785.

R14 14. Valkanova V, Ebmeier KP, Allan CL. CRP, IL-6 and depression: a systematic review and meta-analysis of longitudinal studies. *Journal of affective disorders.* 2013;150(3):736-744.

R15 15. Wium-Andersen MK, Orsted DD, Nielsen SF, Nordestgaard BG. Elevated C-reactive protein levels, psychological distress, and depression in 73, 131 individuals. *JAMA psychiatry.* 2013;70(2):176-184.

R16 16. Stewart JC, Rand KL, Muldoon MF, Kamarck TW. A prospective evaluation of the directionality of the depression-inflammation relationship. *Brain, behavior, and immunity.* 2009;23(7):936-944.

R17 17. Grosse L, Carvalho LA, Wijkhuijs AJ, et al. Clinical characteristics of inflammation-associated depression: Monocyte gene expression is age-related in major depressive disorder. *Brain, behavior, and immunity.* 2014.

R18 18. Ford DE, Erlinger TP. Depression and C-reactive protein in US adults: data from the Third National Health and Nutrition Examination Survey. *Archives of internal medicine.* 2004;164(9):1010-1014.

R19 19. Jones DR, Macias C, Barreira PJ, Fisher WH, Hargreaves WA, Harding CM. Prevalence, severity, and co-occurrence of chronic physical health problems of persons with serious mental illness. *Psychiatric services.* 2004;55(11):1250-1257.

R20 R21 R22 R23 R24 R25 R26 R27 R28 R29 R30 R31 R32 R33 R34 R35 R36 R37 R38 R39

20. Liukkonen T, Rasanen P, Jokelainen J, et al. The association between anxiety and C-reactive protein (CRP) levels: results from the Northern Finland 1966 birth cohort study. *European psychiatry : the journal of the Association of European Psychiatrists.* 2011;26(6):363-369.
21. Moussavi S, Chatterji S, Verdes E, Tandon A, Patel V, Ustun B. Depression, chronic diseases, and decrements in health: results from the World Health Surveys. *Lancet.* 2007;370(9590):851-858.
22. Carvalho LA, Bergink V, Sumaski L, et al. Inflammatory activation is associated with a reduced glucocorticoid receptor alpha/beta expression ratio in monocytes of inpatients with melancholic major depressive disorder. *Translational psychiatry.* 2014;4:e344.
23. Carvalho LA, Juruena MF, Papadopoulos AS, et al. Clomipramine in vitro reduces glucocorticoid receptor function in healthy subjects but not in patients with major depression. *Neuropsychopharmacology : official publication of the American College of Neuropsychopharmacology.* 2008;33(13):3182-3189.
24. Raison CL, Rutherford RE, Woolwine BJ, et al. A randomized controlled trial of the tumor necrosis factor antagonist infliximab for treatment-resistant depression: the role of baseline inflammatory biomarkers. *JAMA psychiatry.* 2013;70(1):31-41.
25. Hofman A, Grobbee DE, de Jong PT, van den Ouweland FA. Determinants of disease and disability in the elderly: the Rotterdam Elderly Study. *European journal of epidemiology.* 1991;7(4):403-422.
26. Beekman AT, Deeg DJ, Van Limbeek J, Braam AW, De Vries MZ, Van Tilburg W. Criterion validity of the Center for Epidemiologic Studies Depression scale (CES-D): results from a community-based sample of older subjects in The Netherlands. *Psychol Med.* 1997;27(1):231-235.
27. Luijendijk HJ, van den Berg JF, Dekker MJ, et al. Incidence and recurrence of late-life depression. *Arch Gen Psychiatry.* 2008;65(12):1394-1401.
28. Blazer D, Burchett B, Service C, George LK. The association of age and depression among the elderly: an epidemiologic exploration. *Journal of gerontology.* 1991;46(6):M210-215.
29. Ikram MA, van Oijen M, de Jong FJ, et al. Unrecognized myocardial infarction in relation to risk of dementia and cerebral small vessel disease. *Stroke; a journal of cerebral circulation.* 2008;39(5):1421-1426.
30. Azar R, Nolan RP, Stewart DE. Listening to the heart-brain talk: persistent depressive symptoms are associated with hsCRP in apparently healthy individuals at high risk for coronary artery disease. *European journal of preventive cardiology.* 2012;19(4):857-863.
31. Gimeno D, Kivimaki M, Brunner EJ, et al. Associations of C-reactive protein and interleukin-6 with cognitive symptoms of depression: 12-year follow-up of the Whitehall II study. *Psychol Med.* 2009;39(3):413-423.
32. Kivimaki M, Shipley MJ, Batty GD, et al. Long-term inflammation increases risk of common mental disorder: a cohort study. *Molecular psychiatry.* 2014;19(2):149-150.
33. Matheny ME, Miller RR, Shardell MD, et al. Inflammatory cytokine levels and depressive symptoms in older women in the year after hip fracture: findings from the Baltimore Hip Studies. *Journal of the American Geriatrics Society.* 2011;59(12):2249-2255.
34. Deverts DJ, Cohen S, DiLillo VG, et al. Depressive symptoms, race, and circulating C-reactive protein: the Coronary Artery Risk Development in Young Adults (CARDIA) study. *Psychosomatic medicine.* 2010;72(8):734-741.
35. Matsushima J, Kawashima T, Nabeta H, et al. Association of inflammatory biomarkers with depressive symptoms and cognitive decline in a community-dwelling healthy older sample: a 3-year follow-up study. *Journal of affective disorders.* 2015;173:9-14.
36. Krogh J, Benros ME, Jorgensen MB, Vesterager L, Elfving B, Nordentoft M. The association between depressive symptoms, cognitive function, and inflammation in major depression. *Brain, behavior, and immunity.* 2014;35:70-76.
37. Cortese GP, Barrientos RM, Maier SF, Patterson SL. Aging and a peripheral immune challenge interact to reduce mature brain-derived neurotrophic factor and activation of TrkB, PLCgamma1, and ERK in hippocampal synaptoneurosomes. *The Journal of neuroscience : the official journal of the Society for Neuroscience.* 2011;31(11):4274-4279.

38. Felger JC, Lotrich FE. Inflammatory cytokines in depression: neurobiological mechanisms and therapeutic implications. *Neuroscience.* 2013;246:199-229.
39. Lotrich FE, Albusaysi S, Ferrell RE. Brain-derived neurotrophic factor serum levels and genotype: association with depression during interferon-alpha treatment. *Neuropsychopharmacology : official publication of the American College of Neuropsychopharmacology.* 2013;38(6):985-995.
40. Campbell S, Macqueen G. The role of the hippocampus in the pathophysiology of major depression. *Journal of psychiatry & neuroscience : JPN.* 2004;29(6):417-426.
41. Hoglinger GU, Rizk P, Muriel MP, et al. Dopamine depletion impairs precursor cell proliferation in Parkinson disease. *Nature neuroscience.* 2004;7(7):726-735.
42. Beekman AT, Penninx BW, Deeg DJ, de Beurs E, Geerling SW, van Tilburg W. The impact of depression on the well-being, disability and use of services in older adults: a longitudinal perspective. *Acta psychiatrica Scandinavica.* 2002;105(1):20-27.
43. Raedler TJ. Inflammatory mechanisms in major depressive disorder. *Current opinion in psychiatry.* 2011;24(6):519-525.
44. Schaap LA, Pluijm SM, Deeg DJ, Visser M. Inflammatory markers and loss of muscle mass (sarcopenia) and strength. *The American journal of medicine.* 2006;119(6):526 e529-517.
45. Becking K, Boschloo L, Vogelzangs N, et al. The association between immune activation and manic symptoms in patients with a depressive disorder. *Translational psychiatry.* 2013;3:e314.
46. Chard MD, Calvin J, Price CP, Cawston TE, Hazleman BL. Serum alpha 1 antichymotrypsin concentration as a marker of disease activity in rheumatoid arthritis. *Annals of the rheumatic diseases.* 1988;47(8):665-671.
47. Lu Y, Feng L, Feng L, Nyunt MS, Yap KB, Ng TP. Systemic inflammation, depression and obstructive pulmonary function: a population-based study. *Respiratory research.* 2013;14:53.

APPENDIX of CHAPTER 2.1

Supplementary material 1

Longitudinal analysis of the association between serum vitamin D levels and incident major depressive disorder with Cox regression analysis 5 years follow-up time (*N* = 2 466)

Vitamin D serum levels	*events*	Major depressieve disorder			
		Model 1		Model 2	
		HR (95%CI)	*p*	*HR* (95%CI)	*P*
Continuously					
Vitamin D √*SD* nmol/l[a]	91	0.96 (0.75 to 1.23)	0.77	0.92 (0.80 to 1.05)	0.53
Cut-off					
<37.5 nmol/L	45	0.90 (0.58 to 1.41)	0.66	1.03 (0.65 to 1.63)	0.91
>37.5 nmol/L	46	**Reference**		**Reference**	
<50 nmol/L	61	0.90 (0.56 to 1.43)	0.65	0.96 (0.75 to 1.22)	0.86
>50 nmol/L	30	**Reference**		**Reference**	
<75 nmol/L	86	1.65 (0.66 to 4.14)	0.29	1.51 (0.94 to 2.42)	0.37
>75 nmol/L	5	**Reference**		**Reference**	
Quartiles					
<28,57 nmol/L	25	0.68 (0.41 to 1.12)	0.13	0.74 (0.44 to 1.27)	0.29
28,58-43,81 nmol/L	29	1.27 (0.81 to 1.98)	0.30	1.31 (0.99 to 1.06)	0.25
43,82-63,21 nmol/L	28	1.62 (1.03 to 2.54)	0.039	1.37 (0.86 to 2.18)	0.20
>63.21 nmol/L	9	**Reference**		**Reference**	

Abbreviations: *SD*, standard deviation; *HR*, hazard ratio; CI, confidence interval.

[a] Standard Deviation of the square root of the Vitamin D serum levels (nmol/L)

Model 1. Adjusted for gender, age and baseline depressive symptoms; Model 2. Additionally adjusted for: body mass index, alcohol consumption, smoking status, marital status and activity of daily living score.

Longitudinal analysis of the association between serum vitamin D levels and incident major depressive disorder with Cox regression analysis 2 years follow-up time (*N* = 2 466)

Vitamin D serum levels	*events*	Major depressieve disorder			
		Model 1		Model 2	
		HR (95%CI)	*p*	*HR* (95%CI)	*p*
Continuously					
Vitamin D √*SD* nmol/l[a]	32	1.10 (0.70 to 1.75)	0.67	0.88 (0.68 to 1.15)	0.63
Cut-off					
<37.5 nmol/L	16	0.87 (0.41 to 1.86)	0.72	1.24 (0.81 to 1.91)	0.61
>37.5 nmol/L	16	**Reference**		**Reference**	
<50 nmol/L	21	0.64 (0.28 to 1.47)	0.29	1.1 (0.70 to 1.72)	0.83
>50 nmol/L	11	**Reference**		**Reference**	
<75 nmol/L	30	0.72 (0.16 to 3.22)	0.66	0.70 (0.32 to 1.52)	0.65
>75 nmol/L	2	**Reference**		**Reference**	
Quartiles					
<28,57 nmol/L	10	0.77 (0.35 to 1.69)	0.51	1.20 (0.73 to 1.97)	0.71
28,58-43,81 nmol/L	9	0.99 (0.45 to 2.19)	0.99	1.02 (0.46 to 2.27)	0.97
43,82-63,21 nmol/L	10	1.63 (0.74 to 3.61)	0.23	1.33 (0.86 to 2.06)	0.51
>63.21 nmol/L	3	**Reference**		**Reference**	

Abbreviations: *SD*, standard deviation; *HR*, hazard ratio; CI, confidence interval.

[a] Standard Deviation of the square root of the Vitamin D serum levels (nmol/L)

Model 1. Adjusted for gender, age and baseline depressive symptoms Model 2. Additionally adjusted for: body mass index, alcohol consumption, smoking status, marital status and activity of daily living score.

CHAPTER 3

The genetics of depression

CHAPTER 3.1

A rare *Asn396Ser* variant in the *LIPG* gene associated with depressive symptoms

Amin Najaf, Jovanova Olivera, Adams Hieab, Dehghan Abbas, Kavousi Maryam, et al.

Molecular Psychiatry. 2017 Apr; 22(4):537-543.

ABSTRACT

Despite a substantial genetic component, efforts to identify common genetic variation underlying depression have largely been unsuccessful. In the current study we aimed to identify rare genetic variants that might have large effects on depression in the general population. Using high coverage exome-sequencing, we studied the exonic variants in 1 265 individuals from the Rotterdam study (RS) who were assessed for depressive symptoms. We identified a missense *Asn396Ser* mutation (rs77960347) in the endothelial lipase (*LIPG*) gene, occurring with an allele frequency of 1% in the general population, that was significantly associated with depressive symptoms (*p value* = $5.2*10^{-08}$, β = 7.2). Replication in three independent datasets (*N* = 3 612) confirmed the association of *Asn396Ser* (*p value* = $7.1*10^{-03}$, β = 2.55) with depressive symptoms. *LIPG* is predicted to have enzymatic function in steroid biosynthesis, cholesterol biosynthesis and thyroid hormone metabolic processes. The *Asn396Ser* variant is predicted to have a damaging effect on the function of *LIPG*. Within the discovery population, carriers also showed an increased burden of white matter lesions (*p value* = $3.3*10^{-02}$) and a higher risk of Alzheimer's disease (odds ratio = 2.01; *p value* = $2.8*10^{-02}$) compared to the non-carriers. Together, these findings implicate the *Asn396Ser* variant of *LIPG* in the pathogenesis of depressive symptoms in the general population.

INTRODUCTION

Depression is a common psychiatric illness with a substantial heritability (h^2 = 40-50%).[1] Despite decades of research, its genetic origin remains largely obscure. Previously candidate gene studies suggested a role of various genes in the monoaminergic neurotransmission pathways. These include *MTHFR, 5-HTTLPR, HTR1A, HTR2A, TPH1* and *BDNF*. However, these findings were not replicated in recent large case-control studies based on array-based genome-wide association studies (GWAS). The hypothesis-free large-scale GWAS have led to the discovery of hundreds of genetic variants contributing to the risk of a wide diversity of common human traits and diseases[2,3], however, the Psychiatric Genetics Consortium (PGC) Working Group for Major Depressive Disorder (MDD) did not identify a genome-wide significant locus for depression that could be replicated.[3] Also PGC could not replicate the two common genetic variants that were recently implicated in MDD in a GWAS in the Han-Chinese population.[4,5] The failure to elucidate the genes underlying MDD is mainly attributed to the fact that MDD is more common, less heritable and more heterogeneous compared to other psychiatric diseases like Schizophrenia[6], genes for which were successfully identified by the same consortium.[7] High heterogeneity of the disease phenotype and inclusion of a substantial number of non-genetic patients that develop a major depression due to life-events dilutes the effects of the genetic variants and diminishes statistical power.[3] To overcome the issues with statistical power relating to heterogeneity, large-scale GWAS (N > 30 000) of sub-domains of depression and quantitative endophenotypes have been performed.[8,9] Endophenotypes are heritable quantitative factors[10,11], that have been successfully used to unravel the genetics of other heterogeneous diseases with major misclassification such as hypertension.[12] For depression, however, these were also not successful.[8,9]

Two important lessons learned from the array-based GWAS in general. First, common variants have small effects and consortia of mega-sizes are needed to detect these. Second, a substantial part of the heritability is not accounted for by the common variants. The variance of MDD explained by common genomic variation across the entire genome is estimated to be ~21%[6,13] and 4% for quantitative outcomes.[8] These estimates suggest there may be various alternative genetic mechanisms that explain the heritability such as rare variants. Also the search for rare variants has a long history starting with linkage analysis – a method studying co-segregation of chromosomal regions with the disease in families. These regions presumably harbour rare-causal variants that have relatively large effects on the disease risk.[14] The identified linked regions are followed up with various fine-mapping techniques to identify the putative causal variant.[14] Linkage of depression has been reported to large genomic regions including 15q25[15,16], 17p[17], 7p[18], 8p[17], 12q23[19,20], 18q[18], 1p36[20] and 13q31[1,20]. The presumed underlying rare-causal variants were never identified. Notably such rare variants are usually not present on microarrays used for large-scale GWAS and despite recent improvements in genotype imputation methods and

the reference haplotypes used for genetic imputations[21], are difficult to impute with sufficient confidence. An alternative approach is next generation sequencing (NGS). This is a highly effective method for identifying rare variants with relatively large effects. Such variants may be anticipated in exonic or coding regions.

In this study, we performed exome sequencing in a large population-based cohort to search for rare exonic variants associated with depressive symptoms measured using the Centre for Epidemiologic Studies Depression (CES-D)[22], as an endophenotype of MDD. Findings were replicated in three independent studies. As endophenotypes are not 1:1 related to disease, we further studied the relation of the variants associated with depressive symptoms to a range of other diseases including Alzheimer's disease, changes in the brain structure, thyroid pathology, and metabolic changes.

METHODS

Study populations

Our discovery population consists of subjects from the Rotterdam study (RS). RS is a prospective, population-based study from the well-defined district of Ommoord within the city of Rotterdam, designed to investigate the occurrence and determinants of diseases in the elderly.[23] The RS cohort was initially defined in 1990 among 7 983 persons who underwent a home interview and extensive physical examination at baseline and during follow-up examinations occurring every 3 - 4 years (RS-I). The cohort was further extended in 2000 (RS-II) and 2005 (RS-III), establishing a total of 14 926 participants.[23] The Rotterdam Study was approved by the Medical Ethics Committee of the Erasmus MC and by the Ministry of Health, Welfare and Sport of the Netherlands. All participants provided written informed consent to participate in the study and to obtain information from their treating physicians.

Participants from RS-I were assessed for depressive symptoms using the 20-item version of the CES-D scale. CES-D is a self-report measure of symptoms experienced during the prior week. It has shown to be relatively stable over time (82% of older adults had stable CES-D scores over four measurement rounds in 10 years)[9,24,25] and covers the following major dimensions of depression: depressed mood, feelings of guilt and worthlessness, feelings of helplessness and hopelessness, psychomotor retardation, loss of appetite, and sleep disturbance.[9] The total score ranges from 0 - 60, with higher scores indicating a greater burden of depressive symptoms. The heritability of the CES-D is estimated to be 0.24[26] and it detects current MDD cases with high sensitivity and specificity.[27,28] A score of >16 is considered to indicate a possible clinical depression.[29] In order to

maximize the number of eligible participants, data from the third follow-up visit was used in this study.

Since 2005, magnetic resonance imaging (MRI) of the brain has been acquired at 1.5 Tesla during RS study examinations. The MRI protocol incorporates several sequences, including a T1-weighted sequence, a proton density weighted sequence, and a fluid-attenuated inversion recovery (FLAIR) sequence. These sequences were automatically processed using custom-developed multimodal algorithms to quantify brain tissue volumes.[30,31] The processed images distinguished between grey matter, white matter, white matter lesions, and cerebrospinal fluid, and the volumes of each of these tissues were calculated by multiplying the number of voxels with that label with the volume of an individual voxel, expressed in ml.

Various other disease related to depressive systems have been characterized in the discover cohort. Alzheimer's disease (AD) was characterized using the National Institute of Neurological Disorders and Stroke–Association Internationale pour la Recherche et l'Enseignement en Neurosciences - Alzheimer's Disease and Related Disorders Association (NINCDS-ADRDA).[32] Thyroid stimulating hormone (TSH) and free thyroxine (FT4) measurements were performed in serum samples stored at - 80°C in samples collected between 1997 and 2000. Thyroxine, thyrotropin, and thyroid perioxidase antibodies were quantified by electrochemiluminescence immunoassay (ECLIA, Roche).[33]

Independent replications were performed in the Erasmus Rucphen Family (ERF) study. ERF is a family-based study that includes inhabitants of a genetically isolated community in the South-West of the Netherlands, ascertained as part of the Genetic Research in Isolated Population (GRIP) program.[34] The ERF cohort includes approximately 3 000 living descendants of 22 founder couples who had at least six children baptized in the community church. All data were collected between 2002 and 2005. The population shows minimal immigration and high inbreeding and therefore the frequency of many rare alleles is increased in this population.[34] All participants provided informed consent. The Medical Ethical Committee of the Erasmus University Medical Centre approved the ERF study. All study participants were assessed for depressive symptoms using Hospital Anxiety and Depression Scale (HADS) and CES-D. HADS is a 14-item scale of which 7 items relate to anxiety and 7 items to depression. For depression, the total score ranges from 0 to 21, with higher scores indicating a greater burden of depressive symptoms. A cut off of ≥ 8 is considered indicative of likely clinical depressive episodes.[35]

Subjects were extensively profiled for metabolomics using either plasma or serum samples collected at the time of visit and stored at - 80°C. Metabolic profiling was performed either by mass spectrometry (MS) or nuclear magnetic resonance (NMR), using five metabolite platforms:

one set by the AbsoluteIDQTM p150 Kit (Biocrates Life Sciences AG) (*N* = 937)[36], two sets by either liquid chromatography coupled MS or Electro-Spray Ionization coupled MS[37,38] (*N* = 2 416 and *N* = 822 respectively), two sets measured by NMR using small molecular compounds[39] (*N* = 2 416, measured by Center for Proteomics and Metabolomics, Leiden University Medical Center, Leiden, the Netherlands) and lipoprotein extraction windows (*N* = 2 609, measured by Center for Proteomics and Metabolomics, Leiden University Medical Center, Leiden, the Netherlands and deconvoluted by the commercial algorithm of Bruker Corporation NMR service protocol). The majority of the blood samples (97%) were obtained during fasting. In total, 2 864 participants from the ERF study were profiled for metabolomics on at least one of the five platforms and 563 metabolites were annotated, which consisted of lipoprotein sub-particles, amino acids, a large array of (lyso)phosphatidylcholines, sphingolipids, ceramides, triglycerides and low-molecular-weight compounds.

Exome sequencing

In the Rotterdam study, exomes of randomly selected subset of 2 628 individuals from the RS-I population were sequenced at the Human Genotyping facility of the department of Internal Medicine, at Erasmus MC, the Netherlands. The sequencing was performed at an average depth of 54x using the Nimblegen SeqCap EZ V2 capture kit on an Illumina Hiseq2000 sequencer using the TrueSeq Version 3 protocol. The sequence reads were aligned to hg19 using Burrows-Wheeler Aligner (BWA).[40] Subsequently, the aligned reads were processed further using Picard's MarkDuplicates, SAMtools[41] and the Indel Realignment and Base Quality Score Recalibration tools from Genome Analysis Toolkit (GATK).[42] Genetic variants were called using the HaplotypeCaller from GATK. Samples with low concordance to genotyping array (< 95%), or that differed 4 standard deviations from the mean on either the number of detected variants per sample, transition to transversion ratio or high heterozygote to homozygote ratio and low call rate (< 90%) were removed from the data. Single Nucleotide Variants (SNVs) with a low call rate (< 90%) and out of Hardy-Weinberg equilibrium (HWE; p-value < 10^{-8}) were also removed from the data. The final dataset consisted of 600 806 SNVs in 2 356 individuals. Of these, 1 265 were assessed for depressive symptoms and were used in this study. The average age at examination was 72.3 years and the majority were women (59%; Supplementary Table 1). File handling and formatting was done using VCFtools[43] and PLINK[44]. Annotation of the variants was performed using SeattleSeq annotation138 (http://snp.gs.washington.edu/SeattleSeqAnnotation138/).

Exomes of 1 336 randomly selected individuals from the ERF study cohort were sequenced at the Center for Biomics of the Cell Biology department, at the Erasmus MC, The Netherlands. Sequencing was done at a median depth of 57x using the Agilent version V4 capture kit, on an Illumina Hiseq2000 sequencer, using the TruSeq Version 3 protocol. The sequence reads were

aligned to the human genome build 19 (hg19), using Burrows-Wheeler Aligner (BWA) and the NARWHAL pipeline.[40,45] Subsequently, the aligned reads were processed further, using the IndelRealigner, MarkDuplicates and Table Recalibration tools from the Genome Analysis Toolkit (GATK)[42] and Picard (http://picard.sourceforge.net). This was necessary to remove systematic biases and to recalibrate the PHRED quality scores in the alignments. After processing, genetic variants were called, using the Unified Genotyper tool from the GATK.[42] For each sample, at least 4 Gigabases of sequence was aligned to the genome. Functional annotations were also performed using the SeattleSeq annotation 138 database. About 1.4 million SNVs were called. After removing variants with a low quality, out of HWE (*p value* < 10^{-6}) and low call rate (< 99%), and samples with a low call rate (< 90%), we retrieved 543 954 very high quality SNVs in 1 327 individuals. Of these, 1 247 individuals (60% women, mean age = 48.5 years) were assessed for depressive symptoms and included in the replications analysis.

Exome-chip genotyping

Study participants from the ERF cohort whose exomes were not sequenced (*N* = 1 527) were genotyped on the Illumina Infinium HumanExome BeadChip, version 1.1, which contains over 240 000 exonic variants selected from multiple sources together spanning 12 000 samples from multiple ethnicities. Calling was performed with GenomeStudio. We removed subjects with a call rate < 0.95, IBS > 0.99 and heterozygote ratio > 0.60. Ethnic outliers identified using a principal component analysis with 1 000 Genomes data (Supplementary Figure 1) and individuals with sex discrepancies were also removed. The SNVs that were monomorphic in our sample or had a call rate < 0.95 were removed. After quality control we retrieved about 70 000 polymorphic SNVs in 1 515 subjects. Of these, 840 individuals (54% women, mean age = 50 years) were assessed for depressive symptoms and included in the replication analysis.

In the RS cohort, 3 163 (1 949 independent) samples were genotyped using the Illumina Human Exome BeadChip v1.0. Genotype calling was performed according to the CHARGE joint calling protocols.[46] For replication, we used 1 525 individuals (52.6% women, mean age = 73.7 years) who were assessed for depressive symptoms but not exome sequenced.

DATA ANALYSIS

Single variant quantitative trait association analysis of depressive symptoms was performed using the seqMeta (v1.4) library of the R software in RS. We included only missense, nonsense, frameshift, or essential splice site bi-allelic variants that had a minor allele count > 7 in the RS discovery set. In the ERF replication samples, for sequence data we used a linear mixed effects

model in seqMeta (V1.4) using the kinship matrix estimated from the pedigree data in the kinship2 library of R. For the exome-chip data, we used FamScore test implemented in RVtests (http://zhanxw.github.io/rvtests/) using the kinship matrix estimated from the genotyped data to adjust for familial relationships. Analysis of the RS exome-chip data was performed using seqMeta library of the R software. All analyses were adjusted for age and sex. Meta-analysis of the replication studies was performed in R using the library rmeta. Association analyses with other dichotomous and quantitative traits were performed using logistic and linear regression analysis respectively, adjusted for age and sex in PLINK v1.07[44] (http://pngu.mgh.harvard.edu/purcell/plink/) or SPSS v21. Analysis of the metabolomics data in the ERF cohort was performed using linear regression adjusted for age and sex. Skewed metabolites were transformed using either natural logarithm or inverse-normal transformation. We used polygenic residuals of the metabolites after regressing out the effect of relationships in a linear mixed model in GenABEL. We used Bonferroni corrected *p*-value thresholds in ERF (0.05/563 metabolites = $8.8*10^{-05}$) to declare significance.

RESULTS

Genome-wide association and quantile-plots are provided in Supplementary Figures 2 and 3. We observed significant association of depressive symptoms with a low frequency variant (rs77960347-G; MAF = 1%, *N* (carriers) = 60; *p value* = $5.2*10^{-08}$) on chromosome 18 (Table 1a, Figure 1). The estimated additive effect of the minor allele (G) on depressive symptoms was large (β = 7.2) (Figures 2 and 3), suggesting that each allele yields an increase of depressive symptoms by 7.2 units. Rs77960347 is a highly conserved missense variant in the LIPG gene (*Asn396Ser* phastCons = 1) and predicted to be damaging (polyPhen = 1). In the replication cohort ERF, rs77960347-G was observed with a similar frequency to the discovery sample. The replication sample demonstrated a significant association (*p value* = $7.1*10^{-03}$) of rs77960347-G with the depressive symptoms, and also a large effect (β = 2.55) (Table 1b, Figure 3).

Table1a. Significant variants from exome-wide association analysis in the RS (discovery)

Name	Chr	Position	minor/ major	RD	GQ	gene	*p*	maf	*N*	beta§	se	β‡	95% CI	Residue Change	function	polyPhen score^{ss}	score PhastCons	CADD score
rs77960347	18	47109955	G/A	18	44	LIPG	**5.2*10^{-08}**	0.011	1265	7.20	1.32	0.151	4.6-9.8	Asn,Ser	missense	1	1	19.61

§ Effect of the variant on CES-D refers to the minor allele
‡ Standardized beta
^{ss} SeattleSeq annotation database 138
RD: mean read-depth for the variant
GQ: mean Phred-scale genotype quality of the variant
CADD: Combined annotation dependent depletion score

Table1b. Results of replication

Name	ERF sequence							ERF chip					RS chip					meta replication		
	RD	GQ	*p*	maf	*N*	beta§	se	*p*	maf	*N*	beta§	se	*p*	maf	*N*	beta§	se	*p*	beta§	se
rs77960347	39	94	0.075	0.009	1247	3.94	2.21	0.277	0.014	840	2.74	1.94	0.104	0.010	1525	2.03	1.25	**7.1*10^{-03}**	2.55	0.95

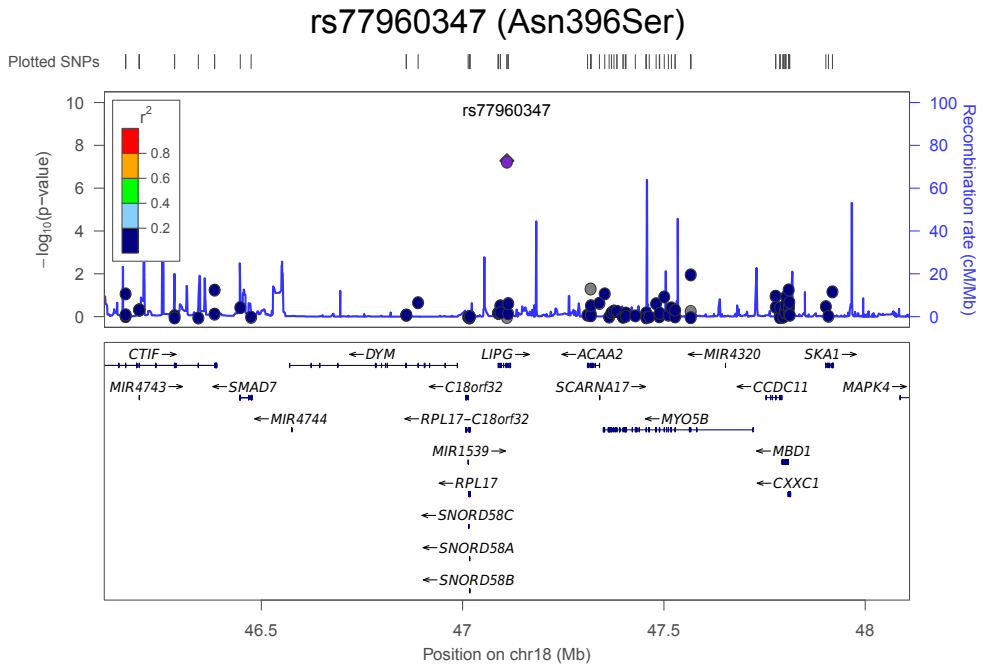
§ Effect of the variant on CES-D refers to the minor allele

Apart from having large effects on depressive symptoms, the carriers of rs77960347-G on average had a smaller brain volume (both white and grey matter) and a higher burden of white matter lesions (*p value* = $3.3*10^{-02}$; based on the subset of 10/60 carriers with an MRI available), compared to non-carriers (Table 2).

Table 2. Brain imaging characteristics of carriers and non-carriers of *Asn396Ser*

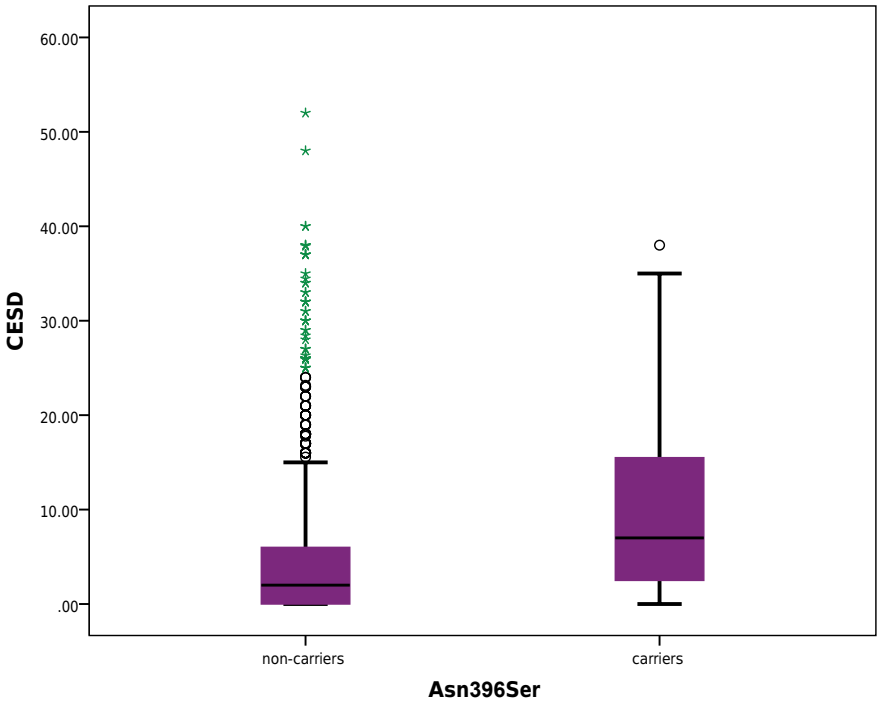
	A/A (*N*=398)	A/G (*N*=10)	*p* value
Demographics			
Age, years	79.0 (4.8)	79.0 (4.8)	0.97
Females	222 (55.8%)	7 (70.0%)	0.37
Brain volumetry, mL			
Intracranial volume	1137.2 (119.6)	1136.5 (93.8)	0.99
Grey matter volume	508.6 (52.1)	501.2 (49.8)	0.66
White matter volume	378.9 (55.1)	377.9 (51.1)	0.96
Cerebrospinal fluid volume	247.4 (51.2)	255.0 (47.8)	0.64
White matter lesion volume	8.3 (3.9 - 17.1)	24.6 (5.4 - 34.5)	**0.03**

Figure 1. Regional association plot for the top hit (*Asn396Ser*).



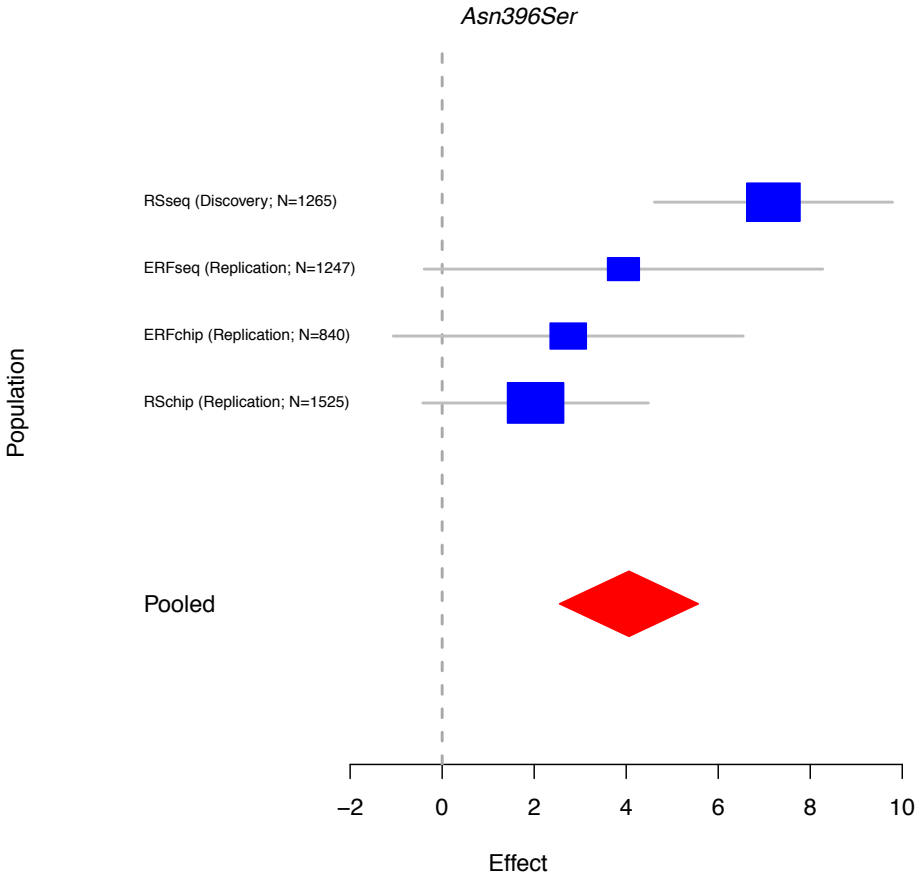
Vertical axis shows the strength of association in terms of negative principal logarithm of the association *p*-value. Dots represent the SNVs tested in the region. Top variant rs77960347 is indicated in purple. Colors of the dots indicate the extent of linkage disequilibrium with the top hit. All genes in the region are shown below the horizontal axis.

Figure 2. Box plot for the carriers and non-carriers of the minor allele (G).



Vertical axis shows the CES-D score. Purple box region shows the interquartile range. The horizontal bar in the box indicates the median. Lower and upper whiskers indicate the first and the fourth quartiles respectively. Circles indicate outliers and stars indicate extreme values.

Figure 3. Forest plot for the top hit (*Asn396Ser*).



Vertical axis represents all the studies. Horizontal axis shows the size of the effect of the variant. Size of the blue boxes indicates the precision. The horizontal lines adjacent the blue boxes indicate the confidence interval for the study. The red diamond indicates the summary effect estimate. The width of the diamond indicates the confidence interval for the summary effect.

Given that the *LIPG* gene is expressed in the thyroid gland and has been implicated in several cardiometabolic disorders, including dyslipidemia[47], atherosclerosis[48,49], cardiovascular diseases[50], and inflammation[51], we additionally tested the association of *Asn396Ser* with these conditions and their related phenotypes (Table 3). No association with thyroid functioning, cardiovascular outcomes, inflammation, atherosclerosis, or hypertension was observed (Table 3). Total cholesterol was increased in carriers (*p value* = $5.3*10^{-03}$; Table 3). The association of *Asn396Ser* with depressive symptoms was relatively decreased after adjusting for total cholesterol (adjusted β = *7.0,* adjusted *p value* = $5.3*10^{-06}$), using a slightly reduced sample size (*N* = 1 112). Furthermore, we observed a two-fold increase in the prevalence of Alzheimer's disease (AD) in carriers of

the minor allele (14/60 carriers) compared to non-carriers (*p value* = $2.8*10^{-02}$; Table 3). The association of *Asn396Ser* with depressive symptoms remained significant (*p value* = $3.7*10^{-08}$) after adjusting for AD. However, the association with AD disappeared (*p value* = 0.46) after adjusting for depressive symptoms. Metabolomic profiles of carriers and non-carriers of *Asn396Ser* in the ERF study demonstrated significantly increased levels of high density lipoprotein (HDL) sub-particles and Apolipoprotein A1 (APOA1) (*p value* < $6.7*10^{-05}$; Supplementary Table 2).

Table 3. Association results of various diseases with *Asn396Ser*

Disease	Phenotype	Effect (OR)§	SE	*p* value
Cardiovascular				
	Myocardial infarction (MI)	-0.32	0.48	0.51
	Coronary heart disease (CHD)	-0.18	0.49	0.72
	Coronary artery calcification (CAC)	-0.27	0.58	0.64
Inflammation				
	C-reactive protein (CRP)	-0.06	0.13	0.65
Thyroid function				
	Thyroid stimulating hormone (TSH)	0.05	0.17	0.79
	Free Thyroxine (FT4; pmol/l)	-0.72	0.49	0.14
Cholesterol				
	High density lipoprotein (HDL; mg/dl)	6.96	2.94	**0.02**
	Low density lipoprotein (LDL; mg/dl)	15.16	7.42	**0.04**
	Total cholesterol (TC; mg/dl)	21.09	7.57	**$5.0*10^{-03}$**
	Triglycerides (TG; mg/dl)	-0.05	0.10	0.59
Atherosclerosis				
	Intima media thickness (IMT)	-0.03	0.03	0.31
Hypertension				
	Diastolic blood pressure	0.20	2.29	0.93
	Systolic blood pressure	0.01	4.21	0.99
	Hypertension	0.27(1.31)	0.52	0.60
Neurological				
	Alzheimer's disease (AD)	0.70(2.01)	0.32	**0.03**
	Cognitive function (G-factor)	-0.10	0.28	0.72

§ all effects/ORs are reported for the minor allele (G)

DISCUSSION

Using a comprehensive exome-based association study of depressive symptoms, we have identified a rare (MAF = 1%) *Asn396Ser* coding variant (rs77960347-G) in the *LIPG* gene in chromosome 18q21. This variant is conserved and predicted to have damaging effects on the LIPG protein. The estimated effect of *Asn396Ser* on depressive symptoms is large and significantly replicates in three independent samples. Carriers of the rare allele (G) manifest smaller brains and more white-matter lesions compared to non-carriers. When considering pleiotropic effects, the variant is nominally associated with blood cholesterol levels and an increased risk of AD.

Rs77960347 is relatively less frequent in the general population (1000Genomes MAF = 0.5%; 10/2174 alleles) compared to the cohorts used in this study. Comparing across ethnicities, it is more frequent in non-Finnish Europeans (1.35% in non-Finnish Europeans versus 0.35% in Finnish and 0.18% in Africans). In the RS cohort, we observed 60 carriers of the minor allele (G). The variant overlaps both transcripts of the endothelial lipase (*LIPG*) gene. Common variants in *LIPG* have consistently been associated with blood cholesterol levels.[52] Previous reports on the association of cholesterol levels with depression have been contradictory.[53] In our study, we observed significantly increased levels of serum cholesterol in carriers of *Asn396Ser* compared to non-carriers. Notably however, the association with cholesterol was not nearly as strong as with the depressive symptoms. Furthermore, the lipidomics/metabolomics profiles of the carriers and non-carriers in the ERF cohort demonstrated significantly increased levels of HDL sub-particles and plasma APOA1.

Carriers of rs77960347-G had a higher burden of white matter lesions than non-carriers. White matter lesions in patients with depression have been well recognized in the elderly.[54-57] The severity of subcortical white matter lesions in the elderly has been associated with the presence of depressive symptoms and to a history of late-onset depression.[57] White matter lesions are also frequently found in AD patients.[58-60] In this context, the higher prevalence of AD in carriers of *Asn396Ser* is interesting. Depression is also one of the most widely studied risk factors for AD.[61] Previous reports suggest that the one-month prevalence of major depressive episodes in AD patients is as high as 25%.[61] While there is a general consensus about the existing complex relationship between the two diseases, the direction of causality remains unclear. However, a few longitudinal epidemiological studies suggest that people with a history of depression are at an increased risk of AD.[62,63] It is of interest that the association of genetic variant loses its significance after adjusting from depressive symptoms but its association to depressive symptoms remains significant when adjusting for AD. This findings is compatible with a mechanism in which the depressive symptoms are an intermediate in the association between *Asn396Ser* and AD. Given the marginal level of significance of the association to AD, this hypothesis requires further confirmation in other cohort studies.

LIPG is highly expressed in the thyroid gland (AUC = 0.96, *p value* = $1*10^{-49}$) and is predicted to be involved in the metabolism of steroids, cholesterol, thyroid hormone, tryptophan, butanoate (butyrate), vitamins A and D, and the platelet amyloid precursor protein (APP) pathway (http://129.125.135.180:8080/GeneNetwork/?gene=LIPG).[64] Hypothalamus-pituitary-thyroid (HPT) axis abnormalities have been documented in some depressed patients.[65] Thyroid dysfunction, including both hypothyroidism and hyperthyroidism, are known to be associated with mood disorders.[66-68] In our study we did not find evidence of thyroid function abnormalities in the carriers of the minor allele. The predicted involvement of *LIPG* in the metabolism of

butanoate (butyrate) and tryptophan is notable given that tryptophan is a precursor of serotonin and depletion in plasma tryptophan has been previously reported in depressed individuals.[69,70] Sodium butyrate is also known to reverse depressive-like behaviors in animals.[71]

Our study is the first to have used exome-sequencing on a large scale with a specific aim to identify rare variants with large effects on depressive symptoms. One of the strengths of this study is the use of a quantitative depression scale for gene discovery. The CES-D detects cases of MDD with high sensitivity and specificity[28] and has proven to be relatively stable over time.[25,72] In association studies quantitative endophenotypes provide an improved power to detect genetic variants compared to a dichotomous trait. This is likely to be especially true for trait such as MDD, for which the severity and duration of illness can be highly heterogeneous.[73] Another strength of our study is the use of high coverage sequencing (> 50 x on average) necessary to reliably detect rare variants.

In conclusion, we have discovered a novel missense variant rs77960347-G (*Asn396Ser*) in the *LIPG* gene that confers a large effect on depressive symptoms in the general population. We have also confirmed that this association is not confounded or mediated by the presence of an underlying somatic disorder. Moreover, the combination of the predicted functioning of *LIPG*, the association of rs77960347-G with white matter lesions, and its pleiotropic effects on various phenotypes linked with major depressive disorder substantiates *LIPG* as a candidate gene for depression that warrants further genetic and functional investigation.

REFERENCES

1. Lohoff FW. Overview of the genetics of major depressive disorder. *Curr Psychiatry Rep.* 2010;12(6):539-546.
2. Visscher PM, Brown MA, McCarthy MI, Yang J. Five years of GWAS discovery. *Am J Hum Genet.* 2012;90(1):7-24.
3. Major Depressive Disorder Working Group of the Psychiatric GC, Ripke S, Wray NR, et al. A mega-analysis of genome-wide association studies for major depressive disorder. *Mol Psychiatry.* 2013;18(4):497-511.
4. CONVERGE. consortium. Sparse whole-genome sequencing identifies two loci for major depressive disorder. *Nature.* 2015;523:588–591.
5. Sullivan PF. Genetics of disease: Associations with depression. *Nature.* 2015;523(7562):539–540.
6. Levinson DF, Mostafavi S, Milaneschi Y, et al. Genetic studies of major depressive disorder: why are there no genome-wide association study findings and what can we do about it? *Biol Psychiatry.* 2014;76(7):510-512.
7. Schizophrenia Working Group of the Psychiatric Genomics C. Biological insights from 108 schizophrenia-associated genetic loci. *Nature.* 2014;511(7510):421-427.
8. Demirkan A, Lahti J, Direk N, et al. Somatic, positive and negative domains of the CES-D scale: a meta-analysis of genome-wide association studies. *Psychol Med.* 2015;accepted.
9. Hek K, Demirkan A, Lahti J, et al. A genome-wide association study of depressive symptoms. *Biol Psychiatry.* 2013;73(7):667-678.
10. Gottesman, II, Gould TD. The endophenotype concept in psychiatry: etymology and strategic intentions. *Am J Psychiatry.* 2003;160(4):636-645.
11. Hall MH, Smoller JW. A new role for endophenotypes in the GWAS era: functional characterization of risk variants. *Harv Rev Psychiatry.* 2010;18(1):67-74.
12. International Consortium for Blood Pressure Genome-Wide Association S, Ehret GB, Munroe PB, et al. Genetic variants in novel pathways influence blood pressure and cardiovascular disease risk. *Nature.* 2011;478(7367):103-109.
13. Cross-Disorder Group of the Psychiatric Genomics C, Lee SH, Ripke S, et al. Genetic relationship between five psychiatric disorders estimated from genome-wide SNPs. *Nat Genet.* 2013;45(9):984-994.
14. Ott J, Wang J, Leal SM. Genetic linkage analysis in the age of whole-genome sequencing. *Nat Rev Genet.* 2015;16(5):275-284.
15. Holmans P, Zubenko GS, Crowe RR, et al. Genomewide significant linkage to recurrent, early-onset major depressive disorder on chromosome 15q. *Am J Hum Genet.* 2004;74(6):1154-1167.
16. Levinson DF, Evgrafov OV, Knowles JA, et al. Genetics of recurrent early-onset major depression (GenRED): significant linkage on chromosome 15q25-q26 after fine mapping with single nucleotide polymorphism markers. *Am J Psychiatry.* 2007;164(2):259-264.
17. Holmans P, Weissman MM, Zubenko GS, et al. Genetics of recurrent early-onset major depression (GenRED): final genome scan report. *Am J Psychiatry.* 2007;164(2):248-258.
18. Camp NJ, Lowry MR, Richards RL, et al. Genome-wide linkage analyses of extended Utah pedigrees identifies loci that influence recurrent, early-onset major depression and anxiety disorders. *Am J Med Genet B Neuropsychiatr Genet.* 2005;135B(1):85-93.
19. Abkevich V, Camp NJ, Hensel CH, et al. Predisposition locus for major depression at chromosome 12q22-12q23.2. *Am J Hum Genet.* 2003;73(6):1271-1281.
20. McGuffin P, Knight J, Breen G, et al. Whole genome linkage scan of recurrent depressive disorder from the depression network study. *Hum Mol Genet.* 2005;14(22):3337-3345.

21. McCarthy S, Das S, Kretzschmar W, Durbin R, Abecasis G, Marchini J. A reference panel of 64,976 haplotypes for genotype imputation. *bioRxiv.* 2015.
22. Radloff LS. The CES-D scale: a self report depression scale for research in the general population. *Appl Pshycol Measurement.* 1977(3):385-401.
23. Hofman A, Darwish Murad S, van Duijn CM, et al. The Rotterdam Study: 2014 objectives and design update. *Eur J Epidemiol.* 2013;28(11):889-926.
24. Harlow SD, Goldberg EL, Comstock GW. A longitudinal study of the prevalence of depressive symptomatology in elderly widowed and married women. *Arch Gen Psychiatry.* 1991;48(12):1065-1068.
25. Kuchibhatla MN, Fillenbaum GG, Hybels CF, Blazer DG. Trajectory classes of depressive symptoms in a community sample of older adults. *Acta Psychiatr Scand.* 2012;125(6):492-501.
26. Lopez-Leon S, Choy WC, Aulchenko YS, et al. Genetic factors influence the clustering of depression among individuals with lower socioeconomic status. *PLoS One.* 2009;4(3):e5069.
27. Lyness JM, Noel TK, Cox C, King DA, Conwell Y, Caine ED. Screening for depression in elderly primary care patients. A comparison of the Center for Epidemiologic Studies-Depression Scale and the Geriatric Depression Scale. *Arch Intern Med.* 1997;157(4):449-454.
28. Beekman AT, Deeg DJ, Van Limbeek J, Braam AW, De Vries MZ, Van Tilburg W. Criterion validity of the Center for Epidemiologic Studies Depression scale (CES-D): results from a community-based sample of older subjects in The Netherlands. *Psychol Med.* 1997;27(1):231-235.
29. Radloff LS. The CED-D scale: A self-report depression scale for research in the general population. *Applied Psychological Measurement.* 1977;1:385-401.
30. de Boer R, Vrooman HA, van der Lijn F, et al. White matter lesion extension to automatic brain tissue segmentation on MRI. *Neuroimage.* 2009;45(4):1151-1161.
31. Vrooman HA, Cocosco CA, van der Lijn F, et al. Multi-spectral brain tissue segmentation using automatically trained k-Nearest-Neighbor classification. *Neuroimage.* 2007;37(1):71-81.
32. Lambert JC, Ibrahim-Verbaas CA, Harold D, et al. Meta-analysis of 74,046 individuals identifies 11 new susceptibility loci for Alzheimer's disease. *Nat Genet.* 2013;45(12):1452-1458.
33. Chaker L, Heeringa J, Dehghan A, et al. Normal Thyroid Function and the Risk of Atrial Fibrillation: the Rotterdam Study. *J Clin Endocrinol Metab.* 2015;100(10):3718-3724.
34. Pardo LM, MacKay I, Oostra B, van Duijn CM, Aulchenko YS. The effect of genetic drift in a young genetically isolated population. *Ann Hum Genet.* 2005;69(Pt 3):288-295.
35. Bjelland I, Dahl AA, Haug TT, Neckelmann D. The validity of the Hospital Anxiety and Depression Scale. An updated literature review. *J Psychosom Res.* 2002;52(2):69-77.
36. Draisma HH, Pool R, Kobl M, et al. Genome-wide association study identifies novel genetic variants contributing to variation in blood metabolite levels. *Nat Commun.* 2015;6:7208.
37. Gonzalez-Covarrubias V, Beekman M, Uh HW, et al. Lipidomics of familial longevity. *Aging Cell.* 2013;12(3):426-434.
38. Demirkan A, van Duijn CM, Ugocsai P, et al. Genome-wide association study identifies novel loci associated with circulating phospho- and sphingolipid concentrations. *PLoS Genet.* 2012;8(2):e1002490.
39. Demirkan A, Henneman P, Verhoeven A, et al. Insight in genome-wide association of metabolite quantitative traits by exome sequence analyses. *PLoS Genet.* 2015;11(1):e1004835.
40. Li H, Durbin R. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics.* 2009;25(14):1754-1760.
41. Li H, Handsaker B, Wysoker A, et al. The Sequence Alignment/Map format and SAMtools. *Bioinformatics.* 2009;25(16):2078-2079.
42. McKenna A, Hanna M, Banks E, et al. The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res.* 2010;20(9):1297-1303.

43. Danecek P, Auton A, Abecasis G, et al. The variant call format and VCFtools. *Bioinformatics.* 2011;27(15):2156-2158.
44. 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):559-575.
45. Brouwer RW, van den Hout MC, Grosveld FG, van Ijcken WF. NARWHAL, a primary analysis pipeline for NGS data. *Bioinformatics.* 2012;28(2):284-285.
46. Grove ML, Yu B, Cochran BJ, et al. Best practices and joint calling of the HumanExome BeadChip: the CHARGE Consortium. *PLoS One.* 2013;8(7):e68095.
47. Yasuda T, Ishida T, Rader DJ. Update on the role of endothelial lipase in high-density lipoprotein metabolism, reverse cholesterol transport, and atherosclerosis. *Circ J.* 2010;74(11):2263-2270.
48. Kluz J, Adamiec R. [The role of endothelial lipase in the pathogenesis of atherosclerosis]Rola lipazy srodblonkowej w patogenezie miazdycy. *Pol Arch Med Wewn.* 2006;115(2):148-156.
49. Huang J, Qian HY, Li ZZ, et al. Role of endothelial lipase in atherosclerosis. *Transl Res.* 2010;156(1):1-6.
50. Paradis ME, Lamarche B. Endothelial lipase: its role in cardiovascular disease. *Can J Cardiol.* 2006;22 Suppl B:31B-34B.
51. Broedl UC, Jin W, Rader DJ. Endothelial lipase: a modulator of lipoprotein metabolism upregulated by inflammation. *Trends Cardiovasc Med.* 2004;14(5):202-206.
52. Global Lipids Genetics C, Willer CJ, Schmidt EM, et al. Discovery and refinement of loci associated with lipid levels. *Nat Genet.* 2013;45(11):1274-1283.
53. van Reedt Dortland AK, Giltay EJ, van Veen T, van Pelt J, Zitman FG, Penninx BW. Associations between serum lipids and major depressive disorder: results from the Netherlands Study of Depression and Anxiety (NESDA). *J Clin Psychiatry.* 2010;71(6):729-736.
54. O'Brien JT, Ames D. White matter lesions in depression and Alzheimer's disease. *Br J Psychiatry.* 1996;169(5):671.
55. Brown FW, Lewine RJ, Hudgins PA, Risch SC. White matter hyperintensity signals in psychiatric and nonpsychiatric subjects. *Am J Psychiatry.* 1992;149(5):620-625.
56. Coffey CE, Wilkinson WE, Weiner RD, et al. Quantitative cerebral anatomy in depression. A controlled magnetic resonance imaging study. *Arch Gen Psychiatry.* 1993;50(1):7-16.
57. de Groot JC, de Leeuw FE, Oudkerk M, Hofman A, Jolles J, Breteler MM. Cerebral white matter lesions and depressive symptoms in elderly adults. *Arch Gen Psychiatry.* 2000;57(11):1071-1076.
58. Prins ND, van Dijk EJ, den Heijer T, et al. Cerebral white matter lesions and the risk of dementia. *Arch Neurol.* 2004;61(10):1531-1534.
59. Fazekas F, Chawluk JB, Alavi A, Hurtig HI, Zimmerman RA. MR signal abnormalities at 1.5 T in Alzheimer's dementia and normal aging. *AJR Am J Roentgenol.* 1987;149(2):351-356.
60. Yoshita M, Fletcher E, Harvey D, et al. Extent and distribution of white matter hyperintensities in normal aging, MCI, and AD. *Neurology.* 2006;67(12):2192-2198.
61. Tsuno N, Homma A. What is the association between depression and Alzheimer's disease? *Expert Rev Neurother.* 2009;9(11):1667-1676.
62. Andersen K, Lolk A, Kragh-Sorensen P, Petersen NE, Green A. Depression and the risk of Alzheimer disease. *Epidemiology.* 2005;16(2):233-238.
63. Dal Forno G, Palermo MT, Donohue JE, Karagiozis H, Zonderman AB, Kawas CH. Depressive symptoms, sex, and risk for Alzheimer's disease. *Ann Neurol.* 2005;57(3):381-387.
64. Fehrmann RS, Karjalainen JM, Krajewska M, et al. Gene expression analysis identifies global gene dosage sensitivity in cancer. *Nat Genet.* 2015;47(2):115-125.
65. Bahls SC, de Carvalho GA. [The relation between thyroid function and depression: a review. A relacao entre a funcao tireoidiana e a depressao: uma revisao. *Rev Bras Psiquiatr.* 2004;26(1):41-49.

66. Bauer M, Goetz T, Glenn T, Whybrow PC. The thyroid-brain interaction in thyroid disorders and mood disorders. *J Neuroendocrinol.* 2008;20(10):1101-1114.
67. Joffe RT, Sokolov ST. Thyroid hormone treatment of primary unipolar depression: a review. *Int J Neuropsychopharmacol.* 2000;3(2):143-147.
68. Medici M, Direk N, Visser WE, et al. Thyroid function within the normal range and the risk of depression: a population-based cohort study. *J Clin Endocrinol Metab.* 2014;99(4):1213-1219.
69. Ogawa S, Fujii T, Koga N, et al. Plasma L-tryptophan concentration in major depressive disorder: new data and meta-analysis. *J Clin Psychiatry.* 2014;75(9):e906-915.
70. Lindseth G, Helland B, Caspers J. The effects of dietary tryptophan on affective disorders. *Arch Psychiatr Nurs.* 2015;29(2):102-107.
71. Resende WR, Valvassori SS, Reus GZ, et al. Effects of sodium butyrate in animal models of mania and depression: implications as a new mood stabilizer. *Behav Pharmacol.* 2013.
72. Harlow SD, Goldberg EL, Comstock GW. A longitudinal study of risk factors for depressive symptomatology in elderly widowed and married women. *Am J Epidemiol.* 1991;134(5):526-538.
73. van der Sluis S, Posthuma D, Nivard MG, Verhage M, Dolan CV. Power in GWAS: lifting the curse of the clinical cut-off. *Mol Psychiatry.* 2013;18(1):2-

CHAPTER 3.2

Non-synonymous variation in *NKPD1* increases depressive symptoms

Amin Najaf, Belonogova Nadezhda, Jovanova Olivera, Brouwer Rutger, van Rooij et al.

Biological Psychiatry. 2017 Apr; 81(8):702-707.

ABSTRACT

Background: Despite high heritability, little success was achieved in mapping genetic determinants of depression-related traits by means of genome-wide association studies (GWAS).

Methods: To identify genes associated with depressive symptomology, we performed a gene-based association analysis of non-synonymous variation captured using exome-sequencing and exome-chip genotyping in a genetically isolated population from the Netherlands (*N* = 1 999). Finally, we reproduced our significant findings in an independent population-based cohort (*N* = 1 604).

Results: We detected significant association of depressive symptoms with a gene *NKPD1* (*p value* = 3.7×10^{-08}). Non-synonymous variants in the gene explained 0.9% of sex-age-adjusted variance of depressive symptoms in the discovery study, which is translated into 3.8% of the total estimated heritability (h^2 = 0.24). Significant association of depressive symptoms with *NKPD1* gene was also observed (*N* = 1 604, *p value* = 1.5×10^{-03}) in the independent replication sample despite little overlap with the discovery cohort in the set of non-synonymous genetic variants observed in the *NKPD1* gene. Meta-analysis of the discovery and replication studies improved the association signal (*p value* = 1.0×10^{-09}).

Conclusion: Our study suggests that non-synonymous variation in the gene *NKPD1* affects depressive symptoms in the general population. *NKPD1* is predicted to be involved in the *de-novo* synthesis of sphingolipids, which have been implicated in the pathogenesis of depression.

INTRODUCTION

Major depressive disorder (MDD) is the most prevalent psychiatric disease[1] and is known to be influenced by both genetic and non-genetic factors. The heritability of MDD is estimated to be ~ 40%.[1-3] Up until recently all efforts to identify common genetic variation involved in depression using large scale genome-wide association studies (GWAS) were unsuccessful.[4,5] Two common variants on chromosome 10 were detected, one near the *SIRT1-DNAJC12* locus (minor allele frequency (MAF) = 0.45; *p value* = 2.4×10^{-10}), the other in an intron of the *LHPP* gene (MAF = 0.26; *p value* = 6.4×10^{-12}), each with a small effect on the disease risk (odds ratios of 1.15 and 0.84). Despite successful replication in an independent sample of Chinese origin, the two variants could not be reproduced in the European MDD case/control samples.[6,7]

The limited success of GWAS in case of depression has been attributed to the specifics of the disease.[1] The patients with MDD have heterogeneous phenotypes, they differ from one another in frequency, severity and the duration of episodes of MDD.[5] Those with a family history of MDD more often have clinically severe illness, earlier age of onset, and suffer higher rates of recurrence.[8] Because MDD is more prevalent and less heritable than other psychiatric traits, for instance, schizophrenia[1], it has been speculated that common genetic variants involved in MDD have effects even smaller than the effects of genetic variants observed for schizophrenia.[1,9] The proportion of variance of MDD explained by common genetic variants is estimated to be 21%[10] asking for a sample size of more than 75K for the discovery of such variants.[1]

One of the approaches that was successfully followed for other complex and prevalent outcomes that are liable to misclassification (e.g., hypertension) is to study the genetics of quantitative traits of liability (e.g. blood pressure).[11] Following this reasoning, large-scale attempts were made to identify common genetic variants associated with depressive symptoms and somatic complaints in depression ($N > 30\,000$). However, these efforts also failed.[2,12] Of note is that, the percentage of variance explained by common genetic variation was estimated to be rather low (< 4%).[12]

Besides common variants, rare variants may be involved in MDD. Recently, two studies identified rare variants in a gene *ZNF34*[13] and Cav2-adaptor gene set[14] to be nominally associated with depression. As these variants may be private to individual (sub-) families even at a single locus, different variants in the same gene may be involved in the disease (allelic heterogeneity). Statistical approaches to aggregate rare variants across genes can help overcome the problem of allelic heterogeneity.[15] Further increase in power of statistical association analysis may be achieved by using founder populations for gene discovery.[15] Because of their unique characteristics including genetic, environmental and cultural homogeneity, and enrichment of some rare alleles due to genetic drift, founder populations are considered more powerful than general populations for the discovery of rare variants.[15,16]

In the current study we aim to identify genetic variants implicated in depressive symptomology using exome-sequencing and exome-chip genotyping in a genetically isolated European population. We conduct single variant and gene-based analyses to discover variants and genes associated with depressive symptoms and replicate our findings into a population-based prospective cohort study from the same source population.

METHODS

Discovery population

The Erasmus Rucphen Family (ERF) study is a cross-sectional cohort including 3 000 living descendants (age range 18 - 96 years) of 22 couples who lived in the middle of 18th century in an isolated village in the Southwest of the Netherlands and had at least 6 children baptized in the community church. Until the last few decades descendants of these founders have lived in social isolation with minimal immigration (less than 5%). From the year 1848, the population has expanded from 700 up to 20 000 inhabitants.[17] 77% of the fathers and 79% of the mothers in this population have inbreeding coefficient greater than zero. The participants are not selected for any disease or outcome. Detailed information regarding the ERF isolate can be found elsewhere.[17-19] The study protocol was approved by the medical ethics board of the Erasmus MC Rotterdam, the Netherlands. Written informed consents were provided by all the subjects participating in the study. Symptoms of depression were assessed with the Center for Epidemiological Studies Depression Scale (CES-D) in 2,353 participants.[20] The questionnaire is a valid and reliable self-report measure of current symptoms of depression.[20,21] The scale consists of 20 items with total scores ranging from 0 to 60. Higher scores correspond to more depressive symptoms. A CES-D score of >16 is indicative of possible depression.[20] The descriptive data of ERF study are provided in Supplementary Table 1.

Exome sequencing

Exomes of 1 336 participants from the ERF pedigree were sequenced at the Center for Biomics, Erasmus MC Rotterdam, the Netherlands. Sequencing was performed at a median depth of 57 x (mean depth = 74x) using Agilent SureSelect V4 capture kit on an Illumina Hiseq2000 sequencer and the TruSeq Version 3 protocol. Sequence reads were aligned to the human genome build 19, using Burrows-Wheeler Aligner[22] and the NARWHAL pipeline.[23] Subsequently, the aligned reads were processed further, using the IndelRealigner, MarkDuplicates and Table Recalibration tools from the Genome Analysis Toolkit (GATK). For each sample, at least 4 Gigabases of sequence was aligned to the genome. Genetic variants were called using the Unified Genotyper tool from

the GATK[24] v2.3. About 1.4 million Single Nucleotide Variants (SNVs) were called. Genotype concordance checks were performed with the available genome-wide array data. The SeattleSeq annotation 138 database (http://snp.gs.washington.edu/SeattleSeqAnnotation138/) was used to perform functional annotations of the identified SNVs. After removal of individuals with low concordance, inconsistent genomic and pedigree kinship estimate and ethnic-outliers detected using principal components analysis (Supplementary Figure 1), and filtering out variants with low quality (QUAL < 30), > 5% Mendelian errors per sample and out of Hardy-Weinberg equilibrium (p *value* < 10^{-08}) we retrieved 528 617 polymorphic SNVs in 1 308 individuals. Of these, 1 209 had depressive symptoms measured.

Exome chip

Study participants from the ERF cohort whose exomes were not sequenced (N = 1 527) were genotyped on the Illumina Infinium HumanExome BeadChip, version 1.1, which contains over 240 000 exonic variants selected from multiple sources together spanning 12 000 samples from multiple ethnicities. Calling was performed with GenomeStudio and the ZCall variant calling tool[25](Broad Institute). We removed subjects with a call rate < 0.95, IBS > 0.99, heterozygote ratio > 0.60, XXY individuals and ethnic-outliers (Supplementary Figure 1). SNPs that were monomorphic in our sample or had a call rate < 0.95 were removed from the dataset. After quality control we retrieved 71 320 genetic variants in 1 512 individuals. Of these, 794 had depressive symptoms measured and 62 246 variants were polymorphic within this sample.

DATA ANALYSIS

Single variant association analysis of depressive symptoms was performed using a linear-mixed model (*mmscore* function) in GenABEL (30) incorporating genomic kinship matrix to account for relatedness. Gene-based association analysis of depressive symptoms including only nonsynonymous variants (missense, missense near splice, stop-gained, stop-gained near splice, stop-lost, stop-lost near splice mutations) was performed using the famSKAT-O test in the FFBSKAT function of the FREGAT R-package[26] (http://mga.bionet.nsc.ru/soft/FREGAT/). FamSKAT-O is a powerful tool for family-based analyses.[27] It combines a sequence kernel association test (SKAT) with a classical burden test and retains the advantages of both. SKAT compares the average similarity of a set of genetic variants in the analyzed region for each pair of individuals with pairwise phenotypic similarities. Pairwise genetic similarity is measured by using a kernel weighting function, which reduces the information on multiple genetic variants for a pair of individuals into a single scalar factor. Weights for genetic variants were assigned as a standard function of the MAF with β distribution parameters of (1, 25)[28] and the range for ρ was

set from 0 to 1 with an increment of 0.1.[29] Genomic kinship was used to correct for relatedness. The kinship matrix was calculated with the IBS function of GenABEL [30] using autosomal genetic variants with a MAF ≥ 5%. Since the two ERF subsamples had little overlapping in individuals (4 individuals in common) they were analyzed separately. Of 88 565 nonsynonymous variants in exome sequence data, 85 858 formed 13 707 gene regions with > 1 polymorphic genetic variants (ranging from 2 to 222 variants, median 5). In exome chip data, 47 469 nonsynonymous variants were present and 43 563 of them were within 9 677 gene regions with > 1 polymorphic genetic variants (ranging from 2 to 123 variants, median 3). Meta-analysis of the two ERF subsamples (as well as of the discovery and replication samples) was performed using Fisher's method [31], which has been shown to have a high power in regional meta-analysis.[32] All analyses were adjusted for age and sex (Supplementary Table 2). Exome-wide significance for gene-based analysis was set at *p value* < 2.5×10^{-6} corresponding to a Bonferroni correction for ~20 000 protein-coding genes in the genome.[33] The proportion of total variance explained by a single region was estimated with the 'return.variance.explained' option of FFBSKAT.

Replication study

Significant findings from the discovery population were selected for replication in the independent population-based Rotterdam Study (RS) study. RS is a prospective, population-study from a well-defined district in Rotterdam (Ommoord) that investigates the occurrence and determinants of diseases in the elderly.[34] The cohort was initially defined in 1990 among approximately 7 900 persons who underwent a home interview and extensive physical examination at the baseline and during follow-up rounds every 3-4 years. The cohort was extended in 2000 and 2005. RS is an outbred population, predominantly of Dutch origin. The Medical Ethics Committee of the Erasmus Medical Center, Rotterdam, approved the study. Written informed consent was obtained from all study participants. Participants from RS-I are assessed at four follow-up visits for depressive symptoms using the 20 item version of the CES-D scale. Because of maximum participation, data from the third visit was used in this study.

Exomes of 2 628 individuals from the RS-I population were sequenced at the Human Genotyping facility of the Internal Medicine department, at the Erasmus MC, the Netherlands, at an average depth of 54x using the Nimblegen SeqCap EZ V2 capture kit on an Illumina Hiseq2000 sequencer and the TruSeq Version 3 protocol. The sequenced reads were aligned to human genome build 19 using Burrows-Wheeler Aligner.[22] The aligned reads were processed further using Picard's MarkDuplicates, SAMtools[35], and the Indel Realignment and Base Quality Score Recalibration tools from GATK.[24] Genetic variants were called using the HaplotypeCaller and genotypeGVCFs tools of GATK. Samples with low concordance to genotyping array (< 95%), or that were 4 standard deviations from mean on either the number of detected variants per sample, transition to

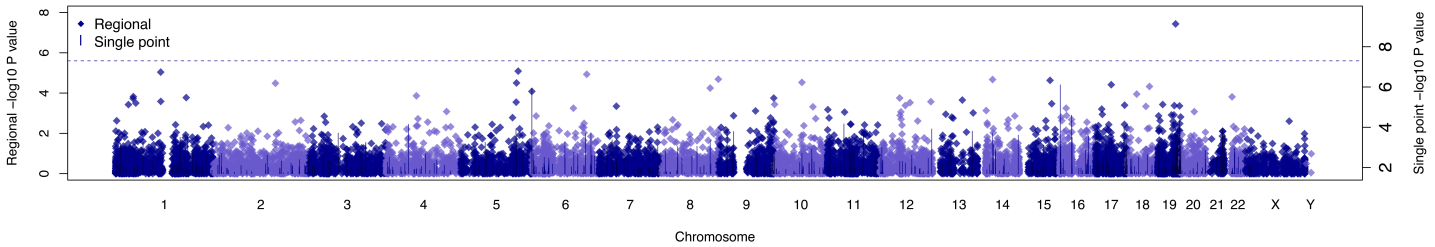
transversion ratio or heterozygous to homozygous ratio and low call rate (< 90%) were removed from the data. Initial variant QC was done using the Variant Quality Score Recalibration tool from GATK, additionally SNVs with a low call rate (< 90%) and out of HWE (*p value* < 10^{-08}) were also removed from the data. The final dataset consisted of 600,806 SNVs in 2,356 individuals. File handling and formatting was done using vcftools and PLINK[36] (http://pngu.mgh.harvard.edu/purcell/plink/). Annotation of the variants was performed using SeattleSeq annotation 138.

The SeqMeta library (https://cran.r-project.org/web/packages/seqMeta/index.html) for the R statistical analysis software was used to perform gene-based SKAT analysis of depressive symptoms for the genes identified in the discovery cohort. As in the discovery cohort only nonsynonymous mutations were used in the SKAT analysis. The analysis was adjusted for age and sex.

RESULTS

The discovery sample (ERF) included 1 999 individuals. The percentage of women (57%) and the mean age (49 years) were similar to that of the whole ERF study (Supplementary Table 1). The heritability of depressive symptoms was estimated at 0.24. Results of the single variant and gene-based analysis are illustrated in Figure 1 and quantile-quantile (QQ) plots are provided in Supplementary Figure 2. None of the single variants surpassed the standard genome-wide significance threshold (Figure 1). Genome-wide results of gene-based analysis are provided in Supplementary Table 3. Significant association of depressive symptoms was observed with the gene *NKPD1* (*p value* = 3.7 x 10^{-08}) (Supplementary Table 3). Nine non-synonymous variants were observed in the *NKPD1* gene in the ERF study (Supplementary Table 4). Except for one variant, minor alleles of all non-synonymous variants were associated with higher CES-D scores (Supplementary Table 4). Two variants were relatively frequent (rs28469095, MAF = 0.11 and rs117934605, MAF = 0.13). The variant that showed the lowest *p* value (rs75291769, *p value* = 2.5 x 10^{-04}) was four times more frequent in ERF (MAF = 0.046) compared to 1 000 genomes (0.012). The region covering the *NKPD1* showed good coverage in sequencing with a mean read depth of 46.5x. Non-synonymous variants in *NKPD1* gene explained 0.9% of the age- and sex-adjusted variance of CES-D in the ERF sample, which is translated into 3.8 % of the estimated heritability (0.24).

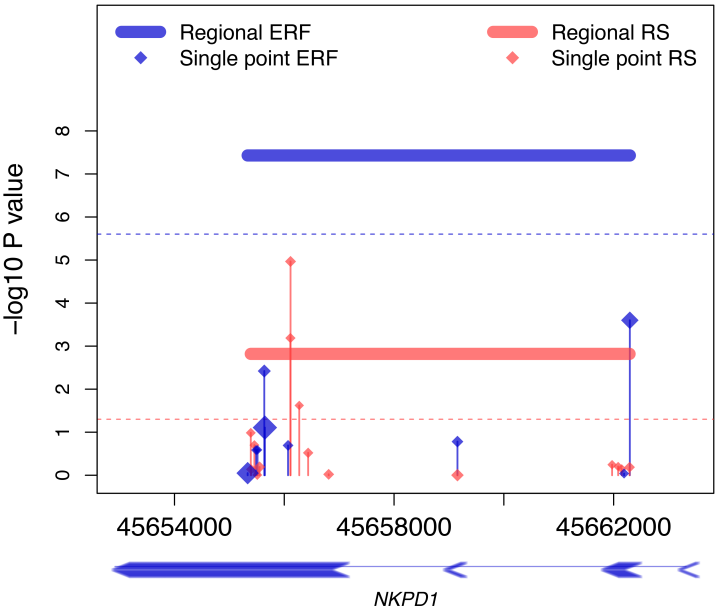
Figure 1. Results of exome-wide association analysis of depressive symptoms.



Dotted line indicates the exome-wide significance level. Horizontal axis represents the chromosome and vertical axis represents the strength of association as the negative principal logarithm of the association *p-value*. Each dot represents a gene, vertical lines represent individual *p-values* for single variants with MAF ≥ 1%.

Significant association of *NKPD1* with depressive symptoms was also observed in the replication cohort (RS) (*p value* = 1.5 x 10^{-03}). A total of 15 nonsynonymous variants were tested in the *NKPD1* gene in the RS with MAFs varying from 0.03% to 0.7%. (Supplementary Table 5). The association signal improved to *p value* = 2.9 x 10^{-05} when we limited the analyses to mutations that were predicted damaging by PolyPhen (10 mutations (9 missense, 1 stop-gained) within *NKPD1* gene; Supplementary Table 5). Figure 2 compares the results of exome single-point and gene-based analysis for *NKPD1* in the two samples. Different variants were identified within *NKPD1* gene in ERF and RS. In ERF, the second *NKPD1* exon gave the strongest association signal while the third and fourth exons contained only modest signals. In the replication, association signals were present in the fourth exon only but the *p* values were stronger compared to that seen in ERF. When meta-analyzing the discovery and replication studies the significance of association increased to a *p value* of 1.0 x 10^{-09}.

Figure 2. Single-point and gene-based association signals of *NKPD1* gene in the discovery sample (ERF, blue) and the replication sample (RS, pink).



Dotted blue and red lines indicate the regional statistically significant levels for discovery and replication, respectively. Size of the points is proportional to the MAF.

None of the previously identified genes i.e., *SIRT-DNAJC12*, *LHPP*, *ZSNF34* or Cav2 adaptor genes showed evidence of association in our study (see Supplementary Table 3).

DISCUSSION

We have identified and replicated association of (rare) non-synonymous variants in *NKPD1* gene with depressive symptoms. These coding variants explained 0.9% of the age- and sex-adjusted variance and 3.8% of heritability of depressive symptoms in the ERF population. To our knowledge this is the first study that explores the gene-based association of non-synonymous variants for depressive symptoms using large scale whole exome-sequencing.

To date, little success was achieved in mapping genetic determinants of depression-related traits by means of GWAS and whole genome/exome sequencing.[1,2,4] We did not find evidence of association of previously reported genes including *SIRT-DNAJC12*, *LHPP*, *ZSNF34* and Cav2 adaptor gene set in our study. In previously published GWAS of MDD and depressive symptoms, only five intronic variants were mapped on to the *NKPD1* gene, of which three showed modest association signals (*p value* < 0.1; Supplementary Tables 6-7). These variants may be in low linkage disequilibrium with coding variants not typed in these studies.

Our study had several strengths that ultimately led to the discovery of a novel gene associated with depressive symptoms. In association studies, quantitative measurements give better chances to detect genetic variants compared to the dichotomous trait, especially for traits like depression, which represents an arbitrarily selected extreme of the continuum of varying severity and duration.[37] Secondly, we used high coverage exome sequencing (average depth = 74x) in a large discovery set that allowed us to determine rare variants with high precision. Finally, our discovery is based within a genetically isolated population. Samples from genetic isolates are expected to have reduced genetic variation and shifted allele frequencies due to long-time genetic drift.[19] Rare alleles can accumulate in large numbers in these populations, thus empowering the association analysis.[15] In our study, the missense variant that showed the strongest association within the *NKPD1* gene (rs75291769/exm1479956, *p value* = 2.5 x 10^{-04}) had a frequency of 4.6% in ERF (Supplementary Table 3) but was monomorphic in RS despite the fact that both populations go to the same source population. Yet, several other variants within *NKPD1* showed modest associations with depressive symptoms in RS, with the strongest signal in exon 4 of *NKPD1*, 6K apart from the strongest signal obtained in ERF. None of these single-point association signals achieved genome-wide significance, illustrating the advantage of a regional gene-based association analysis over a single point study of rare variants.

NKPD1 encodes NTPase KAP family P-loop domain–containing protein 1, and it was found to express strongly in skin and, to a lesser extent, in cerebellum, other brain tissues and thyroid (http://www.gtexportal.org/home/gene/NKPD1).[38] NTPases are nucleoside-triphosphatases that hydrolyze nucleoside triphosphates to nucleoside diphosphates. It has been shown that

membrane NTPase activity is substantially altered in Lesch–Nyhan disease fibroblasts.[39] Lesch–Nyhan syndrome (also known as juvenile gout) is caused by a defect of the purine metabolism and its symptoms include mental retardation, dystonia, spasticity, delayed motor development and a compulsive form of self-injurious behavior.[40] Functional prediction of *NKPD1* suggests that this gene is involved in the de novo biosynthesis of sphingolipids (http://129.125.135.180:8080/GeneNetwork/?gene=NKPD1). De novo synthesis of sphingolipids results in the formation of ceramides.[41-46] Sphingolipid metabolism is also proposed to be a therapeutic target for MDD.[47,48] Some antidepressants (amitriptyline and fluoxetine) are known to inhibit sphingomyelin phosphodiesterase 1 acid lysosomal (SMPD1).[49] SMPD1 is a membrane protein facing the lysosomal lumen that converts sphingomyelin into ceramide.[47,49] Earlier, we found various sphingomyelin levels in the blood to be associated with depressive symptoms.[50]

In summary, we have discovered association of non-synonymous variation in the gene *NKPD1* with depressive symptoms in the general population. The variants explain 0.9% of the age- and sex-adjusted variance of depressive symptoms, which translates into 3.8% of the estimated genetic variance of depressive symptoms. The predicted involvement of *NKPD1* in the biosynthesis of sphingolipids corroborate the possible role of *NKPD1* in depression.

REFERENCES

1. Levinson DF, Mostafavi S, Milaneschi Y, et al. Genetic studies of major depressive disorder: why are there no genome-wide association study findings and what can we do about it? *Biol Psychiatry.* 2014;76(7):510-512.
2. Hek K, Demirkan A, Lahti J, et al. A genome-wide association study of depressive symptoms. *Biol Psychiatry.* 2013;73(7):667-678.
3. Sullivan PF, Neale MC, Kendler KS. Genetic epidemiology of major depression: review and meta-analysis. *Am J Psychiatry.* 2000;157(10):1552-1562.
4. Dunn EC, Brown RC, Dai Y, et al. Genetic determinants of depression: recent findings and future directions. *Harv Rev Psychiatry.* 2015;23(1):1-18.
5. Flint J, Kendler KS. The genetics of major depression. *Neuron.* 2014;81(3):484-503.
6. consortium C. Sparse whole-genome sequencing identifies two loci for major depressive disorder. *Nature.* 2015;523(7562):588-591.
7. Sullivan PF. Genetics of disease: Associations with depression. *Nature.* 2015;523(7562):539-540.
8. Weissman MM, Wickramaratne P, Nomura Y, Warner V, Pilowsky D, Verdeli H. Offspring of depressed parents: 20 years later. *Am J Psychiatry.* 2006;163(6):1001-1008.
9. Schizophrenia Working Group of the Psychiatric Genomics C. Biological insights from 108 schizophrenia-associated genetic loci. *Nature.* 2014;511(7510):421-427.
10. Cross-Disorder Group of the Psychiatric Genomics C, Lee SH, Ripke S, et al. Genetic relationship between five psychiatric disorders estimated from genome-wide SNPs. *Nat Genet.* 2013;45(9):984-994.
11. International Consortium for Blood Pressure Genome-Wide Association S, Ehret GB, Munroe PB, et al. Genetic variants in novel pathways influence blood pressure and cardiovascular disease risk. *Nature.* 2011;478(7367):103-109.
12. Demirkan A, Lahti J, Direk N, et al. Somatic, positive and negative domains of the CES-D scale: a meta-analysis of genome-wide association studies. *Psychol Med.* 2015;accepted.
13. Subaran RL, Odgerel Z, Swaminathan R, Glatt CE, Weissman MM. Novel variants in ZNF34 and other brain-expressed transcription factors are shared among early-onset MDD relatives. *Am J Med Genet B Neuropsychiatr Genet.* 2016;171(3):333-341.
14. Pirooznia M, Wang T, Avramopoulos D, Potash JB, Zandi PP, Goes FS. High-throughput sequencing of the synaptome in major depressive disorder. *Mol Psychiatry.* 2016;21(5):650-655.
15. Hatzikotoulas K, Gilly A, Zeggini E. Using population isolates in genetic association studies. *Brief Funct Genomics.* 2014;13(5):371-377.
16. Heutink P, Oostra BA. Gene finding in genetically isolated populations. *Hum Mol Genet.* 2002;11(20):2507-2515.
17. Aulchenko YS, Heutink P, Mackay I, et al. Linkage disequilibrium in young genetically isolated Dutch population. *Eur J Hum Genet.* 2004;12(7):527-534.
18. Lopez-Leon S, Choy WC, Aulchenko YS, et al. Genetic factors influence the clustering of depression among individuals with lower socioeconomic status. *PLoS One.* 2009;4(3):e5069.
19. Pardo LM, MacKay I, Oostra B, van Duijn CM, Aulchenko YS. The effect of genetic drift in a young genetically isolated population. *Ann Hum Genet.* 2005;69(Pt 3):288-295.
20. Radloff LS. The CES-D scale: a self report depression scale for research in the general population. *Appl Pshycol Measurement* 1977(3):385-401.
21. Weissman MM, Sholomskas D, Pottenger M, Prusoff BA, Locke BZ. Assessing depressive symptoms in five psychiatric populations: a validation study. *Am J Epidemiol.* 1977;106(3):203-214.

22. Li H, Durbin R. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics.* 2009;25(14):1754-1760.
23. Brouwer RW, van den Hout MC, Grosveld FG, van Ijcken WF. NARWHAL, a primary analysis pipeline for NGS data. *Bioinformatics.* 2012;28(2):284-285.
24. McKenna A, Hanna M, Banks E, et al. The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res.* 2010;20(9):1297-1303.
25. Goldstein JI, Crenshaw A, Carey J, et al. zCall: a rare variant caller for array-based genotyping: genetics and population analysis. *Bioinformatics.* 2012;28(19):2543-2545.
26. Belonogova NM, Svishcheva GR, Axenovich TI. FREGAT: an R library for Family REGional Association Tests. *European Journal of Human Genetics.* 2015;23:473.
27. Jiang D, McPeek MS. Robust rare variant association testing for quantitative traits in samples with related individuals. *Genet Epidemiol.* 2014;38(1):10-20.
28. Wu MC, Lee S, Cai T, Li Y, Boehnke M, Lin X. Rare-variant association testing for sequencing data with the sequence kernel association test. *Am J Hum Genet.* 2011;89(1):82-93.
29. Lee S, Wu MC, Lin X. Optimal tests for rare variant effects in sequencing association studies. *Biostatistics.* 2012;13(4):762-775.
30. Aulchenko YS, Ripke S, Isaacs A, van Duijn CM. GenABEL: an R library for genome-wide association analysis. *Bioinformatics.* 2007;23(10):1294-1296.
31. Fisher RA. Combining independent tests of significance. *American Statistician.* 1948;2(5):30.
32. Zakharov S, Wang X, Liu J, Teo YY. Improving power for robust trans-ethnic meta-analysis of rare and low-frequency variants with a partitioning approach. *Eur J Hum Genet.* 2015;23(2):238-244.
33. Pennisi E. Genomics. ENCODE project writes eulogy for junk DNA. *Science.* 2012;337(6099):1159, 1161.
34. Hofman A, Brusselle GG, Darwish Murad S, et al. The Rotterdam Study: 2016 objectives and design update. *Eur J Epidemiol.* 2015;30(8):661-708.
35. Li H, Handsaker B, Wysoker A, et al. The Sequence Alignment/Map format and SAMtools. *Bioinformatics.* 2009;25(16):2078-2079.
36. 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):559-575.
37. van der Sluis S, Posthuma D, Nivard MG, Verhage M, Dolan CV. Power in GWAS: lifting the curse of the clinical cut-off. *Mol Psychiatry.* 2013;18(1):2-3.
38. Fehrmann RS, Karjalainen JM, Krajewska M, et al. Gene expression analysis identifies global gene dosage sensitivity in cancer. *Nat Genet.* 2015;47(2):115-125.
39. Pinto CS, Jinnah HA, Shirley TL, Nyhan WL, Seifert R. Altered membrane NTPase activity in Lesch-Nyhan disease fibroblasts: comparison with HPRT knockout mice and HPRT-deficient cell lines. *J Neurochem.* 2005;93(6):1579-1586.
40. Nyhan WL, Wong DF. New approaches to understanding Lesch-Nyhan disease. *N Engl J Med.* 1996;334(24):1602-1604.
41. Causeret C, Geeraert L, Van der Hoeven G, Mannaerts GP, Van Veldhoven PP. Further characterization of rat dihydroceramide desaturase: tissue distribution, subcellular localization, and substrate specificity. *Lipids.* 2000;35(10):1117-1125.
42. Merrill AH, Jr. Characterization of serine palmitoyltransferase activity in Chinese hamster ovary cells. *Biochim Biophys Acta.* 1983;754(3):284-291.
43. Merrill AH, Jr., Williams RD. Utilization of different fatty acyl-CoA thioesters by serine palmitoyltransferase from rat brain. *J Lipid Res.* 1984;25(2):185-188.
44. Pewzner-Jung Y, Ben-Dor S, Futerman AH. When do Lasses (longevity assurance genes) become CerS (ceramide synthases)?: Insights into the regulation of ceramide synthesis. *J Biol Chem.* 2006;281(35):25001-25005.

45. Pruett ST, Bushnev A, Hagedorn K, et al. Biodiversity of sphingoid bases ("sphingosines") and related amino alcohols. *J Lipid Res.* 2008;49(8):1621-1639.

46. Zitomer NC, Mitchell T, Voss KA, et al. Ceramide synthase inhibition by fumonisin B1 causes accumulation of 1-deoxysphinganine: a novel category of bioactive 1-deoxysphingoid bases and 1-deoxydihydroceramides biosynthesized by mammalian cell lines and animals. *J Biol Chem.* 2009;284(8):4786-4795.

47. Osherovich L. Depressing sphingolipids. *SciBX.* 2013;6(27).

48. Gulbins E, Palmada M, Reichel M, et al. Acid sphingomyelinase-ceramide system mediates effects of antidepressant drugs. *Nat Med.* 2013;19(7):934-938.

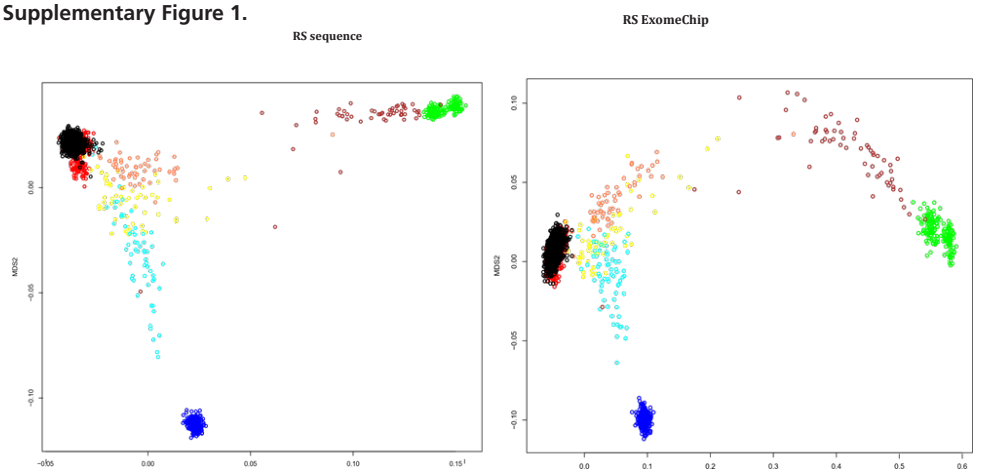
49. Kornhuber J, Tripal P, Reichel M, et al. Functional Inhibitors of Acid Sphingomyelinase (FIASMAs): a novel pharmacological group of drugs with broad clinical applications. *Cell Physiol Biochem.* 2010;26(1):9-20.

50. Demirkan A, Isaacs A, Ugocsai P, et al. Plasma phosphatidylcholine and sphingomyelin concentrations are associated with depression and anxiety symptoms in a Dutch family-based lipidomics study. *J Psychiatr Res.* 2013;47(3):357-362.

APPENDIX of CHAPTER 3

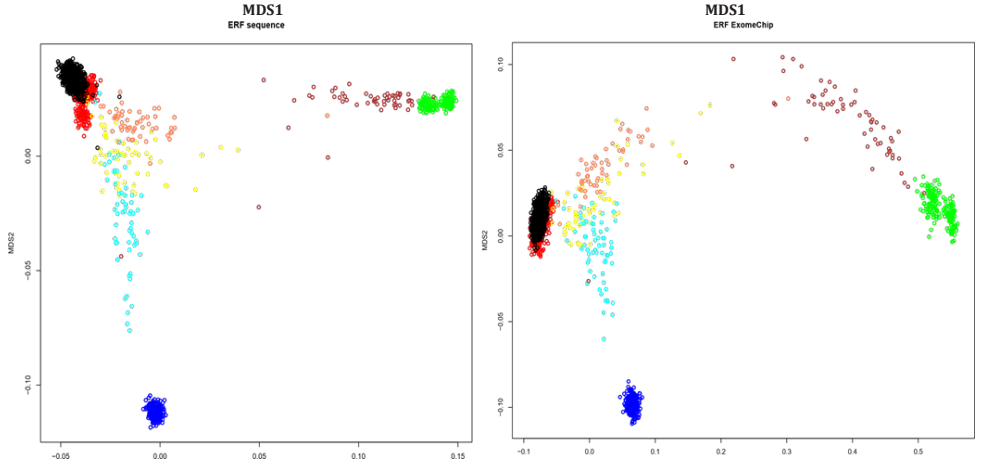
Exome-sequencing in a large population-based study reveals a rare *Asn396Ser* variant in the *LIPG* gene associated with depressive symptoms

Supplementary Figure 1.



Ethnicity information provided by the principal components analysis. Black circles represent the samples analyzed, red: Caucasian, blue: Chinese, green: African, yellow: Colombian (AMR), cyan: Mexican (AMR), coral: Puerto Rican (AMR), brown: African (AMR).

Supplementary Figure 2.



Genome-wide association plot. Horizontal axis shows all chromosomes. Vertical axis shows the strength of association in terms of negative principal logarithm of the association p-value. Each dot represents a SNV. The horizontal red dotted line indicates the genome-wide significance threshold.

Supplementary Figure 3.

Quantile-quantile plot for the genome-wide association results. Horizontal axis shows the expected chi-square distribution and vertical axis shows the observed chi-square distribution. Red line indicates the expected distribution black dotted line indicates the observed distribution.

Supplement Table 1

	RS (seqeunce)	RS (exomechip)	ERF (sequence)	ERF (exomechip)
N	1265	1525	1247	840
Age mean(*SD*)	72.3 (7.1)	73.7 (7.2)	48.5 (14.6)	50.0 (15.8)
% women	59.1	52.6	60	54
CES-D mean(*SD*)	4.5 (6.9)	4.7 (7.1)	11.7 (10.3)	9.2 (9.0)

Supplement Table 2

Name	BETA	SE	t-value	p-value	*N*
M-HDL-ApoA1	2.8811	0.6684	4.31	1.70E-05	2186
L-HDL-cholesterol	1.2789	0.3148	4.06	5.01E-05	2187
M-HDL-phospholipids	2.3569	0.5854	4.03	5.86E-05	2187
HDL-phospholipids	9.0311	2.2490	4.02	6.13E-05	2186
plasma-ApoA1	9.7158	2.4312	4.00	6.65E-05	2186
L-HDL-phospholipids	2.2543	0.5746	3.92	9.01E-05	2186
M-HDL-Free cholesterol	0.2622	0.0679	3.86	0.000116	2188
L-HDL-Free cholesterol	0.3130	0.0813	3.85	0.000121	2188
M-HDL-cholesterol	1.1438	0.3017	3.79	0.000154	2187
L-HDL-ApoA1	2.3032	0.6189	3.72	0.000203	2186
XL-HDL-Free cholesterol	0.7118	0.1937	3.67	0.000244	2177
Ethanol	-0.2681	0.0731	-3.67	0.000249	1991
HDL-cholesterol	5.0928	1.4056	3.62	0.000298	2186
L-HDL-ApoA2	0.4994	0.1394	3.58	0.000346	2186
XL-HDL-phospholipids	0.1480	0.0415	3.57	0.000369	2178
Phosphatidylcholine diacyl (38:7)	0.2853	0.0825	3.46	0.00055	2258
HDL-ApoA1	9.0328	2.6479	3.41	0.000658	2185
M-HDL-ApoA2	0.6115	0.1831	3.34	0.000852	2185
Phosphatidylcholine diacyl (36:4)	0.8600	0.2624	3.28	0.001061	2262
S-VLDL-cholesterol	0.2625	0.0830	3.16	0.001583	2187
HDL-Free cholesterol	1.1654	0.3752	3.11	0.001921	2186
S-VLDL-phospholipids	0.2874	0.0947	3.03	0.002441	2181
Ceramide C24:1	0.3180	0.1059	3.00	0.002766	648
Phosphatidylcholine diacyl (40:4)	0.0147	0.0050	2.95	0.003218	2251
Phosphatidylethanolamine 38:4	1.4875	0.5070	2.93	0.003463	647
HDL-APOA2	1.6880	0.5989	2.82	0.004871	2185
Lysophosphatidylcholine(18:2)	-0.2889	0.1049	-2.75	0.005927	2258
Phosphatidylethanolamine(38:4)	0.0361	0.0134	2.69	0.00715	2254
Phosphatidylcholine diacyl (32:0)	0.0441	0.0164	2.69	0.007191	2260
Sphingomyeline C 16:0	22.8932	8.4976	2.69	0.007241	647
Phosphatidylcholine diacyl C 38:0	1.9246	0.7309	2.63	0.008662	649
M-VLDL-cholesterol	0.9207	0.3597	2.56	0.010539	2181
Isoleucine	-8489.8937	3340.3312	-2.54	0.011108	2009
Sphingomyeline C 16:1-OH	1.0739	0.4217	2.55	0.011114	651
XL-HDL-ApoA2	0.1398	0.0552	2.53	0.01134	2172
XL-LDL-triglycerides	0.1755	0.0702	2.50	0.012455	2186
Glycine	57814.1927	23167.4814	-2.50	0.012659	1995

Name	BETA	SE	t-value	p-value	*N*
Phosphatidylcholine diacyl (34:3)	0.0700	0.0281	2.49	0.01276	2259
Phosphatidylcholine diacyl C42:4	0.0265	0.0108	2.45	0.014385	794
Phosphatidylethanolamine 38:5	0.4631	0.1887	2.45	0.014397	647
glutamine	-991.2650	409.1073	-2.42	0.015481	2015
Ornithine	-178.3774	73.9146	-2.41	0.015898	2006
L-VLDL-cholesterol	0.1080	0.0449	2.40	0.016391	2178
Lysophosphatidylcholine(22:6)	-0.0120	0.0050	-2.40	0.016637	2255
XL-HDL-cholesterol	0.1235	0.0516	2.39	0.016766	2173
Phosphatidylethanolamine 38:2	0.0397	0.0167	2.39	0.017297	647
Lysophosphatidylcholine(18:1)	-0.1436	0.0609	-2.36	0.018419	2256
Lysine	-177.0992	75.8465	-2.33	0.019643	2014
Phosphatidylcholine diacyl (34:2)	0.8670	0.3737	2.32	0.020426	2261
Phosphatidylethanolamine 36:4	0.2993	0.1336	2.24	0.02542	648
Phosphatidylethanolamine 34:2	0.2988	0.1355	2.21	0.02779	647
Phosphatidylcholine diacyl C 38:1	3.0173	1.3726	2.20	0.028281	651
M-VLDL-Free cholesterol	0.2980	0.1361	2.19	0.02863	2182
Phosphatidylcholine alkyl-acyl C42:5	0.3665	0.1674	2.19	0.028907	648
Hexadecenoylcarnitine	0.0050	0.0023	2.19	0.029	794
Phosphatidylcholine diacyl C34:3	2.3860	1.0976	2.17	0.030018	794
S-LDL-Free cholesterol	-0.3569	0.1645	-2.17	0.030139	2185
M-LDL-Free cholesterol	-0.5330	0.2460	-2.17	0.030399	2187
plasma-Free cholesterol	2.4840	1.1572	2.15	0.031931	2188
Phosphatidylcholine diacyl (34:4)	0.0055	0.0026	2.14	0.03224	2251
Phosphatidylcholine diacyl (36:3)	0.4598	0.2146	2.14	0.032254	2263
Phosphatidylethanolamine 38:6	0.3097	0.1444	2.14	0.032328	648
Phosphatidylcholine diacyl C40:5	1.5985	0.7519	2.13	0.033813	791
Lysophosphatidylcholine(O-16:1)	-0.0037	0.0018	-2.12	0.034018	2260
Phosphatidylcholine diacyl (38:4)	0.3551	0.1692	2.10	0.035946	2262
Leucine	-4708.2738	2255.4395	-2.09	0.036967	2014
S-LDL-cholesterol	-1.6397	0.7943	-2.06	0.039114	2185
Lysophosphatidylcholine(O-18:1)	-0.0029	0.0014	-2.05	0.040424	2257
Phosphatidylcholine diacyl C38:5	6.8767	3.3750	2.04	0.041927	792
XS-VLDL-phospholipids	0.0422	0.0208	2.03	0.042923	2186
Phosphatidylcholine diacyl (34:1)	0.6373	0.3152	2.02	0.043282	2260
XXL-LDL-cholesterol	163.1383	80.9920	2.01	0.044106	2182
Phosphatidylcholine diacyl (32:1)	0.1196	0.0596	2.01	0.044668	2249
Lysophosphatidylcholine(18:0)	-0.1969	0.0980	-2.01	0.044669	2262
plasma-ApoA2	1.0394	0.5181	2.01	0.04495	2186
XXL-LDL-Free cholesterol	162.9651	81.4730	2.00	0.045598	2185
Phosphatidylcholine diacyl C 34:4	-0.5293	0.2647	-2.00	0.045958	651
Phosphatidylethanolamine 34:1	0.2943	0.1474	2.00	0.046312	648
XXL-LDL-ApoB	151.1427	76.2554	1.98	0.047598	2186
Decadienylcarnitine	0.0057	0.0029	1.98	0.0477	793
Phosphatidylethanolamine(34:2)	0.1340	0.0677	1.98	0.047836	2204
Phosphatidylethanolamine alkyl-acyl 40:3	0.1621	0.0819	1.98	0.048181	648
Phosphatidylethanolamine 40:6	0.2902	0.1466	1.98	0.048245	646

Name	BETA	SE	t-value	p-value	*N*
Phosphatidylethanolamine 40:5	0.2829	0.1432	1.98	0.048645	648
IDL-ApoB	0.3683	0.1871	1.97	0.049135	2182
M-LDL-cholesterol	-1.7971	0.9265	-1.94	0.052548	2187
Phosphatidylcholine alkyl-acyl C40:5	0.5589	0.2880	1.94	0.052754	650
Sphingomyeline C 18:0	3.8163	1.9669	1.94	0.052775	651
L-VLDL-Free cholesterol	0.0783	0.0406	1.93	0.05388	2180
Phosphatidylcholine diacyl (38:2)	0.0119	0.0062	1.93	0.054025	2259
Phosphatidylcholine diacyl C32:2	0.5833	0.3035	1.92	0.05496	794
Hydroxytetradecenoylcarnitine	0.0020	0.0010	1.92	0.055369	792
XXL-LDL-triglycerides	123.5163	64.6331	1.91	0.056131	2180
IDL-Free cholesterol	0.3210	0.1681	1.91	0.056353	2187
Phosphatidylcholine diacyl C 42:5	0.1727	0.0907	1.90	0.05733	650
Phosphatidylethanolamine 36:2	0.2449	0.1288	1.90	0.057706	648
S-LDL-phospholipids	-0.8458	0.4468	-1.89	0.058478	2184
Sphingomyeline C 18:2	0.3667	0.1937	1.89	0.05877	650
Sphingomyeline C 24:1	9.8264	5.2068	1.89	0.059578	649
Phosphatidylcholine diacyl C40:6	4.0630	2.1548	1.89	0.059717	792
plasma-cholesterol	8.8493	4.7207	1.87	0.060986	2188
M-VLDL-phospholipids	0.5778	0.3095	1.87	0.062069	2183
Ceramide C23:0	0.1458	0.0783	1.86	0.063143	648
Phosphatidylcholine alkyl-acyl C42:6	0.2532	0.1368	1.85	0.064659	651
Phosphatidylcholine diacyl C36:5	0.1639	0.0890	1.84	0.065981	790
XL-HDL-triglycerides	0.1156	0.0629	1.84	0.066065	2170
XXL-LDL-phospholipids	146.5294	79.9350	1.83	0.066923	2180
Phosphatidylethanolamine 42:7	0.2279	0.1247	1.83	0.068173	645
Phosphatidylcholine diacyl C 40:1	0.2408	0.1321	1.82	0.068711	651
Phosphatidylethanolamine 36:3	0.2568	0.1410	1.82	0.068956	650
Phosphatidylethanolamine 40:4	0.0392	0.0216	1.82	0.069951	648
Lysophosphatidylcholine C 18:3	-0.1599	0.0883	-1.81	0.070663	650
Phosphatidylcholine diacyl C42:5	0.0496	0.0276	1.80	0.072586	792
Phosphatidylcholine diacyl C40:4	0.4455	0.2481	1.80	0.072998	792
Phosphatidylethanolamine 38:1	0.0364	0.0204	1.79	0.073877	649
Phosphatidylcholine diacyl (38:5)	0.1528	0.0858	1.78	0.0753	2257
Phosphatidylcholine diacyl C 34:2	66.9234	37.6360	1.78	0.075844	648
M-HDL-triglycerides	0.0839	0.0472	1.78	0.07592	2182
Pimelylcarnitine	0.0058	0.0033	1.76	0.078856	794
Hydroxysphingomyeline C22:2	-0.9502	0.5416	-1.75	0.079769	794
Phosphatidylcholine diacyl C 40:3	0.2123	0.1223	1.74	0.083094	650
L-HDL-triglycerides	0.0753	0.0435	1.73	0.083145	2175
Phosphatidylcholine diacyl C 36:4	28.5307	16.5341	1.73	0.0849	650
Lysophosphatidylcholine 16:1	-0.4205	0.2441	-1.72	0.085425	648
S-LDL-ApoB	-0.6864	0.3994	-1.72	0.085814	2186
Hydroxybutyrate_2	141.3856	82.4395	1.72	0.086496	2005
IDL-phospholipids	0.6283	0.3705	1.70	0.090089	2187
Methylglutarylcarnitine	0.0028	0.0017	1.70	0.090163	783
Valine	22356.0277	13187.6231	-1.70	0.090187	2013

Name	BETA	SE	t-value	p-value	*N*
XL-HDL-ApoA1	0.2082	0.1251	1.66	0.096193	2175
S-VLDL-Free cholesterol	0.0473	0.0285	1.66	0.09699	2187
IDL-cholesterol	1.3646	0.8228	1.66	0.097369	2183
Phosphatidylcholine diacyl (32:2)	0.0133	0.0080	1.66	0.0979	2257
Methionine	-3860.7325	2344.4098	-1.65	0.099759	2014
Lysophosphatidylcholine(20:3)	-0.0109	0.0067	-1.62	0.104468	2256
Phosphatidylcholine diacyl C42:6	0.0573	0.0356	1.61	0.107882	790
Decenoylcarnitine	0.0259	0.0161	1.61	0.107923	791
Phosphatidylethanolamine 36:1	0.2314	0.1442	1.61	0.108925	649
Phosphatidylcholine alkyl-acyl C34:0	0.3015	0.1899	1.59	0.112722	650
M-LDL-phospholipids	-0.8264	0.5275	-1.57	0.117373	2187
Creatinine	3421.9775	2224.6016	1.54	0.124147	2005
Lysophosphatidylcholine C 20:3	-0.4603	0.3033	-1.52	0.129503	647
Phosphatidylcholine alkyl-acyl C36:2	1.7172	1.1367	1.51	0.131344	648
Sphingomyeline C 16:1	2.2652	1.5040	1.51	0.132524	648
S-VLDL-triglycerides	0.0863	0.0574	1.50	0.132933	2186
lysoPhosphatidylcholine acyl C18:2	-3.3201	2.2083	-1.50	0.13312	792
Hydroxytetradecadienylcarnitine	0.0010	0.0007	1.50	0.134354	794
Alpha-ketoglutarate	140.5460	93.9052	1.50	0.134634	2007
Phosphatidylcholine diacyl C 34:3	-2.9436	1.9661	-1.50	0.134845	649
Phosphatidylethanolamine 34:3	0.2035	0.1374	1.48	0.13897	646
Phosphatidylcholine diacyl C32:0	0.8981	0.6076	1.48	0.139798	793
Phosphatidylcholine diacyl C36:6	0.1238	0.0849	1.46	0.145278	791
L-VLDL-phospholipids	0.0728	0.0501	1.45	0.146333	2182
Phosphatidylethanolamine 42:5	0.0084	0.0058	1.45	0.146905	648
Phosphatidylcholine acyl-alkyl C42:0	0.0355	0.0246	1.44	0.149189	793
Glutaconylcarnitine	0.0018	0.0012	1.44	0.150427	793
Phosphatidylethanolamine 38:3	0.1137	0.0797	1.43	0.154356	645
lysoPhosphatidylcholine acyl C18:1	-1.7766	1.2508	-1.42	0.15589	793
Creatine	-107.2443	75.7401	-1.42	0.156945	2014
Phosphatidylcholine alkyl-acyl C38:4	1.1903	0.8465	1.41	0.160173	648
Lysophosphatidylcholine C15:0	-0.1681	0.1196	-1.41	0.160322	650
Phosphatidylcholine diacyl C34:4	0.2012	0.1446	1.39	0.164438	794
Triglycerides(54:1)	-0.0162	0.0117	-1.38	0.167058	2237
lysoPhosphatidylcholine acyl C17:0	-0.1123	0.0812	-1.38	0.167282	794
PE-plasmologen 18:0/18:2	0.5911	0.4284	1.38	0.16814	650
XL-LDL-phospholipids	0.8290	0.6038	1.37	0.169927	2187
Sphingomyeline C 24:2	3.6515	2.6928	1.36	0.175559	650
Phosphatidylcholine diacyl C38:3	4.1113	3.0512	1.35	0.17823	794
Phosphatidylcholine diacyl (36:5)	0.0810	0.0604	1.34	0.179706	2242
Phosphatidylcholine diacyl C 34:1	29.0610	21.7588	1.34	0.182149	647
Phosphatidylcholine alkyl-acyl C38:5	1.4855	1.1193	1.33	0.184943	647
Sphingomyeline C 23:0	1.8779	1.4273	1.32	0.188718	651
Phosphatidylcholine diacyl C24:0	-0.0702	0.0533	-1.32	0.188782	732
Methanol	-0.0700	0.0534	-1.31	0.19043	2011
M-LDL-ApoB	-0.4892	0.3767	-1.30	0.194213	2187

Name	BETA	SE	t-value	p-value	*N*
Phosphatidylcholine alkyl-acyl C38:3	0.5011	0.3857	1.30	0.194315	650
PE-plasmologen 16:0/18:2	0.3073	0.2406	1.28	0.201867	648
Sphingomyeline C 14:0	1.4318	1.1269	1.27	0.204338	650
Phosphatidylcholine diacyl C34:1	12.5520	9.9108	1.27	0.205706	793
Carnitine	-98.4857	77.9250	-1.26	0.206429	2011
lysoPhosphatidylcholine acyl C24:0	-0.0208	0.0166	-1.25	0.210078	794
Phosphatidylcholine alkyl-acyl C34:3	1.0013	0.7990	1.25	0.210552	650
Alanine	28500.7753	22802.4272	-1.25	0.211481	2014
M-VLDL-triglycerides	0.5922	0.4742	1.25	0.21185	2182
Phosphatidylcholine diacyl C30:0	0.3928	0.3157	1.24	0.213793	793
Phosphatidylcholine alkyl-acyl C36:4	1.6566	1.3316	1.24	0.213923	648
Phosphatidylcholine diacyl C 40:2	0.1421	0.1143	1.24	0.213948	647
Ceramide C24:0	0.2999	0.2417	1.24	0.215263	648
Hexadecadienylcarnitine	0.0009	0.0007	1.23	0.21772	793
Serine	-5.2465	4.2630	-1.23	0.218797	794
Phosphatidylcholine diacyl C38:6	6.4927	5.2831	1.23	0.219451	793
Phosphatidylethanolamine 36:5	0.1543	0.1258	1.23	0.220495	646
Phosphatidylcholine diacyl C36:1	3.4804	2.8498	1.22	0.222347	792
Sphingomyeline C 23:1	1.5702	1.2873	1.22	0.223007	650
HDL-Triglycerides	0.1254	0.1033	1.21	0.224612	2181
Phosphatidylcholine diacyl C 38:7	0.3707	0.3055	1.21	0.225516	649
Phosphatidylcholine alkyl-acyl C40:6	0.4005	0.3318	1.21	0.227858	650
Phosphatidylcholine diacyl (38:3)	0.0851	0.0708	1.20	0.229341	2261
S-HDL-Free cholesterol	0.1211	0.1015	1.19	0.232627	2183
Sphingomyeline C 22:1	2.9411	2.4818	1.19	0.23643	650
1,5-Anhydrosorbitol	-246.0655	209.3522	-1.18	0.239987	2012
Phosphatidylcholine diacyl C 32:2	-0.7894	0.6719	-1.17	0.240492	648
Lysophosphatidylcholine C 20:4	-0.1275	0.1085	-1.17	0.240499	646
Tryptophan	-2.9284	2.4992	-1.17	0.241656	794
VLDL-cholesterol	1.6616	1.4209	1.17	0.242357	2178
Lactate	-9454.6311	8112.9518	-1.17	0.244006	2008
Phosphatidylcholine alkyl-acyl C36:1	0.8637	0.7434	1.16	0.245757	648
Lysophosphatidylcholine(20:4)	-0.0211	0.0184	-1.15	0.250424	2260
Phosphatidylcholine acyl-alkyl C44:6	0.0722	0.0637	1.13	0.257079	793
Phosphatidylcholine diacyl C36:3	7.7691	6.8679	1.13	0.258307	794
Hydroxysphingomyeline C14:1	-0.2966	0.2624	-1.13	0.258645	794
Phosphatidylcholine diacyl C36:4	10.2810	9.1160	1.13	0.259742	794
Lysophosphatidylcholine C 22:5	-0.0729	0.0653	-1.12	0.264636	647
Hydroxybutyrate_2	0.1030	0.0935	1.10	0.270762	1983
Decanoylcarnitine	0.0829	0.0754	1.10	0.272104	789
Lysophosphatidylcholine C 22:4	-0.1038	0.0945	-1.10	0.272391	650
Hydroxyhexadecenoylcarnitine	0.0009	0.0009	1.09	0.27388	782
Phosphatidylcholine acyl-alkyl C38:0	0.1570	0.1435	1.09	0.274208	792
LDL-Triglycerides	0.7844	0.7175	1.09	0.274425	2185
XL-LDL-ApoB	0.3159	0.2902	1.09	0.27651	2188
Spingomyeline C16:1	-0.7776	0.7164	-1.09	0.278095	794

Name	BETA	SE	t-value	p-value	*N*
Lipids (CH=CH*CH2CH2)	-80.2993	74.1875	-1.08	0.279213	2000
Phosphatidylcholine diacyl C38:4	6.2041	5.7969	1.07	0.284834	793
Sphingomyeline C dih 22:0	95.4900	89.4785	1.07	0.286286	649
lysoPhosphatidylcholine acyl C20:3	-0.1631	0.1540	-1.06	0.289769	794
Phosphatidylcholine diacyl C40:3	0.0364	0.0345	1.05	0.291782	791
Betaine	1400.0068	1334.6173	1.05	0.294306	2012
Hydroxyhexadecanoylcarnitine	0.0006	0.0006	1.05	0.294761	792
Glutamine	-20.3888	19.4510	-1.05	0.29486	794
TMAO_Betaine	78.7771	76.2238	1.03	0.301495	2007
Sphingomyeline C dih 23:0	81.9513	79.3116	1.03	0.301855	651
Phosphatidylcholine diacyl C 30:0	-0.6958	0.6735	-1.03	0.301949	650
Hexenoylcarnitine	0.0010	0.0010	1.03	0.303113	794
PE-plasmologen 18:1/18:1	0.0744	0.0725	1.03	0.305148	651
S-HDL-ApoA2	-0.3972	0.3873	-1.03	0.305208	2184
Octenoylcarnitine	0.0131	0.0128	1.02	0.308574	785
Phosphatidylcholine diacyl C 36:3	11.6641	11.4826	1.02	0.3101	649
Phosphatidylcholine diacyl C40:1	0.0187	0.0184	1.02	0.310131	793
Phosphatidylcholine diacyl C 36:2	22.3630	22.0639	1.01	0.311173	648
Phosphatidylcholine diacyl C42:0	0.0319	0.0315	1.01	0.311565	794
Ceramide Glu C16:0	0.0330	0.0332	0.99	0.320287	648
PE-plasmologen 18:0/18:1	0.0984	0.0989	0.99	0.320344	651
Hydroxybutyrylcarnitine	0.0034	0.0034	0.99	0.320506	775
Phosphatidylcholine diacyl C32:1	0.0990	0.1002	0.99	0.323103	788
Formate	-0.0367	0.0372	-0.99	0.323332	2006
Phosphatidylcholine acyl-alkyl(O-34:3)	0.0116	0.0118	0.98	0.325258	2259
Phosphatidylcholine diacyl C 32:0	1.2006	1.2219	0.98	0.326184	649
Phosphatidylcholine alkyl-acyl C32:1	-0.2789	0.2847	-0.98	0.327617	650
Phosphatidylethanolamine(38:2)	0.1365	0.1399	0.98	0.32941	2258
Phosphatidylcholine diacyl C 38:4	9.4543	9.7055	0.97	0.330362	649
Threonine	-4.9665	5.1092	-0.97	0.331314	794
Hydroxysphingomyeline C16:1	-0.1327	0.1368	-0.97	0.332164	794
Tyrosine	-3.0380	3.1432	-0.97	0.334073	794
S-HDL-phospholipids	0.4825	0.5000	0.96	0.334696	2185
Phosphatidylcholine diacyl C34:2	11.3734	11.8155	0.96	0.336049	794
S-LDL-triglycerides	-0.0468	0.0487	-0.96	0.336649	2183
Phosphatidylcholine diacyl C 38:6	6.8689	7.2302	0.95	0.34245	650
VLDL-ApoB	0.2725	0.2876	0.95	0.343335	2181
Lysophosphatidylcholine(16:0)	-0.2401	0.2538	-0.95	0.344153	2262
Phosphatidylcholine diacyl C38:1	0.1023	0.1083	0.94	0.344977	794
Ceramide C16:0	0.0748	0.0792	0.94	0.34533	650
Sphingomyeline C 18:1	1.1630	1.2321	0.94	0.345566	651
Phosphatidylethanolamine 32:2	0.1064	0.1128	0.94	0.345996	650
Hydroxysphingomyeline C24:1	-0.0697	0.0745	-0.94	0.349675	794
Triglycerides(54:2)	-0.0510	0.0548	-0.93	0.352231	2245
Phosphatidylcholine alkyl-acyl C36:3	0.5918	0.6371	0.93	0.353257	648
Lysophosphatidylcholine C 20:5	-0.0612	0.0661	-0.93	0.354892	647

Name	BETA	SE	t-value	p-value	*N*
Tyrosine	-1013.0963	1095.8059	-0.92	0.355326	2014
L-VLDL-triglycerides	0.0727	0.0790	0.92	0.357169	2181
Phosphatidylethanolamine(36:3)	0.0013	0.0014	0.92	0.358499	2254
Acetylcarnitine	0.4078	0.4482	0.91	0.363168	794
lysoPhosphatidylcholine acyl C28:1	-0.0203	0.0224	-0.91	0.365177	792
Sphingomyeline C 15:0	0.5819	0.6483	0.90	0.369739	650
XL-LDL-Free cholesterol	0.2323	0.2589	0.90	0.369799	2188
XL-VLDL-cholesterol	0.2079	0.2319	0.90	0.370136	2184
Lysophosphatidylcholine C 22:6	-0.0627	0.0707	-0.89	0.376067	646
Phosphatidylcholine alkyl-acyl C34:1	0.7240	0.8297	0.87	0.383226	649
Triglycerides(56:2)	-0.0041	0.0047	-0.86	0.388454	2236
Phosphatidylcholine acyl-alkyl(O-44:5)	0.0020	0.0024	0.86	0.392304	2259
Glycine	-14.1780	16.6154	-0.85	0.393748	787
Phosphatidylcholine diacyl C 32:1	-0.1200	0.1415	-0.85	0.396773	648
Phosphatidylcholine acyl-alkyl C32:1	0.0983	0.1161	0.85	0.397347	794
Phosphatidylcholine acyl-alkyl(O-38:7)	-0.0019	0.0023	-0.84	0.398223	2256
Lysophosphatidylcholine C 18:2	-3.0736	3.6445	-0.84	0.399337	650
Spingomyeline C24:0	1.0422	1.2376	0.84	0.399976	793
plasma-Triglycerides	0.0368	0.0439	0.84	0.402157	2177
Phosphatidylcholine diacyl C 40:7	0.5245	0.6292	0.83	0.40477	650
XL-LDL-cholesterol	0.7663	0.9225	0.83	0.406276	2188
Phosphatidylcholine diacyl C 38:5	4.3620	5.2496	0.83	0.406325	649
Phosphatidylethanolamine 32:0	0.0145	0.0175	0.83	0.406514	647
Phosphatidylcholine acyl-alkyl C42:2	0.0261	0.0316	0.83	0.40845	794
Sphingomyeline C 20:0	1.9004	2.2998	0.83	0.408903	650
Phosphatidylcholine acyl-alkyl(O-36:6)	-0.0411	0.0499	-0.83	0.409286	2244
lysoPhosphatidylcholine acyl C20:4	-0.3328	0.4112	-0.81	0.418531	793
Hydroxyoctadecenoylcarnitine	0.0006	0.0007	0.81	0.418969	794
Phosphatidylcholine diacyl C 30:1	0.1613	0.1994	0.81	0.419051	648
XL-VLDL-Free cholesterol	0.0324	0.0402	0.81	0.419238	2183
Lysophosphatidylethanolamine(18:0)	-0.0077	0.0095	-0.81	0.42	2252
Phosphatidylcholine diacyl (40:6)	-0.0437	0.0543	-0.80	0.421561	2251
Sphingomyeline C 22:0	2.6844	3.3666	0.80	0.425526	649
Phosphatidylcholine acyl-alkyl(O-36:5)	0.0162	0.0204	0.80	0.425807	2260
Lysophosphatidylcholine(16:1)	-0.0064	0.0081	-0.79	0.429678	2248
Ceramide(d18:1/24:0)	0.0059	0.0075	0.79	0.431545	2261
Octadecadienylcarnitine	0.0034	0.0043	0.79	0.432248	794
Phosphatidylcholine diacyl C 42:4	0.0771	0.0991	0.78	0.436625	650
Spingomyeline C20:2	-0.0332	0.0428	-0.78	0.438471	794
Ceramide C18:0	0.0141	0.0183	0.77	0.44146	649
Ceramide Glu C24:1	0.0331	0.0438	0.75	0.450606	649
Lysophosphatidylcholine(14:0)	-0.0041	0.0055	-0.75	0.451357	2261
Triglycerides(51:4)	0.0043	0.0057	0.75	0.453446	2248
PE-plasmologen 18:1/20:4	-0.3952	0.5304	-0.75	0.456418	650
Phosphatidylcholine diacyl C40:2	0.0155	0.0209	0.74	0.456985	791
Spingomyeline C26:0	-0.0095	0.0127	-0.74	0.457176	793

Name	BETA	SE	t-value	p-value	*N*
PE-plasmologen 18:1/18:2	0.1514	0.2040	0.74	0.458306	650
Acetate	-0.0420	0.0568	-0.74	0.459246	1977
Phosphatidylcholine acyl-alkyl C34:1	0.3079	0.4166	0.74	0.460141	793
Phosphatidylcholine diacyl C 40:4	0.3484	0.4720	0.74	0.460747	649
Phosphatidylcholine acyl-alkyl C42:4	0.0322	0.0438	0.74	0.461432	794
Phosphatidylcholine diacyl C36:2	7.7112	10.4923	0.73	0.462594	794
Triglycerides(54:7)	0.0101	0.0137	0.73	0.463333	2246
Lipids (CH2)	-55.1358	75.4446	-0.73	0.46498	2001
Methionine	-1.1058	1.5218	-0.73	0.467651	794
Sphingomyeline(d18:1/20:1)	-0.0038	0.0052	-0.73	0.468248	2263
Myoinositol	-56.8363	78.7998	-0.72	0.470824	2009
Phosphatidylcholine diacyl C42:1	0.0103	0.0143	0.72	0.471821	794
XS-VLDL-cholesterol	0.0078	0.0109	0.72	0.472844	2187
Spingomyeline C18:1	-0.4150	0.5809	-0.71	0.475098	794
Hydroxysphingomyeline C22:1	-0.4984	0.7032	-0.71	0.478706	794
Proline	7.4088	10.5659	0.70	0.483387	789
Triglycerides(56:7)	0.0435	0.0626	0.69	0.487798	2246
Phosphatidylcholine diacyl C 38:2	0.6817	0.9963	0.68	0.494078	651
Sphingomyeline(d18:1/18:2)	0.0011	0.0016	0.68	0.49474	2260
Phosphatidylcholine acyl-alkyl C34:0	0.0471	0.0692	0.68	0.496504	794
Tetradecadienylcarnitine	0.0022	0.0032	0.68	0.49868	789
Ceramide C22:0	0.0630	0.0936	0.67	0.501078	649
Phosphatidylcholine diacyl C26:0	0.0167	0.0249	0.67	0.501168	731
Triglycerides(46:1)	-0.0194	0.0289	-0.67	0.502614	2211
Phosphatidylcholine acyl-alkyl(O-34:1)	0.0039	0.0059	0.67	0.506041	2259
lysoPhosphatidylcholine acyl C14:0	0.0644	0.0970	0.66	0.507031	793
Triglycerides(52:1)	-0.0500	0.0761	-0.66	0.510788	2242
Phosphatidylcholine acyl-alkyl C44:4	0.0116	0.0176	0.66	0.512032	794
XS-LDL-ApoB	0.4366	0.6696	0.65	0.514404	2183
Nonaylcarnitine	0.0482	0.0741	0.65	0.515314	791
VLDL-Free cholesterol	0.3656	0.5673	0.64	0.519371	2182
Lysophosphatidylcholine C 18:1	-1.4611	2.2729	-0.64	0.520543	648
Arginine	-2.4807	3.8999	-0.64	0.524899	794
Triglycerides(53:1)	-0.0023	0.0037	-0.63	0.530546	2207
XXL-VLDL-triglycerides	-0.0580	0.0933	-0.62	0.534289	2174
L-LDL-triglycerides	-0.0254	0.0414	-0.61	0.539607	2185
Phosphatidylcholine diacyl (38:6)	0.0850	0.1389	0.61	0.540482	2254
Dodecenoylcarnitine	0.0053	0.0086	0.61	0.542218	788
Triglycerides(54:6)	0.0385	0.0635	0.61	0.544317	2246
Phosphatidylcholine diacyl (40:5)	0.0085	0.0141	0.60	0.545483	2256
Proline	-509.4497	845.3949	-0.60	0.546831	2010
Phosphatidylcholine diacyl C28:1	-0.0960	0.1596	-0.60	0.547744	794
Triglycerides(50:3)	0.0418	0.0697	0.60	0.548103	2251
Triglycerides(52:4)	0.0358	0.0604	0.59	0.553434	2253
Phosphatidylcholine alkyl-acyl C32:0	0.1918	0.3237	0.59	0.553767	650
Phosphatidylcholine acyl-alkyl C36:4	-0.5615	0.9509	-0.59	0.555019	793

Name	BETA	SE	t-value	p-value	*N*
Triglycerides(51:3)	0.0376	0.0644	0.58	0.559294	2239
Sphingomyeline(d18:1/24:0)	0.0128	0.0222	0.58	0.565135	2260
Glutarylcarnitine (Hydroxyhexanoylcarnitine)	0.0010	0.0017	0.57	0.565601	788
Lipids_CH2CO	-44.0181	76.6504	-0.57	0.565849	2001
Sphingomyeline C 20:1	44.9383	78.4664	0.57	0.567041	650
Phosphatidylethanolamine 32:1	0.0841	0.1480	0.57	0.569939	644
Triglycerides(50:2)	0.0387	0.0683	0.57	0.571087	2256
XL-VLDL-phospholipids	0.0330	0.0583	0.57	0.571338	2182
lysoPhosphatidylcholine acyl C28:0	-0.0281	0.0501	-0.56	0.57447	789
Leucine / Isoleucine	-5.6052	10.1280	-0.55	0.580119	793
Phosphatidylcholine diacyl (36:2)	0.1515	0.2756	0.55	0.582477	2262
Dimethylglycine	42.0366	77.2745	0.54	0.586508	2009
S-HDL-ApoA1	-0.5308	0.9847	-0.54	0.589925	2188
VLDL-phospholipids	0.7279	1.3644	0.53	0.593733	2182
Acetone	0.0277	0.0520	0.53	0.594045	2006
Sphingomyeline C 24:3	-0.4168	0.7847	-0.53	0.595494	651
Sphingomyeline(d18:1/23:1)	-0.0047	0.0091	-0.52	0.603933	2262
Phosphatidylethanolamine 40:3	0.0043	0.0083	0.52	0.606339	647
Phosphatidylcholine acyl-alkyl C38:5	-0.4371	0.8491	-0.51	0.606833	794
Phenylalanine	-0.9544	1.9295	-0.49	0.620986	794
Lipids_CH3	-36.2720	74.0097	-0.49	0.624118	2004
Spingomyeline C16:0	-2.0123	4.1675	-0.48	0.62933	794
Phosphatidylcholine diacyl (36:1)	0.0367	0.0762	0.48	0.63038	2254
Octadecenoylcarnitine	0.0242	0.0502	0.48	0.630581	792
Phosphatidylethanolamine 42:6	-0.0033	0.0068	-0.48	0.630713	650
Ornithine	-1.9861	4.1330	-0.48	0.630966	794
Sphingomyeline C dih 18:0	-0.1815	0.3783	-0.48	0.631499	651
Sphingomyeline(d18:1/17:0)	-0.0010	0.0022	-0.48	0.632139	2253
Phosphatidylcholine acyl-alkyl C38:2	0.0483	0.1011	0.48	0.633098	794
Hexose	-90.3905	190.0176	-0.48	0.634422	792
Dodecanedioylcarnitine	0.0009	0.0018	0.47	0.635353	792
lysoPhosphatidylcholine acyl C26:0	0.0321	0.0678	0.47	0.636075	788
plasma ApoB	0.7405	1.6156	0.46	0.646764	2188
lysoPhosphatidylcholine acyl C18:0	-0.6846	1.4943	-0.46	0.646978	794
Triglycerides(56:5)	0.0227	0.0498	0.46	0.648325	2125
Phosphatidylcholine diacyl C 40:6	1.4246	3.1226	0.46	0.648383	649
Triglycerides(48:2)	-0.0453	0.1000	-0.45	0.650715	2244
Sphingomyeline C dih 24:0	-40.0546	88.4365	-0.45	0.650758	649
Phosphatidylcholine acyl-alkyl C40:0	0.1407	0.3151	0.45	0.655347	792
Phosphatidylcholine alkyl-acyl C36:5	0.4329	0.9732	0.44	0.6566	650
Hydroxypropionylcarnitine	8.9460	20.1257	0.44	0.656798	789
Sphingomyeline(d18:1/22:1)	-0.0091	0.0206	-0.44	0.657658	2263
Phosphatidylcholine acyl-alkyl C30:0	0.0087	0.0198	0.44	0.659196	794
Phosphatidylcholine alkyl-acyl C34:2	0.4665	1.0574	0.44	0.65923	650
Propenoylcarnitine	0.0002	0.0005	0.44	0.662192	794

Name	BETA	SE	t-value	p-value	*N*
Phosphatidylcholine diacyl C 40:5	0.5817	1.3474	0.43	0.666077	647
Spingomyeline C24:1	-0.9208	2.1338	-0.43	0.666183	794
XS-VLDL-triglycerides	0.0136	0.0319	0.43	0.670875	2183
Phosphatidylcholine acyl-alkyl(O-36:2)	-0.0015	0.0035	-0.42	0.672867	2259
Lysophosphatidylcholine C 22:0	0.0302	0.0729	0.41	0.678915	651
Pyruvate_Oxaloacetate	4071.9179	9939.2325	0.41	0.682083	2010
Phosphatidylethanolamine(O-38:5)	-0.0054	0.0131	-0.41	0.6829	2253
XXL-VLDL-phospholipids	-0.0266	0.0660	-0.40	0.686687	2177
Triglycerides(54:5)	0.0229	0.0570	0.40	0.688598	2247
Triglycerides(52:3)	0.0181	0.0458	0.39	0.692943	2260
LDL-cholesterol	-1.4978	3.8434	-0.39	0.696798	2188
Phosphatidylcholine acyl-alkyl(O-34:2)	-0.0051	0.0131	-0.39	0.699704	2255
Ceramide C20:0	0.0054	0.0143	0.38	0.705939	649
Hydroxyhexadecadienylcarnitine	-0.0003	0.0007	-0.38	0.706733	794
Triglycerides(56:8)	0.0087	0.0233	0.37	0.709907	2249
Sphingomyeline(d18:1/24:1)	-0.0202	0.0544	-0.37	0.710105	2261
lysoPhosphatidylcholine acyl C16:0	-1.4995	4.0408	-0.37	0.710679	794
IDL-Triglycerides	0.0197	0.0532	0.37	0.710729	2178
L-LDL-Free cholesterol	0.0965	0.2650	0.36	0.7158	2188
PE-plasmologen 18:1/22:6	0.0685	0.1883	0.36	0.716139	650
Phosphatidylcholine diacyl C 36:5	1.1213	3.1010	0.36	0.717769	645
Phosphatidylcholine acyl-alkyl C42:1	0.0072	0.0206	0.35	0.724643	794
PE-plasmologen 16:0/20:4	-0.2042	0.5801	-0.35	0.724947	648
Phosphatidylcholine acyl-alkyl C30:1	0.0049	0.0139	0.35	0.726045	791
Triglycerides(46:2)	-0.0048	0.0136	-0.35	0.726556	2223
Phosphatidylcholine alkyl-acyl C40:4	0.0731	0.2099	0.35	0.727829	648
Triglycerides(50:4)	0.0250	0.0717	0.35	0.727955	2250
Phosphatidylcholine diacyl (40:7)	0.0026	0.0073	0.35	0.728051	2252
Hydroxyisobutyrate_3	-321.4775	925.3928	-0.35	0.728331	2005
PE-plasmologen 18:0/22:6	0.1055	0.3051	0.35	0.729713	648
Sphingomyeline(d18:1/25:0)	0.0004	0.0012	0.34	0.730474	2258
Sphingomyeline(d18:1/18:1)	-0.0030	0.0090	-0.34	0.734911	2259
Phosphatidylcholine acyl-alkyl C36:2	-0.2273	0.6772	-0.34	0.737227	794
PE-plasmologen 18:0/20:4	-0.3108	0.9327	-0.33	0.739019	649
Phosphatidylcholine acyl-alkyl C30:2	-0.0021	0.0064	-0.33	0.740243	794
Spingomyeline C26:1	-0.0090	0.0278	-0.32	0.746477	794
Spingomyeline C18:0	-0.3357	1.0451	-0.32	0.748156	794
Triglycerides(52:5)	0.0210	0.0655	0.32	0.74873	2250
Phosphatidylcholine acyl-alkyl(O-38:6)	-0.0027	0.0083	-0.32	0.74907	2260
Dimethylamine	-27.6135	86.7613	-0.32	0.750313	2010
Phosphatidylcholine acyl-alkyl C42:5	0.0275	0.0893	0.31	0.758343	794
Phosphatidylethanolamine alkyl-acyl 38:7	-0.0059	0.0193	-0.31	0.759982	647
L-LDL-cholesterol	0.3414	1.1199	0.30	0.76054	2188
XL-VLDL-triglycerides	0.0297	0.0975	0.30	0.760755	2181
Sphingomyeline(d18:1/16:1)	-0.0032	0.0109	-0.30	0.765331	2262
Ceramide(d18:1/24:1)	-0.0014	0.0048	-0.30	0.767937	2255

Name	BETA	SE	t-value	p-value	*N*
Valine	-4.0565	13.7895	-0.29	0.768701	794
Phosphatidylcholine acyl-alkyl C38:3	-0.0559	0.1927	-0.29	0.771753	794
Phosphatidylcholine acyl-alkyl C40:1	0.0214	0.0756	0.28	0.777053	794
XS-LDL-phospholipids	-0.1554	0.5548	-0.28	0.779398	2182
XXL-VLDL-cholesterol	-0.1516	0.5422	-0.28	0.779781	2176
Hydroxyvalerylcarnitine (Methylmalonylcarnitine)	-0.0005	0.0017	-0.28	0.780933	791
Triglycerides(46:0)	-0.0055	0.0199	-0.28	0.781351	2157
Phosphatidylcholine acyl-alkyl C40:4	0.0293	0.1069	0.27	0.783972	794
Phosphatidylcholine acyl-alkyl C40:2	-0.0240	0.0882	-0.27	0.785219	794
Carnitine	-0.4899	1.8088	-0.27	0.786592	794
VLDL-Triglycerides	-1.3683	5.1113	-0.27	0.788947	2176
Tetradecenoylcarnitine	0.0018	0.0069	0.27	0.789881	790
Phosphatidylcholine diacyl C 38:3	1.2547	4.7558	0.26	0.791992	647
Phosphatidylcholine acyl-alkyl C38:4	-0.1667	0.6320	-0.26	0.792033	794
Phosphatidylcholine diacyl C 36:1	1.0831	4.1101	0.26	0.792239	647
Succinate	-21.7916	83.8768	-0.26	0.795041	2006
LDL-phospholipids	0.5059	1.9634	0.26	0.79669	2188
Phosphatidylcholine acyl-alkyl C36:3	-0.1038	0.4031	-0.26	0.796911	794
Phosphatidylcholine diacyl C42:2	-0.0035	0.0138	-0.26	0.796979	787
Triglycerides(58:8)	0.0026	0.0101	0.25	0.799365	2235
Phosphatidylcholine alkyl-acyl C38:2	-0.0928	0.3688	-0.25	0.80141	651
Phosphatidylcholine acyl-alkyl C42:3	0.0091	0.0378	0.24	0.809781	794
LDL-Free cholesterol	0.2666	1.1139	0.24	0.810893	2188
PE-plasmologen 18:1/20:5	-0.0353	0.1511	-0.23	0.815332	650
Citrate	-16.3810	70.3089	-0.23	0.815796	2014
Phosphatidylcholine diacyl C 36:0	0.1048	0.4568	0.23	0.818696	650
Triglycerides(50:0)	-0.0213	0.0935	-0.23	0.819562	2232
Phosphatidylcholine acyl-alkyl C32:2	-0.0077	0.0338	-0.23	0.820224	794
lysoPhosphatidylcholine acyl C26:1	-0.0081	0.0359	-0.22	0.822359	794
Phosphatidylcholine diacyl C36:0	0.0335	0.1492	0.22	0.822634	794
L-LDL-ApoB	-0.0767	0.3432	-0.22	0.823214	2187
Triglycerides(50:5)	-0.0013	0.0058	-0.22	0.828031	2242
Sphingomyeline(d18:1/24:2)	-0.0053	0.0245	-0.22	0.828094	2262
Phosphatidylcholine diacyl C 40:0	-0.0453	0.2089	-0.22	0.828306	647
Triglycerides(50:1)	-0.0158	0.0748	-0.21	0.83305	2253
lysoPhosphatidylcholine acyl C16:1	0.0428	0.2042	0.21	0.833967	790
Phosphatidylcholine acyl-alkyl(O-36:3)	0.0007	0.0032	0.21	0.835596	2257
Sphingomyeline C 24:0	0.4460	2.1962	0.20	0.839133	649
Phosphatidylcholine alkyl-acyl C36:0	0.0325	0.1606	0.20	0.83947	650
Sphingomyeline(d18:1/22:0)	0.0069	0.0342	0.20	0.839647	2262
Phosphatidylcholine acyl-alkyl(O-36:4)	-0.0040	0.0202	-0.20	0.843835	2255
S-HDL-cholesterol	-0.0848	0.4409	-0.19	0.84753	2183
PE-plasmologen 16:0/22:5	-0.0273	0.1430	-0.19	0.848578	650
Octanoylcarnitine	0.0029	0.0151	0.19	0.849731	788
Phosphatidylcholine alkyl-acyl C38:1	0.0633	0.3344	0.19	0.850008	651

Name	BETA	SE	t-value	p-value	*N*
Sphingomyeline(d18:1/20:0)	0.0028	0.0152	0.19	0.8514	2263
PE-plasmologen 16:0/20:5	0.0284	0.1532	0.19	0.85289	643
Triglycerides(51:2)	-0.0127	0.0688	-0.19	0.853036	2247
Phosphatidylcholine diacyl C38:0	0.0349	0.1900	0.18	0.85408	794
Phenylalanine	-45.4450	249.6529	-0.18	0.855575	2015
Phosphatidylcholine acyl-alkyl(O-38:5)	-0.0027	0.0152	-0.18	0.860024	2260
Phosphatidylethanolamine(O-38:7)	-0.0024	0.0136	-0.18	0.860276	2241
Triglycerides(56:3)	-0.0093	0.0534	-0.17	0.861199	2238
Phosphatidylethanolamine 34:0	0.0042	0.0263	0.16	0.873179	649
Phosphatidylcholine acyl-alkyl C36:0	-0.0056	0.0375	-0.15	0.881088	792
Triglycerides(48:3)	-0.0139	0.0939	-0.15	0.882245	2241
Phosphatidylcholine acyl-alkyl C36:1	-0.0552	0.3730	-0.15	0.882348	793
Phosphatidylcholine diacyl C 26:0	0.0237	0.1648	0.14	0.885881	648
Triglycerides(51:1)	-0.0021	0.0145	-0.14	0.886137	2232
Triglycerides(58:9)	-0.0013	0.0093	-0.14	0.887377	2230
PE-plasmologen 16:0/18:1	0.0103	0.0738	0.14	0.888968	650
Sphingomyeline(d18:1/15:0)	0.0006	0.0043	0.14	0.889676	2255
Phosphatidylcholine acyl-alkyl C36:5	0.0826	0.6147	0.13	0.89308	792
L-LDL-phospholipids	-0.0737	0.5607	-0.13	0.895395	2187
Phosphatidylcholine diacyl (36:6)	0.0002	0.0015	0.13	0.895829	2239
Butenylcarnitine	0.0001	0.0010	0.13	0.899509	794
Glycerol	-10.2162	81.2641	-0.13	0.899969	2007
LDL-ApoB	-0.1633	1.3267	-0.12	0.902052	2188
Phosphatidylcholine acyl-alkyl C44:3	-0.0009	0.0071	-0.12	0.902938	794
Tiglylcarnitine	-0.0002	0.0014	-0.12	0.906041	794
Phosphatidylcholine acyl-alkyl C34:2	-0.0711	0.6043	-0.12	0.906316	794
Lysophosphatidylcholine C 20:0	8.8906	78.3525	0.11	0.909694	649
Phosphatidylcholine acyl-alkyl C34:3	-0.0429	0.4192	-0.10	0.918444	794
Phosphatidylcholine acyl-alkyl C44:5	0.0081	0.0792	0.10	0.918814	794
Phosphatidylcholine acyl-alkyl(O-38:4)	0.0007	0.0073	0.10	0.922692	2260
Triglycerides(56:6)	0.0054	0.0578	0.09	0.92536	2233
XS-LDL-Free cholesterol	0.0231	0.2501	0.09	0.92629	2184
XS-LDL-cholesterol	0.1048	1.1674	0.09	0.92846	2186
Glutamate	-6.5508	75.4108	-0.09	0.930785	2008
XXL-VLDL-Free cholesterol	0.0052	0.0600	0.09	0.930978	2178
Phosphatidylcholine diacyl C32:3	-0.0025	0.0311	-0.08	0.935986	794
Sphingomyeline(d18:1/14:0)	0.0006	0.0081	0.08	0.937455	2259
Phosphatidylcholine acyl-alkyl C38:6	0.0365	0.4652	0.08	0.937534	793
Sphingomyeline(d18:1/18:0)	0.0015	0.0191	0.08	0.938558	2260
Histidine	0.2177	2.8664	0.08	0.939469	793
PE-plasmologen 16:0/22:6	0.0204	0.3044	0.07	0.9467	646
Phosphatidylethanolamine(O-36:5)	0.0006	0.0096	0.06	0.952598	2253
S-HDL-triglycerides	0.0089	0.1514	0.06	0.953293	2187
Glucose	-0.0010	0.0178	-0.06	0.953938	2004
Triglycerides(48:0)	0.0022	0.0408	0.05	0.956873	2154
Triglycerides(48:1)	0.0030	0.0586	0.05	0.959147	2231

Name	BETA	SE	t-value	p-value	*N*
Phosphatidylcholine diacyl C 34:0	-0.0317	0.6225	-0.05	0.959342	651
Phosphatidylcholine acyl-alkyl C38:1	-0.0024	0.0486	-0.05	0.960266	794
Triglycerides(54:4)	0.0028	0.0580	0.05	0.960843	2251
Sphingomyeline C 22:2	4.3857	89.2959	0.05	0.960843	649
XS-LDL-triglycerides	-0.0085	0.1740	-0.05	0.961046	2182
Alpha-ketoisovalerate	18.3234	397.6434	0.05	0.963251	2014
Sphingomyeline C dih 16:0	0.0304	0.6692	0.05	0.963774	649
PE-plasmologen 18:0/20:5	-0.0073	0.1706	-0.04	0.965662	648
Triglycerides(58:10)	0.0002	0.0040	0.04	0.968962	2200
Lysophosphatidylcholine C 16:0	-0.3769	10.0728	-0.04	0.970164	650
Phosphatidylcholine acyl-alkyl C40:3	-0.0017	0.0500	-0.03	0.973658	794
Sphingomyeline(d18:1/23:0)	-0.0004	0.0131	-0.03	0.973792	2262
Sphingomyeline(d18:1/21:0)	0.0002	0.0054	0.03	0.977014	2259
lysoPhosphatidylcholine acyl C6:0	0.0001	0.0018	0.03	0.977302	792
Triglycerides(52:2)	-0.0013	0.0473	-0.03	0.977692	2262
Sphingomyeline(d18:1/25:1)	-0.0001	0.0025	-0.03	0.978594	2256
Phosphatidylcholine acyl-alkyl C40:6	0.0063	0.2606	0.02	0.980731	794
Phosphatidylcholine acyl-alkyl C40:5	0.0042	0.1736	0.02	0.980733	794
Triglycerides(52:6)	0.0008	0.0616	0.01	0.990163	2248
Lysophosphatidylcholine C 18:0	-0.0364	3.9873	-0.01	0.992723	649
Sphingomyeline(d18:1/16:0)	-0.0007	0.0831	-0.01	0.99333	2261
M-LDL-triglycerides	0.0000	0.0549	0.00	0.99938	2188

APPENDIX of CHAPTER 3.2

Non-synonymous Variation in *NKPD1* Increases Depressive Symptoms in the European Populations

Supplement 1

Table S1. Descriptive data of ERF study (*N* = 2 353)

Characteristics	Median (Range)	Mean (*SD*)
Sex (% women)		56
Age, Years	48.84 (18 – 96)	48.73 (14.96)
CES-D	8.24 (0 – 59)	10.74 (9.65)

Table S2. Association of depressive symptoms (CES-D) with age

	Intercept		**Age**	
	Estimate ± s.e.	*p* value	Estimate ± s.e.	*p* value
Males	2.893 ± 0.928	0.0019	0.131 ± 0.018	8.40×10^{-13}
Females	4.802 ± 0.916	1.85×10^{-07}	0.145 ± 0.018	1.74×10^{-15}

Table S3. Genome-wide results of gene-based analysis
Provided online see URL: http://www.biologicalpsychiatryjournal.com/article/S0006-3223(16)32669-5/addons

Table S4. Genetic variants within *NKPD1* gene in discovery (ERF) sample

Name	Position	MAF	*N*	Estimate of the effect size (β ± s.e.)	*p* value	Function	PolyPhen prediction	MAF 1000G
rs28469095	19:45655333	0.10614	1206	0.086 ± 0.653	0.8950	missense	benign	0.083
exm1479909	19:45655486	0.00063	792	10.536 ± 9.306	0.2576	missense	benign	0.029
19:45655513	19:45655513	0.00041	1209	10.530 ± 9.310	0.2580	missense	Probably damaging	NA
rs11083761	19:45655636	0.00952	1208	5.947 ± 2.048	0.0037	missense	benign	0.0008
rs117934605	19:45655647	0.13215	1207	1.040 ± 0.585	0.0753	missense	benign	0.051
exm1479927	19:45656069	0.00189	792	6.837 ± 5.438	0.2086	missense	benign	0.021
rs144764378,exm1479947	19:45659155	0.00401	1997	3.311 ± 2.388	0.1655	missense	NA	0.0024
19:45662190	19:45662190	0.00083	1207	-0.686 ± 6.573	0.9169	missense	NA	NA
rs75291769,exm1479956	19:45662295	0.04591	1993	2.900 ± 0.792	$2.5x10^{-04}$	missense	NA	0.012

Table S5. Genetic variants within *NKPD1* gene in replication (RS) sample

Name	Position	MAF	*N*	Estimate of the effect size (β ± s.e.)	*p* value	Function	PolyPhen prediction	MAF 1000G
19:45655390	19:45655390	0.00069	1604	7.968 ± 4.893	0.1034	missense	possibly-damaging	NA
19:45655398	19:45655398	0.00103	1604	-1.067 ± 3.094	0.7302	missense	probably-damaging	NA
19:45655456	19:45655456	0.00068	1604	6.274 ± 4.890	0.1995	missense	probably-damaging	NA
19:45655508	19:45655508	0.00066	1604	-0.076 ± 4.889	0.9875	missense	probably-damaging	NA
19:45655550	19:45655550	0.00265	1604	-1.131 ± 2.448	0.6442	missense	probably-damaging	NA
19:45656112	19:45656112	0.00099	1604	10.545 ± 3.093	$6.5x10^{-4}$	missense	probably-damaging	NA
19:45656114	19:45656114	0.00230	1604	10.149 ± 2.306	$1.1x10^{-5}$	missense	probably-damaging	NA
19:45656273	19:45656273	0.00033	1604	15.620 ± 6.915	0.0239	missense	possibly-damaging	NA
19:45656435	19:45656435	0.00126	1604	-3.579 ± 3.466	0.3018	missense	benign	NA
19:45656809	19:45656809	0.00204	1604	-0.144 ± 2.446	0.9530	missense	probably-damaging	NA
19:45662082	19:45662082	0.00031	1604	-3.353 ± 6.911	0.6275	missense	NA	NA
19:45662147	19:45662147	0.00031	1604	-2.557 ± 6.911	0.7114	stop-gained	NA	NA
19:45662290	19:45662290	0.00170	1604	1.182 ± 2.614	0.6510	missense	NA	NA
rs144764378	19:45659155	0.00717	1604	-0.010 ± 1.451	0.9948	missense	NA	0.002
rs150780522	19:45661972	0.00031	1604	-3.940 ± 6.912	0.5687	missense	NA	0.0006

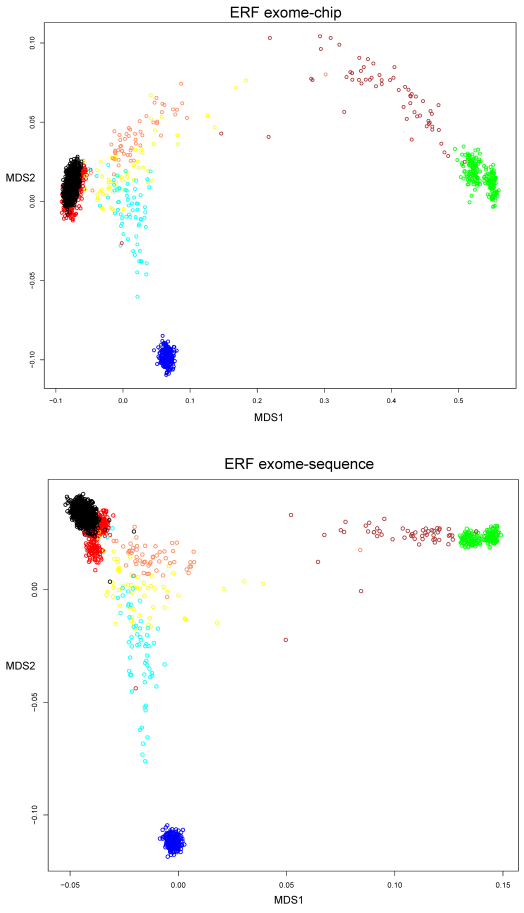
Table S6. Genetic variants mapped to NKPD1 gene in the GWAS of depression performed by Psychiatric genetics consortium (PGC) (1)

Snpid	hg18chr	bp	a1	a2	or	se	*p* value	info	ngt	hg19chr	hg19_bp	Gene	Function	freq1000G
rs10421247	19	50349326	T	C	0.991	0.0258	0.733	0.672	0	chr19	45657486	NKPD1	intron	0.4335
rs10417602	19	50352479	T	C	1.048	0.0402	0.244	0.868	0	chr19	45660639	NKPD1	intron	0.0964
rs8109620	19	50354299	A	G	0.956	0.0252	0.072	0.799	0	chr19	45662459	NKPD1	intron	0.2173

Table S7. Genetic variants mapped to NKPD1 gene in the GWAS of depressive symptoms performed by Hek et al. (2)

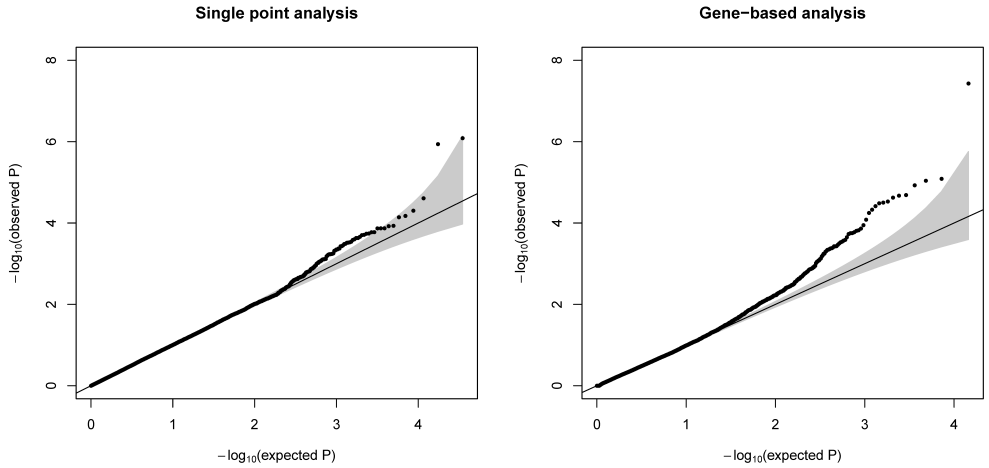
Snpid	A1	A2	Freq1	Freq SE	Min Freq	Max Freq	Zscore	*p* value	HetPVal	hg19chr	hg19_bp	Gene	Function	Freq 1000G
rs8108762	a	g	0.311	0.0323	0.204	0.379	-1.67	0.095	0.667	chr19	45653227	NKPD1	3ÚTR	0.270
rs10421247	t	c	0.539	0.0366	0.462	0.643	-1.54	0.123	0.572	chr19	45657486	NKPD1	intron	0.434
rs10416371	a	c	0.546	0.0362	0.479	0.678	1.79	0.073	0.470	chr19	45660136	NKPD1	intron	0.332

Figure S1.



Ethnicity information provided by the principal components analysis. Black circles represent the samples included in this study, red: Caucasian, blue: Chinese, green: African, yellow: Colombian (AMR), cyan: Mexican (AMR), coral: Puerto Rican (AMR), brown: African (AMR).

Figure S2.



Quantile-quantile (QQ) plots of the single-variant and gene-based association analysis. Observed distribution is depicted on the vertical axis and expected distribution is depicted on the horizontal axis.

REFERENCES SUPPLEMENT

1. Major Depressive Disorder Working Group of the Psychiatric GC. A mega-analysis of genome-wide association studies for major depressive disorder. Mol Psychiatry. 2013;18:497-511.
2. Hek K et al. A genome-wide association study of depressive symptoms. Biol Psychiatry. 2013;73:667-678

CHAPTER 4

The epigenetics of depression

CHAPTER 4.1

DNA-methylation signatures and depressive symptoms

Jovanova Olivera, Nedeljkovic Ivana, Lemaitre Rozzenn, Brody Jennifer, Swenson Brenton, et al.

Under review Biological Psychiatry

ABSTRACT

Depression is determined by the interplay between environmental and genetic risk. Epigenetic mechanisms, like DNA methylation, can mediate the association between depression and it's environmental and genetic risk factors. By performing the largest (*N* = 11 256) cross-ethnic epigenome-wide association study of depressive symptoms, we identified hypermethylation of cg04987734 (*p value* = 1.51 × 10^{-8}, *CDC42BPB* gene), cg12325605 (*p value* = 5.24 × 10^{-09}, *ARHGEF3* gene) and cg14023999 (*p value* = 5.99 × 10^{-08}, intergenic) associated with depressive symptoms. These candidate loci for depression are involved in the regulation of serotonin and dopamine levels in the brain, sphingolipid metabolism, synaptic plasticity, and inflammation.

INTRODUCTION

Depression is one of the most common mental health disorders that results in major disability and shorter life-expectancy[1,2], and it is projected to play a leading role in disease burden by the year 2030.[3,4] Limited understanding of the molecular mechanisms underlying depression is a major bottleneck in the development of innovative treatment, prognostic markers, and prevention strategies.

Studying depression is challenging as it is a heterogeneous disorder with a multifactorial etiology, including genetic and environmental risk factors.[5,6] The contribution of genetics to the risk of depression is substantial with heritability estimates of 40 to 50%.[7] Genome-wide association studies (GWASs) have recently identified numerous rare and common genetic variants associated with depression and related traits.[8-12] However, genetic sequence variation alone does not completely explain an individual's risk for developing depression. Among environmental factors, adverse life-events and stress are major risk factors for depression.[13-15] Converging evidence from animal and human studies suggest that psychosocial stressors trigger depression onset by inducing elevations in pro-inflammatory cytokines.[16] These environmental and psychosocial stressors are also known to influence epigenetic mechanisms, such as DNA methylation.[17] These stress-triggered DNA methylation changes can drive sustained changes in gene expression that may link stress exposures with mental health outcomes.[18,19] Thus, an increasing interest in the role of DNA methylation in depression has emerged.[20]

DNA methylation may be global or tissue-specific.[20,21] Tissues likely to be involved in complex psychiatric disorders, such as brain, are not directly accessible from living patients. The, use of post-mortem brain tissue to study DNA methylation is a possible solution, although obtaining a sufficient sample size is challenging.[22,23] To study differential DNA methylation associated with mental health symptoms on a large scale, peripheral tissues such as blood constitutes a useful proxy for detecting trans-tissue changes and the most appropriate tissue for biomarkers.[22,24] Moderate correlation has been demonstrated between blood and brain tissues at non-tissue specific regulatory regions across the methylome.[25] To date, several studies have assessed the correlation between depression and blood DNA methylation.[26,27] However, these studies have been limited to a small number of DNA methylation sites (CpGs) and/or small samples. For instance, the largest published epigenome-wide association study (EWAS) assessed brain DNA methylation in 76 cases persons who died during a depressive episode and 45 controls.[28]

In the current study, we performed an EWAS of depressive symptoms using whole blood samples of 7 948 individuals of European ethnicity from the Cohorts for Heart and Aging Research in Genomic Epidemiology (CHARGE) consortium. We replicated our findings in 3 308 individuals of African-American and European ancestry. Finally, to identify the affected genes, we performed

a genome-wide gene expression analysis followed by pathway analyses of the differentially methylated and expressed genes.

METHODS

Study populations

The study sample for the discovery analysis included a total of 7 948 participants of European ancestry from nine population-based cohorts of the CHARGE consortium (Table 1): Cardiovascular Health Study (CHS)[29], Framingham Heart Study (FHS)[30], Helsinki Birth Cohort Study (HBCS)[31], Cooperative Health Research in the Augsburg Region (KORA) study[32], two sub-cohorts from Lothian Birth-Cohort born in 1921 (LBC1921)[33] and 1936 (LBC1936)[34], two sub-cohorts from Rotterdam Study (RS-III and RS-BIOS)[35] and Generation Scotland: Scottish Family Health Study (GS) study.[36] These cohorts included community dwelling individuals, who were not selected based on disease status. Informed consent was obtained from all participants. The protocol for each study was approved by the institutional review board of each institution.

Table 1. Characteristics of the individuals in the discovery (*N* = 7 948) and replication cohorts (*N* = 3 308)

Study	*N*	Female (%)	Mean Age (SD)	Current smokers* (%)	Depressive symptoms	Antidepressant medication use (%)
Discovery						
CHS	323	194 (60.1)	75.6 (5.2)	173 (53.6)	CES-D (0-30)	19 (5.9)
FHS	2,722	1,508 (53.6)	58.5 (11.6)	948 (34.8)	CES-D (0-20)	251 (16.1)
HBCS	122	0 (0)	65.2 (2.7)	24 (19.7)	CES-D (0-20)	11 (9.0)
KORA	1,727	882 (51.1)	61.0 (8.9)	250 (14.5)	PHQ-9	82 (4.7)
LBC 1921	432	261 (60.4)	79.1 (0.6)	194 (44.9)	HADS	15 (3.5)
LBC 1936	916	452 (49.3)	69.6 (0.8)	504 (55)	HADS	30 (3.3)
RS III	722	391 (54.2)	59.8 (8.1)	167 (23.1)	CES-D (0-20)	38 (5.3)
RS BIOS	757	319 (42.1)	67.6 (5.9)	78 (10.3)	CES-D (0-20)	51 (6.7)
GS	227	151 (64.5)	52.4 (8.1)	46 (19.7)	SCID¥	44 (18.8)
Total	7,948	4,158 (48.4)	65.4 (5.8)	2,384 (30.6)	-	541 (8.1)
Replication						
ARIC	2,297	1,445 (63)	56.1 (5.7)	584 (25.4)	21-MQ	74 (3.3)
WHI-EMPC	1,011	1,011 (100)	64.6 (7.1)	509 (50.3)	CES-D/DIS	61 (6.0)
Total	3,308	2,456 (74.2)	60.3 (6.4)	1,093 (37.9)	-	135 (4.7)

Characteristics are depicted as mean (SD), unless otherwise specified.

CHC Cardiovascular health cohort, ***FHS*** Framingham Heart Study, ***HBCS*** Helsinki Birth Cohort Study, ***KORA*** Cooperative Health Research in the Augsburg Region, ***LBC*** Lothian Birth Cohort, ***RS*** Rotterdam Study, ***GS*** Generation Scotland Study, and ¥CASE-CONTROL STUDY, ***ARIC*** Atherosclerosis Risk in Communities Study and ***WHI-EMPC*** the Women's Health Initiative - Epigenetic Mechanisms of PM-Mediated Cardiovascular disease***;*** "(in brackets we state number of item of the questionnaires)". * Represents former or current smoker

The replication sample included 3 308 participants from the Atherosclerosis Risk in Communities Study (ARIC)[37] and the Women's Health Initiative - Epigenetic Mechanisms of PM-Mediated Cardiovascular disease (WHI-EMPC).[38,39] More detailed information for each cohort can be found in the Supplementary Text.

Depressive symptoms assessment

Depressive symptoms were measured using self-reported questionnaires or structured interview performed by a trained researcher, psychologist, or psychiatrist at the same time point when blood samples were obtained for DNA methylation quantification (Table 1). Four cohorts (FHS, HBCS, RS-III, and RS-BIOS) assessed depressive symptoms using the 20-item Centre for Epidemiologic Studies Depression (CES-D) scale[40], while CHS used the 10-item CES-D scale. Participants could score from zero to 60 (or 30 for CHS) points, where higher scores indicate more depressive symptoms. WHI-EMPC used a cohort specific CES-D/DIS screening instrument, which is described in detail in the Supplementary Text. The LBC1921 and LBC1936 assessed self-reported depressive symptoms using the Hospital Anxiety and Depression Scale-depression subscale (HADS-D),[41] which consists of seven items. Participants could score from zero to 21. The KORA study used the self-administered Patient Health Questionnaire (PHQ-9)[42] representing a depression module that scores each of the nine Diagnostic and Statistical Manual of Mental Disorders, 4th Edition (DSM-IV) criteria for depression from zero to three. The GS study assessed life-time history of depression using the Structured Clinical Interview for DSM-IV Disorders (SCID).[43] The ARIC study assessed depressive symptoms using the 21-item Maastricht Questionnaire (21-MQ). In all cohorts depressive symptoms were analyzed as continuous variable except for GS, which studied depression status as binary trait.

DNA methylation sample and measurement

In all cohorts, DNA was extracted from whole blood and methylation levels were assessed using the Illumina-Infinium Human Methylation 450K BeadChip (Illumina Inc, San Diego, CA, USA) using standard manufacturer's protocols. The 450K array includes more than 450,000 CpGs and is enriched for genic regions, covering 99% of all genes. DNA methylation data pre-processing, including quality control (QC) and normalization, was conducted per cohort using study-specific methods. In all cohorts, DNA methylation levels were quantified as β-values, which range from zero to one, and indicate the proportion of DNA strands in a sample methylated at a specific CpG. Detailed information about cohort specific DNA extraction, bisulfite conversion, DNA methylation profiling, normalization and QC is described in detail in the Supplementary Text.

Gene expression

For the expression analyses, data from the RS-III cohort was used where whole-blood was collected (PAXGene Tubes – Becton Dickinson). All RNA samples with an RNA Quality Score > 7 were amplified, labelled (Ambion TotalPrep RNA), and hybridized to the Illumina HumanHT12v4 Expression Beadchips as described by the manufacturer's protocol. All samples were scanned on the Illumina iScan System (combined with an AutoLoader) using Illumina iScan image data acquisition software. Illumina GenomeStudio software (version 1.9.0) was used to generate output files for statistical analysis using R. In R Probes were quantile normalized to the median distribution, log2-transformed, and centered to the mean. Further, a sample Z transformation (mean = 0, standard deviation = 1) was applied. 44 877 probes measured over the whole genome. All probes that showed a detection *p value* < 0.05 in > 10% of samples were retained for the analysis (*N* = 21 238).[44]

DATA ANALYSIS

Epigenome-wide association analysis

In all cohorts, the association between depressive symptoms and CpG sites was assessed using linear regression analysis in the R software. In the regression analyses, DNA methylation β values at each CpG site was specified as the dependent variable and the depressive symptoms/depression as the predictor of interest. Further, as DNA methylation patterns significantly associate with age[45], sex[46] and smoking[47], the regression analysis included these as covariates in addition to methylation batch effects and white blood cell composition (imputed or directly measured, depending on availability within each cohort). All cohorts used principal components estimated using genome-wide genotype data to control for population stratification. Familial relationships were also accounted for in the model when applicable (FHS). Cohort specific details of these analyses are provided in the Supplementary Text.

It is unclear whether antidepressant medication use confounds the association between depressive symptoms and DNA methylation.[48,49] Therefore, we performed sensitivity analyses by adjusting the initial model additionally for antidepressant medication use at the time of sample collection.

Finally, we performed sample-size weighted meta-analysis of the cohort-specific results using METAL.[50] Due to the variability of available CpG sites after quality control steps, we excluded CpG sites missing in more than three of the participating cohorts. In total, 484 516 probes were available for meta-analysis. In addition to the gene annotation provided by Illumina based on

RefSeq database for the intergenic sites, the UCSC database (GRCh37/hg19) was explored to further annotate the CpGs. Bonferroni correction was applied to adjust for multiple testing giving a genome-wide significance threshold of 0.05/484 516 methylation sites = 1.03×10^{-7}. All CpG sites suggestive of association ($p\ value \leq 10^{-5}$) were tested for association in the independent replication cohort using the same model as used in the discovery EWAS. Finally, a sample size weighted meta-analysis was performed for all cohorts included in the discovery and replication phases in METAL.

Gene pathway analysis

Pathway analysis was performed for genes containing one or more nominally significant CpG sites from the meta-analysis of discovery and replication cohorts with PANTHER (see URLs). Genes that were included in the pathway analysis are provided in the Supplementary Table 1. Pathway analysis was performed using PANTHER[51] over-representation test using annotations from Reactome version 58 release 2016-12-07. In total 1 775 pathways were tested for over representation resulting in a Bonferroni corrected significance threshold of 0.05/1 775 = 2.82×10^{-05}.

Gene expression analysis

Linear regression analysis was used to associate identified methylation sites with each of the 21 238 expression probes using probes as dependent, CpG site as the independent variable, and age, sex, fasting state, blood cell counts and technical covariates (plate ID and RNA quality score) as covariates. Correction for multiple testing was performed using false discovery rate using Benjamini-Hochberg method implemented the stats library of R. Finally, pathway analysis of all significantly associated genes with CpG sites was performed in PANTHER.[51]

RESULTS

The characteristics of each individual cohort included in this meta-analysis are summarized in **Table 1.** The mean age in the discovery cohorts ranged from 52.4 years (*SD* = 8.1) in GS to 79.1 years (*SD* = 0.57) in LBC1921. Forty-eight percent of the total discovery sample were female. The replication cohort comprised 74% women and had an average age of 60.3 years (*SD* = 6.4) (**Table 1**).

Epigenome-wide association analysis

In the meta-analysis of depressive symptoms of European ancestry, we identified one CpG site on chromosome 14q32.32 (cg04987734, *CDC42BPB*, *p value* = 4.93 × 10^{-8}) that passed the Bonferroni threshold for significance (Table 2, Supplementary Figure 1). Suggestive association was observed at 19 additional CpG sites (Table 2). Adjusting for anti-depressive medication use did not meaningfully change the results (Supplementary Table 2). No inflation in the test statistic was observed (λ = 1.03, Supplementary Figure 2). We tested all 20 CpG sites for association in the replication sample. The top CpG site from the discovery sample (cg04987734) showed nominal association (*p value* = 0.048) with depressive symptoms in the validation data set (Table 2). In addition, significant association of a CpG site (cg12325605; *p value* = 9.17 ×10^{-05}, Table 2) annotated to the *ARHGEF3* gene with depressive symptoms was observed in the replication sample.

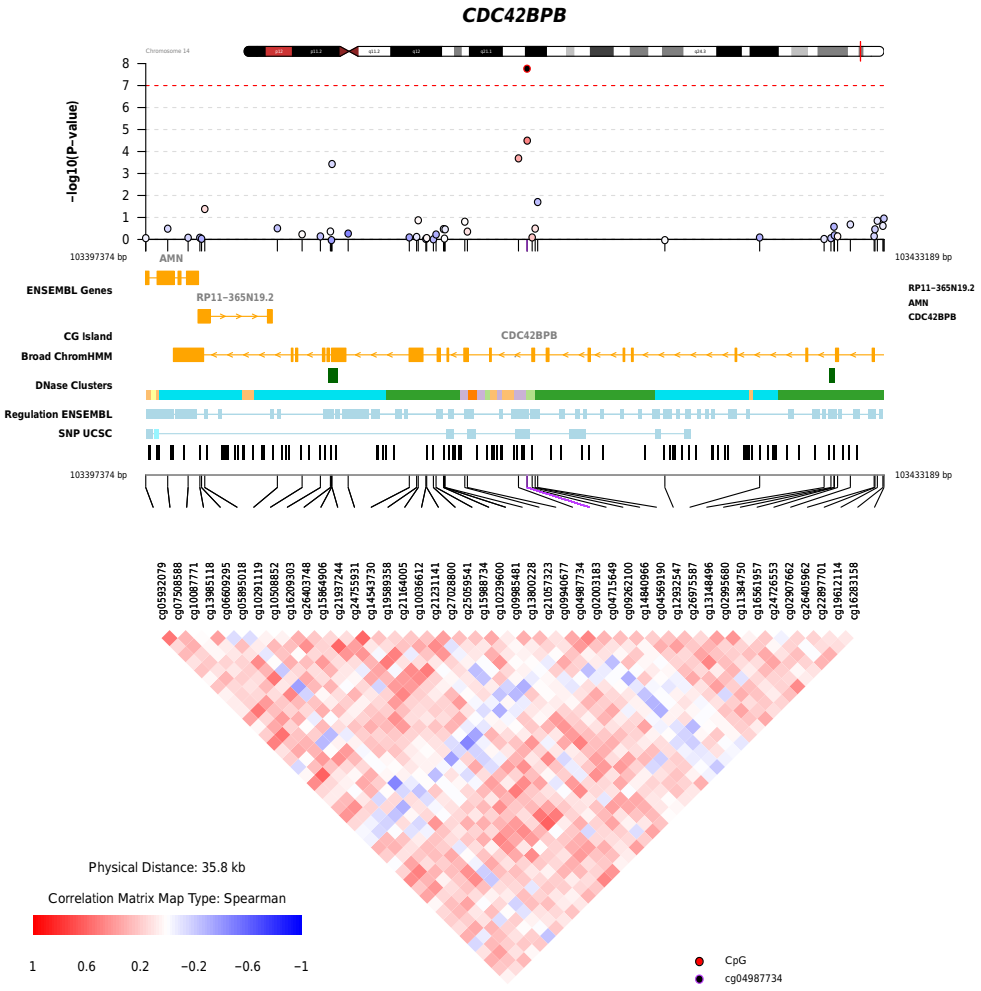
Meta-analysis of discovery and replication cohorts showed a significant association of both cg04987734 (*p value* = 1.51 × 10^{-08}) and cg12325605 (*p value* = 5.24 × 10^{-09}) with depressive symptoms (Table 2; Figures 1 & 2). Also, an additional intergenic CpG site (cg14023999; *p value* = 5.99 × 10^{-08}) at chromosome 15q26.1 locus showed genome-wide significant association with depressive symptoms (Supplementary Table 3). Top results from the meta-analysis of discovery and replication cohorts are provided in the Supplementary Table 3. Genome-wide results are illustrated in Supplementary Figures 3 & 4.

Table 2. Top DNA methylation sites associated with depressive symptoms in the discovery EWAS

				Discovery (*N* = 7 948)		Replication (*N*=3 308)		Meta-analysis (*N*=11 256)
CpG site ID	Chr	Location	Gene symbol	Direction[$]	*P value*	Direction[$]	*P value*	*P value*
cg04987734	14	103415873	*CDC42BPB*	++--+-+++	**4.93×10^{-8}**	+-	4.82×10^{-02}	**1.51×10^{-08}**
cg07012687	17	80195180	*SLC16A3*	+++-+++++	3.47×10^{-7}	+-	1.58×10^{-01}	4.45×10^{-06}
cg08796240	16	70733832	*VAC14*	+++++++++	7.43×10^{-7}	+-	2.56×10^{-01}	1.80×10^{-06}
cg06096336	2	231989800	*PSMD1;HTR2B*	++-++-+++	8.06×10^{-7}	++	3.01×10^{-01}	2.51×10^{-06}
cg16745930	10	100220809	*HPSE2*	---------	1.34×10^{-6}	--	4.01×10^{-01}	6.26×10^{-06}
cg09849319	5	1494983	*LPCATI1*	+++-+-+++	1.81×10^{-6}	+-	4.64×10^{-01}	1.04×10^{-04}
cg17237086	22	40814966	*MKL1*	+++++-+++	3.44×10^{-6}	+-	2.51×10^{-01}	6.10×10^{-06}
cg03985718	2	105924245	*TGFBRAP1*	+++-+-+++	3.61×10^{-6}	++	8.54×10^{-01}	6.53×10^{-05}
cg21098005	20	44538605	*PLTP*	+++++++++	4.36×10^{-6}	+-	9.60×10^{-01}	1.01×10^{-04}
cg16466652	19	6271960	*MLLT1*	+++-+++++	4.39×10^{-6}	+-	3.97×10^{-01}	1.57×10^{-05}
cg07884764	11	64107517	*CCDC88B*	+++-+++++	5.03×10^{-6}	+-	9.99×10^{-01}	1.25×10^{-04}
cg01541347	7	4729920	*FOXK1*	++++++++-	5.64×10^{-6}	--	3.77×10^{-01}	8.46×10^{-04}
cg02341197	21	34185927	*C21or62*	+++++-+++	5.84×10^{-6}	++	2.02×10^{-01}	6.80×10^{-06}
cg01947751	3	196728969	-	+++-+++++	6.23×10^{-6}	-+	6.63×10^{-01}	3.68×10^{-04}
cg13747876	17	80195402	*SLC16A3*	+++++-+++	6.32×10^{-6}	+-	1.04×10^{-01}	2.93×10^{-06}
cg12764201	1	105101123	*CORT;APITD1*	+++++-+++	7.15×10^{-6}	-+	7.20×10^{-01}	7.29×10^{-05}
cg08295111	5	133866097	*PHF15*	+++-+++++	7.87×10^{-6}	+-	5.76×10^{-01}	5.64×10^{-04}
cg18030453	3	45506216	*LARS2*	+++-+-+++	9.16×10^{-6}	+-	3.87×10^{-03}	1.20×10^{-07}
cg12325605	3	56810151	*ARHGEF3*	+++++++++	9.62×10^{-6}	++	**9.17×10^{-05}**	**5.24×10^{-09}**
cg23282441	10	73533927	*C10orf54;CDH23*	++--+++++	9.69×10^{-6}	+-	1.77×10^{-01}	8.63×10^{-06}

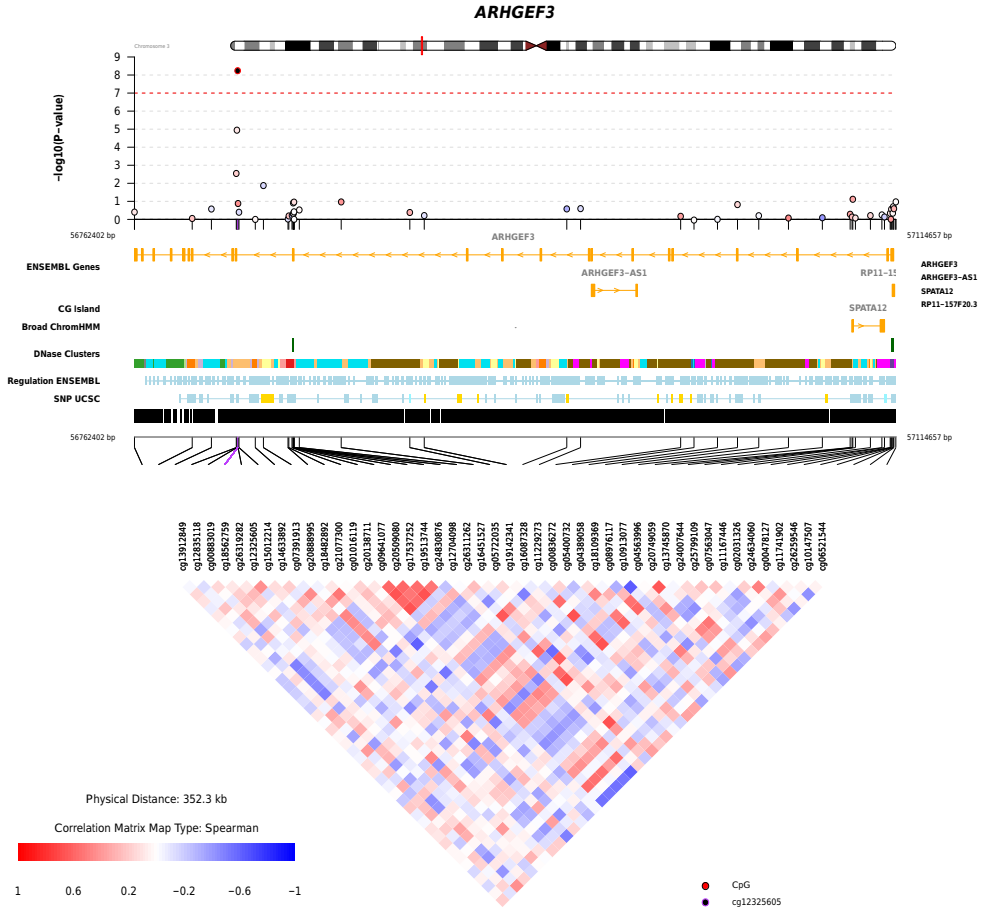
$ Order of cohorts in the direction column: Discovery meta-analysis: CHS, FS, GS, HBCS, KORA, LBC1921, LBC1936, RS_BIOS, RS_III; Replication meta-analysis: WHI-EMPC, ARIC; Discovery and replication samples meta-analysis: CHS, FS, GS, HBCS, KORA, LBC1921, LBC1936, RS_BIOS, RS_III, ARIC, and WHI-EMPC; + hyper-methylation and – hypo-methylation; Bolded *p values* refer to genome-wide statistical significance.

Figure 1: Regional association plot for the top CpG site, Genome 19 (cg04987734).



The horizontal axis depicts the position in base pair for the entire *CDC42BPB* gene region. The Vertical axis indicates the strength of association in terms of negative logarithm of the association *p-value*. Each circle represents CpG site. Red dashed line indicates the genome-wide significance threshold. Below the horizontal axis the figure shows the regulatory information and correlation matrix of other CpG sites in the region with the top hit. Color intensity marks the strength of the correlation and color the direction of the correlation.

Figure 2: Regional association plot for the top CpG site, Genome 19 (cg12325605).



The horizontal axis depicts the position in base pair for the entire *ARHGEF3* gene region. The Vertical axis indicates the strength of association in terms of negative logarithm of the association *p-value*. Each circle represents CpG site. Red dashed line indicates the genome-wide significance threshold. Below the horizontal axis the figure shows the regulatory information and correlation matrix of other CpG sites in the region with the top hit. Color intensity marks the strength of the correlation and color the direction of the correlation.

From blood to brain

Since the brain is the most relevant tissue for psychiatric disorders like depression, we checked the correlation between methylation in blood and various brain regions at the three identified sites using a web-based tool, BECon.[25] Methylation at cg04987734 in the CDC42BPB gene was highly correlated (r = 0.81) between blood and the Brodmann area 7 that spans the medial and lateral walls of the parietal cortex (Supplementary Figure 5). Methylation at the other two sites was negatively correlated with methylation in the Brodmann area 10 than spans anterior prefrontal cortex (cg12325605, r = -0.39; cg14023999, r = -0.42) suggesting strong but reverse methylation patterns in blood and brain (Supplementary Figures 6 & 7).

Gene pathway analysis

Pathway analysis of all genes annotated to at least one nominally significant CpG site ($p < 0.05$) showed significant over-representation of genes involved in axon guidance (1.41 fold increase, *p value* = 5.2×10^{-07}), nerve growth factor (NGF) signaling (1.42 fold increase, *p value* = 1.31×10^{-05}), platelet-derived growth factor (PDGF) signaling (1.47 fold increase, *p value* = 4.5×10^{-09}) and developmental biology (1.32 fold increase, *p value* = 3.95×10^{-06}) (Supplementary Table 4).

Gene expression analysis

Methylation at cg04987734 was significantly associated with increased expression of four genes (Supplementary Table 5). Methylation at cg12325605 was significantly associated with differential expression of 422 genes (Supplementary Table 6) and methylation at cg14023999 was significantly associated with differential expression of two genes (Supplementary Table 7). Pathway analysis of the 422 differentially expressed genes revealed only one significant pathway: immune system (fold enrichment = 2.56, *p value* = 6.93×10^{-10}) and more specifically the adaptive immune system (fold enrichment = 2.99, *p value* = 7.37×10^{-07}) rather than the innate immune system (fold enrichment = 2.53, *p value* = 1.77×10^{-03}) (Supplementary Table 8). Results were more significant when we added the other six genes that were differentially associated with cg04987734 and cg14023999 (Supplementary Table 9).

DISCUSSION

In this large-scale EWAS of depressive symptoms, we identified hypermethylation of three CpG sites (cg04987734, cg12325605 and cg14023999) associated with depressive symptoms. Cg04987734 is annotated to the *CDC42BPB* gene, cg12325605 to the *ARHGEF3* gene, and cg14023999 lies in an intergenic region on chromosome 15q26.1 locus. Pathway analyses of genes harboring all nominally associated methylation sites suggest over representation of genes

involved in axon guidance, nerve growth factor (NGF) and platelet-derived growth factor (PDGF) signaling. The three CpG sites were significantly associated with the differential expression of 428 genes. Pathway analysis of the differentially expressed genes show significant enrichment of genes involved in regulating the immune system.

CDC42BPB encodes a serine/threonine protein kinase that plays a role in the regulation of cytoskeleton reorganization, cell migration and regulation of neurite outgrowth[52] and is highly expressed in the brain. Further, gene co-expression network suggests that it is involved in the synthesis, recycling, salvage and regulation of tetrahydrobiopterin (BH4) and cytotoxic T-lymphocyte-associated protein 4 (CTLA4) inhibitory signaling (see URLs). BH4 is an essential co-factor for the aromatic-amino-acid-hydroxylases and glyceryl-ether mono-oxygenase and regulates nitric-oxide synthase (NOS) activity.[53] BH4 deficiency leads to dopamine and serotonin deficiency in the brain[54], a characteristic pathophysiological hallmark of depression.[55] BH4 has also been reported as an intrinsic regulator of inflammatory pain sensitivity and chronicity in animal studies[56] and implicated in immune response.[57] Inflammation-induced increases in inducible NOS activity can usurp available BH4, resulting in NOS uncoupling and increased oxidative stress affecting dopamine synthesis.[58] Further, CTLA4 is a member of the immunoglobulin superfamily that is expressed by activated T-cells[59] and soluble CTLA4 has been suggested as a biomarker for inflammation.[60] Interestingly, methylation levels at this CpG site (cg04987734) in *CDC42BPB* gene were also previously associated with C-reactive protein (CRP) levels in blood[61]; a marker for inflammation. Hyper-methylation of cg04987734 has been associated with increased expression of *CDC42BPB* in blood (see URLs).[62] Consistent with such an involvement in inflammatory processes, hyper-methylation of cg04987734 was also identified as an epigenetic marker associated with cigarette smoking.[47] Life-style factors such as smoking and alcohol consumption reflect additional environmental effects that may induce epigenetic changes. In our study, however, we controlled for smoking in the regression model; therefore the association between depression and DNA methylation of this CpG site may be independent of smoking habits.

ARHGEF3 encodes for Rho Guanine Nucleotide Exchange Factor 3 protein. The gene is highly expressed in adrenal glands, brain and uterus (see URLs). Genetic variants in *ARHGEF3* have been associated with platelet volume (see URLs). Functional analysis of *ARHGEF3* (see URLs) suggests that this gene is involved in the regulation of innate immune response and sphingolipid metabolism. Earlier we identified rare missense variants in a gene *NKPD1*[11], a gene also predicted to be involved in de novo synthesis of sphingolipids, and plasma sphingolipid levels associated with depressive symptomology.[63]

The third associated CpG site, cg14023999 lies in an intergenic region on chromosome 15q26.1. Hypermethylation of cg14023999 is associated with decreased expression of *SEMA4B* gene in

blood (see URLs). *SEMA4B* gene encodes for Semaphorin 4B protein. Semaphorins are implicated in axon guidance, regulation of cell migration, angiogenesis and immune response.[64-68] Sema4B is believed to function through a direct interaction with post-synaptic density protein PSD-95[69-71] to promote synapse maturation.[69,72,73] The knock-down of Sema4B causes a decrease in GABAergic synapse number[72] suggesting a role in the assembly of excitatory and inhibitory postsynaptic specializations.[73]

Interestingly, all the three top hits are connected directly or through regulatory effects to genes involved in the immune pathway. Further, the pathway analysis of all the differentially expressed genes associated with these three top hits point towards the immune pathway. Together these findings are compatible with the hypothesis that environmental stressors may trigger depression by inducing elevations in pro-inflammatory cytokines[16]. Testing this hypothesis, however, requires further work, as our cross-sectional design does not allow us to establish the direction of association. We can only carefully speculate, whether the epigenetic factors mediate the relation between some environmental exposure(s) and depression, whether the epigenetic factors mediate the relation between the genome and depression; or whether the epigenetic variants are biomarkers of depression but not causally related to the disorder. In contrast to the expression data of our top hits, in the formal pathway analysis of the methylation data, we did not observe a significant overrepresentation of genes involved in the immune pathway (see Supplementary Table 4). Rather, these pathway analysis suggested significant enrichment of genes involved in axon-guidance and growth factor signaling.

This is the largest epigenome-wide study of depressive symptoms reported to date. Our major strength is the sample size that enabled detection of a replicable epigenome-wide significant locus, which suggests that in blood, DNA methylation signatures associated with depression may be subtle and will require large samples to be detected. Using peripheral blood tissue for DNA methylation profiling is a limitation of this study, as DNA methylation is known to be tissue specific.[20] While peripheral blood is not considered to be the most relevant tissue for the pathophysiology of depression, some sites show correlated methylation profiles between-tissues.[20,21] The three sites identified in our study show moderate to high correlation between methylation in blood and various brain regions. Previous studies have also reported epigenetic changes with similar effects in both peripheral tissue and brain.[48] Secondly, we replicated our top hits in the African-American samples. This suggests that depression related differences in DNA methylation may be similar across ethnicities.[74-76] However, the use of a multi-ethnic predominantly African-American replication sample may also have resulted in false negatives due to different genetic background. Third, in these analysis we mostly used quantitative measures of depressive symptoms. Quantitative endo-phenotypes provide powerful alternatives for several complex outcomes, for example, hypertension.[77] This is likely to be especially true for a trait such as

depressive symptoms, for which the severity and duration of illness can be highly heterogeneous.[78] Studies that use depression as an arbitrary selected extreme of the continuum of varying severity, are highly susceptible to misclassification, resulting in the loss of the statistical power. Genome-wide studies of depressive traits, using quantitative endo-phenotypes, have been suggested to improve statistical power.[78] The majority of the cohorts used the CES-D depression scale to assess depressive symptoms in our study. However, the use of different phenotypic measures by different samples typically implies some loss of statistical power due to the heterogeneity in the phenotype assessment. Finally, although we adjusted for potential confounders, the possibility of residual confounding cannot be excluded. Antidepressant medication indicate treated depression but itself may result in epigenetic modifications involved in depression pathophysiology.[49] Antidepressants can thus mediate or confound the relation between DNA methylation and depression. However, in sensitivity analysis additionally adjusted for antidepressant medication, our results did not change.

To summarize, we report the first EWAS of depressive symptoms. We identified hypermethylation of three sites in the genome significantly associated with depressive symptoms. The identified loci are implicated in axon guidance, nerve growth factor, regulation of serotonin and dopamine levels in the brain, sphingolipid metabolism, synaptic plasticity and inflammation. Our findings provide new insights into the molecular mechanisms underlying the complex pathophysiology of depression.

URLs. The expression of genes in various brain regions is available through the genome browser: http://genenetwork.org/webqtl/main.py. The associations between methylation of cg04987734 with the expression of the *CDC42BPB* gene is available at http://genenetwork.nl/biosqtlbrowser/. The expression of the *ARHGEF3* gene is available at https://www.gtexportal.org/home/gene/ARHGEF3#geneExpression. The association between gene variants of ARHGEF3 and platelet activation is available at https://www.ebi.ac.uk/gwas/search?query=arhgef3. Functional analysis of *ARHGEF3* are available at http://129.125.135.180:8080/GeneNetwork/?gene=ARHGEF3. Association between methylation of cg14023999 with expression of *SEMA4B* gene in blood is available at http://genenetwork.nl/biosqtlbrowser/. Pathway analysis was performed using the web based tool PANTHER (www.pantherdb.org). Correlation between methylation in blood and brain was assessed using BECon (https://redgar598.shinyapps.io/BECon/).

REFERENCES

1. Li, D., Zhang, D.-j., Shao, J.-j., Qi, X.-d. & Tian, L. A meta-analysis of the prevalence of depressive symptoms in Chinese older adults. *Archives of Gerontology and Geriatrics* 58, 1-9 (2014).
2. Volkert, J., Schulz, H., Härter, M., Wlodarczyk, O. & Andreas, S. The prevalence of mental disorders in older people in Western countries – a meta-analysis. *Ageing Research Reviews* 12, 339-353 (2013).
3. Alexopoulos, G.S. Depression in the elderly. *Lancet* 365, 1961-70 (2005).
4. Lepine, J.P. & Briley, M. The increasing burden of depression. *Neuropsychiatr Dis Treat* 7, 3-7 (2011).
5. Hidaka, B.H. Depression as a disease of modernity: explanations for increasing prevalence. *J Affect Disord* 140, 205-14 (2012).
6. Kessler, R.C. The effects of stressful life events on depression. *Annu Rev Psychol* 48, 191-214 (1997).
7. Lohoff, F.W. Overview of the genetics of major depressive disorder. *Curr Psychiatry Rep* 12, 539-46 (2010).
8. Saavedra, K., Molina-Marquez, A.M., Saavedra, N., Zambrano, T. & Salazar, L.A. Epigenetic Modifications of Major Depressive Disorder. *Int J Mol Sci* 17(2016).
9. Hyde, C.L. *et al.* Identification of 15 genetic loci associated with risk of major depression in individuals of European descent. *Nat Genet* 48, 1031-6 (2016).
10. Amin, N. *et al.* Exome-sequencing in a large population-based study reveals a rare Asn396Ser variant in the LIPG gene associated with depressive symptoms. *Mol Psychiatry* (2016).
11. Amin, N. *et al.* Nonsynonymous Variation in NKPD1 Increases Depressive Symptoms in the European Populations. *Biol Psychiatry* (2016).
12. Amin, N. *et al.* A rare missense variant in RCL1 segregates with depression in extended families. *Mol Psychiatry* (2017).
13. Agid, O., Kohn, Y. & Lerer, B. Environmental stress and psychiatric illness. *Biomed Pharmacother* 54, 135-41 (2000).
14. Kinderman, P., Schwannauer, M., Pontin, E. & Tai, S. Psychological processes mediate the impact of familial risk, social circumstances and life events on mental health. *PLoS One* 8, e76564 (2013).
15. Kendler, K.S., Karkowski, L.M. & Prescott, C.A. Causal relationship between stressful life events and the onset of major depression. *Am J Psychiatry* 156, 837-41 (1999).
16. Berk, M. *et al.* So depression is an inflammatory disease, but where does the inflammation come from? *BMC Med* 11, 200 (2013).
17. Dahl, C. & Guldberg, P. DNA methylation analysis techniques. *Biogerontology* 4, 233-50 (2003).
18. Rakyan, V.K., Blewitt, M.E., Druker, R., Preis, J.I. & Whitelaw, E. Metastable epialleles in mammals. *Trends Genet* 18, 348-51 (2002).
19. Bind, M.A. *et al.* Air pollution and markers of coagulation, inflammation, and endothelial function: associations and epigene-environment interactions in an elderly cohort. *Epidemiology* 23, 332-40 (2012).
20. Bagot, R.C., Labonte, B., Pena, C.J. & Nestler, E.J. Epigenetic signaling in psychiatric disorders: stress and depression. *Dialogues Clin Neurosci* 16, 281-95 (2014).
21. Lokk, K. *et al.* DNA methylome profiling of human tissues identifies global and tissue-specific methylation patterns. *Genome Biol* 15, r54 (2014).
22. Byrne, E.M. *et al.* Monozygotic twins affected with major depressive disorder have greater variance in methylation than their unaffected co-twin. *Transl Psychiatry* 3, e269 (2013).
23. Sabunciyan, S. *et al.* Genome-wide DNA methylation scan in major depressive disorder. *PLoS One* 7, e34451 (2012).
24. Latham, K.E. Stage-specific and cell type-specific aspects of genomic imprinting effects in mammals. *Differentiation* 59, 269-82 (1995).

25. Edgar, R.D., Jones, M.J., Meaney, M.J., Turecki, G. & Kobor, M.S. BECon: A tool for interpreting DNA methylation findings from blood in the context of brain. *bioRxiv* (2017).
26. Bakusic, J., Schaufeli, W., Claes, S. & Godderis, L. Stress, burnout and depression: A systematic review on DNA methylation mechanisms. *J Psychosom Res* 92, 34-44 (2017).
27. Januar, V., Saffery, R. & Ryan, J. Epigenetics and depressive disorders: a review of current progress and future directions. *Int J Epidemiol* 44, 1364-87 (2015).
28. Nagy, C. *et al.* Astrocytic abnormalities and global DNA methylation patterns in depression and suicide. *Mol Psychiatry* 20, 320-8 (2015).
29. Fried, L.P. *et al.* The Cardiovascular Health Study: design and rationale. *Ann Epidemiol* 1, 263-76 (1991).
30. Kannel, W.B., Feinleib, M., McNamara, P.M., Garrison, R.J. & Castelli, W.P. An investigation of coronary heart disease in families. The Framingham offspring study. *Am J Epidemiol* 110, 281-90 (1979).
31. Eriksson, J.G., Forsen, T., Tuomilehto, J., Osmond, C. & Barker, D.J. Early growth and coronary heart disease in later life: longitudinal study. *BMJ* 322, 949-53 (2001).
32. Wichmann, H.E., Gieger, C., Illig, T. & Group, M.K.S. KORA-gen--resource for population genetics, controls and a broad spectrum of disease phenotypes. *Gesundheitswesen* 67 Suppl 1, S26-30 (2005).
33. Deary, I.J., Gow, A.J., Pattie, A. & Starr, J.M. Cohort profile: the Lothian Birth Cohorts of 1921 and 1936. *Int J Epidemiol* 41, 1576-84 (2012).
34. Deary, I.J. *et al.* The Lothian Birth Cohort 1936: a study to examine influences on cognitive ageing from age 11 to age 70 and beyond. *BMC Geriatr* 7, 28 (2007).
35. Hofman, A. *et al.* The Rotterdam Study: 2016 objectives and design update. *Eur J Epidemiol* 30, 661-708 (2015).
36. Smith, B.H. *et al.* Cohort Profile: Generation Scotland: Scottish Family Health Study (GS:SFHS). The study, its participants and their potential for genetic research on health and illness. *Int J Epidemiol* 42, 689-700 (2013).
37. The Atherosclerosis Risk in Communities (ARIC) Study: design and objectives. The ARIC investigators. *Am J Epidemiol* 129, 687-702 (1989).
38. Design of the Women's Health Initiative clinical trial and observational study. The Women's Health Initiative Study Group. *Control Clin Trials* 19, 61-109 (1998).
39. Whitsel, E.A. AS315 - Epigenetic Mechanisms of PM-mediated CVD Risk.
40. Lewinsohn, P.M., Seeley, J.R., Roberts, R.E. & Allen, N.B. Center for Epidemiologic Studies Depression Scale (CES-D) as a screening instrument for depression among community-residing older adults. *Psychol Aging* 12, 277-87 (1997).
41. Zigmond, A.S. & Snaith, R.P. The hospital anxiety and depression scale. *Acta Psychiatr Scand* 67, 361-70 (1983).
42. Kroenke, K., Spitzer, R.L. & Williams, J.B. The PHQ-9: validity of a brief depression severity measure. *J Gen Intern Med* 16, 606-13 (2001).
43. First, M.B., Spitzer, Robert L, Gibbon Miriam, and Williams, Janet B.W. Structured Clinical Interview for DSM-IV-TR Axis I Disorders, Research Version, Patient Edition. (SCID-I/P) (November 2002.).
44. Peters, M.J. *et al.* The transcriptional landscape of age in human peripheral blood. *Nat Commun* 6, 8570 (2015).
45. Horvath, S. DNA methylation age of human tissues and cell types. *Genome Biol* 14, R115 (2013).
46. Singmann, P. *et al.* Characterization of whole-genome autosomal differences of DNA methylation between men and women. *Epigenetics Chromatin* 8, 43 (2015).
47. Joehanes, R. *et al.* Epigenetic Signatures of Cigarette Smoking. *Circ Cardiovasc Genet* 9, 436-447 (2016).
48. Menke, A. & Binder, E.B. Epigenetic alterations in depression and antidepressant treatment. *Dialogues Clin Neurosci* 16, 395-404 (2014).

49. Lisoway, A.J., Zai, C.C., Tiwari, A.K. & Kennedy, J.L. DNA Methylation and Clinical Response to Antidepressant Medication in Major Depressive Disorder: A Review and Recommendations. *Neurosci Lett* (2017).
50. Willer, C.J., Li, Y. & Abecasis, G.R. METAL: fast and efficient meta-analysis of genomewide association scans. *Bioinformatics* 26, 2190-1 (2010).
51. Mi, H. & Thomas, P. PANTHER pathway: an ontology-based pathway database coupled with data analysis tools. *Methods Mol Biol* 563, 123-40 (2009).
52. Chen, X.Q., Tan, I., Leung, T. & Lim, L. The myotonic dystrophy kinase-related Cdc42-binding kinase is involved in the regulation of neurite outgrowth in PC12 cells. *J Biol Chem* 274, 19901-5 (1999).
53. Taguchi, H. & Armarego, W.L. Glyceryl-ether monooxygenase [EC 1.14.16.5]. A microsomal enzyme of ether lipid metabolism. *Med Res Rev* 18, 43-89 (1998).
54. Scriver, C.R., Eisensmith, R.C., Woo, S.L. & Kaufman, S. The hyperphenylalaninemias of man and mouse. *Annu Rev Genet* 28, 141-65 (1994).
55. Belmaker, R.H. & Agam, G. Major depressive disorder. *N Engl J Med* 358, 55-68 (2008).
56. Tegeder, I. *et al.* GTP cyclohydrolase and tetrahydrobiopterin regulate pain sensitivity and persistence. *Nat Med* 12, 1269-77 (2006).
57. McNeill, E. & Channon, K.M. The role of tetrahydrobiopterin in inflammation and cardiovascular disease. *Thromb Haemost* 108, 832-9 (2012).
58. Felger, J.C. The Role of Dopamine in Inflammation-Associated Depression: Mechanisms and Therapeutic Implications. *Curr Top Behav Neurosci* 31, 199-219 (2017).
59. Wulfing, C., Tunbridge, H.M. & Wraith, D.C. New inhibitory signaling by CTLA-4. *Nat Immunol* 15, 408-9 (2014).
60. Sakthivel, P. *et al.* Circulating soluble CTLA-4 is related to inflammatory markers in the 70 year old population. *Scand J Clin Lab Invest* 70, 237-43 (2010).
61. Ligthart, S. *et al.* DNA methylation signatures of chronic low-grade inflammation are associated with complex diseases. *Genome Biol* 17, 255 (2016).
62. Bonder, M.J. *et al.* Disease variants alter transcription factor levels and methylation of their binding sites. *Nat Genet* 49, 131-138 (2017).
63. Demirkan, A. *et al.* Plasma phosphatidylcholine and sphingomyelin concentrations are associated with depression and anxiety symptoms in a Dutch family-based lipidomics study. *J Psychiatr Res* 47, 357-62 (2013).
64. Fiore, R. & Puschel, A.W. The function of semaphorins during nervous system development. *Front Biosci* 8, s484-99 (2003).
65. Giordano, S. *et al.* The semaphorin 4D receptor controls invasive growth by coupling with Met. *Nat Cell Biol* 4, 720-4 (2002).
66. Kumanogoh, A. & Kikutani, H. Immune semaphorins: a new area of semaphorin research. *J Cell Sci* 116, 3463-70 (2003).
67. Tamagnone, L. & Comoglio, P.M. To move or not to move? Semaphorin signalling in cell migration. *EMBO Rep* 5, 356-61 (2004).
68. Serini, G. *et al.* Class 3 semaphorins control vascular morphogenesis by inhibiting integrin function. *Nature* 424, 391-7 (2003).
69. Burkhardt, C. *et al.* Semaphorin 4B interacts with the post-synaptic density protein PSD-95/SAP90 and is recruited to synapses through a C-terminal PDZ-binding motif. *FEBS Lett* 579, 3821-8 (2005).
70. Inagaki, S. *et al.* Sema4c, a transmembrane semaphorin, interacts with a post-synaptic density protein, PSD-95. *J Biol Chem* 276, 9174-81 (2001).
71. Schultze, W. *et al.* Semaphorin4F interacts with the synapse-associated protein SAP90/PSD-95. *J Neurochem* 78, 482-9 (2001).

72. Paradis, S. *et al.* An RNAi-based approach identifies molecules required for glutamatergic and GABAergic synapse development. *Neuron* 53, 217-32 (2007).
73. Pasterkamp, R.J. & Giger, R.J. Semaphorin function in neural plasticity and disease. *Curr Opin Neurobiol* 19, 263-74 (2009).
74. Klengel, T., Pape, J., Binder, E.B. & Mehta, D. The role of DNA methylation in stress-related psychiatric disorders. *Neuropharmacology* 80, 115-32 (2014).
75. Galanter, J.M. *et al.* Differential methylation between ethnic sub-groups reflects the effect of genetic ancestry and environmental exposures. *Elife* 6(2017).
76. Kader, F. & Ghai, M. DNA methylation-based variation between human populations. *Mol Genet Genomics* 292, 5-35 (2017).
77. International Consortium for Blood Pressure Genome-Wide Association, S. *et al.* Genetic variants in novel pathways influence blood pressure and cardiovascular disease risk. *Nature* 478, 103-9 (2011).
78. van der Sluis, S., Posthuma, D., Nivard, M.G., Verhage, M. & Dolan, C.V. Power in GWAS: lifting the curse of the clinical cut-off. *Mol Psychiatry* 18, 2-3 (2013).

APPENDIX of CHAPTER 4

Supplementary Text

Cardiovascular Health Study Cohort

Description

The Cardiovascular Health Study Cohort (CHS) is a population-based cohort study of risk factors for coronary heart disease and stroke in adults ≥ 65 years conducted across four field centers.[1] The original predominantly European ancestry cohort of 5 201 persons was recruited in 1989 - 1990 from random samples of the Medicare eligibility lists. CHS was approved by institutional review committees at each field center and individuals in the present analysis had available DNA and gave informed consent including consent to use of genetic information for the study of cardiovascular disease.

DNA methylation was measured on a randomly selected subset of 323 European descent participants from study year. The samples were randomly selected among participants without presence of coronary heart disease, congestive heart failure, peripheral vascular disease, valvular heart disease, stroke or transient ischemic attack at study baseline or lack of available DNA at study year 5.

Methylation measurements were performed at the Institute for Translational Genomics and Population Sciences at the Harbor-UCLA Medical Center Institute for Translational Genomics and Population Sciences using the Infinium HumanMethylation450 BeadChip (Illumina Inc, San Diego, CA). Quality control was performed in in the minfi R package.[2-4] (version 1.12.0, http://www.bioconductor.org/packages/release/bioc/html/minfi.html). Samples with low median intensities of below 10.5 (log2) across the methylated and un-methylated channels, samples with a proportion of probes falling detection of greater than 0.5%, samples with QC probes falling greater than 3 standard deviation from the mean, sex check mismatches, or failed concordance with prior genotyping were removed. Methylation values were normalized using the SWAN quantile normalization method.[3] Since white blood cell proportions were not directly measured in CHS they were estimated from the methylation data using the Houseman method.[5] All association analyses were performed in R using linear mixed models with DNA methylation beta values as the outcome. Analyses adjusted for age, gender, current smoking status and antidepressant medication. Additionally chip, chip row and column were adjusted for as random effects.
CHS was approved by institutional review committees at each field center and individuals in the present analysis had available DNA and gave informed consent including consent to use of genetic information for the study of cardiovascular disease.

Framingham Heart Study

Description

The Framingham Heart Study (FHS) is a population-based, prospective study. In 1948, the Original cohort of 5 209 individuals were recruited from Framingham, MA.[6] The offspring cohort was recruited in 1971, including 5 124 offspring and spouses of offspring of the FHS Original cohort.[7] Offspring participants underwent examinations every four years (except eight years between the first and the second examinations).[7] Peripheral whole blood samples were collected from these participants at the eighth examination (2005 - 2008). Buffy coat fractions were obtained and genomic DNA was extracted using the Gentra Puregene DNA extraction kit (Qiagen, Venlo, Netherlands). Bisulfite conversion of genomic DNA was performed with the EZ DNA Methylation Kit (Zymo Research, Irvine, CA). DNA methylation was measured by *Infinium HumanMethylation450 BeadChip* (Illumina, CA) in two separate labs. All participants provided written consent for genetic study.

QC procedures were performed. Bad samples were excluded if the samples were outliers in multidimensional scaling (MDS) analysis, high missing rate (> 1%), poor matching to SNP genotype. Low quality probes were excluded if these probes had high missing rate (> 20%), were mapped to multiple locations, had SNP (MAF > 5% in EUR 1000G) at CpG site or ≤ 10 bp of Single Base Extension.[8] About 440,000 probes remain in ~2,600 FHS participants. A total of 1 680 Offspring participants had both DNA methylation and CES-D phenotype. The sample characteristics are displayed in Table 1 in the manuscript.

Analysis

Two models of association were tested. In order to reduce the batch effects, we analyzed the DNA methylation in each lab: residuals were obtained by regressing a CpG against the batch effects (plate, row and col numbers) and lab-specific principle components that were related to the outcome. Residuals from two labs were put together and then regressed against the CES-D, adjusting for covariates and family structure. Plate ID and family structures were treated as random effects in linear mixed models.

Model 1. Resid (CpG ~ Batch effects + PCs) (lab-specific)
Pooled residuals from two labs ~ CES-D + Gender + Age + smoking + WBC + family structure

Model 2. Resid (CpG ~ Batch effects + PCs) (lab-specific)
Pooled residuals from two labs ~ CES-D + Gender + Age + smoking + WBC + deprx8 + family structure

Helsinki Birth Cohort Study (HBCS)

Description

The Helsinki Birth Cohort Study (HBCS) comprises 13 345 individuals (6 370 women and 6 975 men), born as singletons between 1934 and 1944 in one of the two main maternity hospitals in Helsinki and who were living in Finland in 1971 when a unique personal identification number was allocated to each member of the Finnish population. The HBCS, which has been described in detail elsewhere, has been approved by the Ethics Committee of the National Public Health Institute. Register data were linked with permission from the Finnish Ministry of Social Affairs and Health and the Finnish National Archives.

In 2001 – 2004 at an average age of 61.5 years (*SD* = 2.9 and range = 56.7 – 69.8 years), a randomly selected subsample of the cohort comprising 2 003 individuals (1 075 women and 928 men) was invited to a clinical examination including collection of a blood sample for genetic, epigenetic and biochemical studies and a psychological survey including a measure of depressive symptoms. For 283 participants, extraction of DNA was not successful, or DNA showed gender discrepancy or close relatedness. From the remaining sample of 1 720 individuals, 115 women and 97 men had been evacuated to Sweden or Denmark during the World War II according to the Finnish National Archives' register. To study the effects of early separation from parents on methylome, we selected 83 evacuated men with data on age at and length of evacuation and of father's occupational status and 83 non-evacuated controls matched for sex, birth year and father's occupational status in childhood for the methylation typing. In this study, the analyses are based on 62 evacuated men and 60 non-evacuated controls with full data on CES-D depressive symptoms, covariates, and methylation profiles. For this group the mean age was 63.5 years (*SD* = 2.8).

DNA methylation sample, measurement, normalization and quality control

DNA methylation analysis was performed at the Genetics Core of the Wellcome Trust Clinical Research Facility (Edinburgh, UK). Bisulphite conversion of 500 ng input DNA was carried out using the EZ DNA Methylation Kit (Zymo Research, Freiburg, Germany). Four microlitres of bisulphite-converted DNA was processed using the Infinium HD Assay for Methylation. This was performed using the Illumina Methylation 450 k beadchip and Infinium chemistry (Illumina, Inc., San Diego, CA, USA). Each sample was interrogated on the arrays against 485 000 methylation sites. The arrays were imaged on the Illumina HiScan platform and genotypes were called automatically using GenomeStudio Analysis software version 2011.1.

Quality control pipeline was set up using the R-package minfi, including intensity read outs, normalization, cell type composition estimation, β- and M-value calculation. We excluded samples with low intensity (badSampleCutoff < 10.6) or deviant beta distribution based on visual inspection (N = 5). We did not detect any gender discrepancy. Data were normalized with functional normalization (FunNorm). Of the probes, we excluded those with detection p value > 0.01 in > 50% of samples[9], non-autosomal and non-specific binding probes as well as probes with SNPs in the interval for which the Illumina probe is designed to hybridize and if they were located close (10bp from query site) to a SNP which had a minor allele frequency of ≥ 0.05. Probes located in the X and Y chromosome were also excluded. This yielded a total number of probes of 424 844. Batch effects were identified by inspecting the association of principal components of the methylation levels with possible technical batches using linear regressions and visual inspection of PCA plots using the Bioconductor R package *shinyMethyl* (version 0.99.3). Identified batch effects (i.e. array column) were removed using the Empirical Bayes' (EB) method *ComBat*[10]. Batch corrected M-values after *ComBat* were used for all further statistical analyses.

Depressive symptoms assessment

Frequency of depressive symptoms were self-reported with the 20-item Centre for Epidemiologic Studies Depression scale (CES-D).

Covariates

Age, sex, current smoking status, evacuation status (evacuated in childhood or not), and antidepressant medication (Model 2) were self-reported at the clinical visit. Blood cell proportions were estimated with the Houseman method. First three genotype MDS components were derived from genome-wide genotype data.

Analysis

A linear regression of the CES-D depressive symptoms score and covariates on each individual methylation probe in the array was performed in R.

Cooperative Health Research in the Augsburg Region

Description

The Cooperative Health Research in the Region of Augsburg (KORA) study is an independent population-based cohort from the region of Augsburg, Southern Germany. Whole blood samples were obtained from the KORA F4 survey (examination 2006-2008), a seven-year follow-up study of the KORA S4 cohort. Participants gave written informed consent and the study was approved by the local ethics committee (Bayerische Landesärztekammer).

DNA methylation sample, measurement, normalization and quality control

Whole blood genomic DNA was bisulfite converted using the EZ-96 DNA Methylation Kit (Zymo Research, Orange, CA, USA) according to the manufacturer's procedure, with the alternative incubation conditions recommended when using the Illumina Infinium Methylation Assay. Raw methylation data were extracted using the Illumina GenomeStudio software (version 2011.1, Methylation module 1.9.0). Preprocessing was performed using R (version 3.0.1). Probes with signals from less than three functional beads, and probes with a detection p-value > 0.01 were defined as low-confidence probes. Probes that covered SNPs (MAF in Europeans > 5%) were excluded from the data set. A color bias adjustment was performed using the R package lumi (version 2.12.0) through smooth quantile normalization and background correction based on the negative control probes present on the Infinium HumanMethylation BeadChip. This was performed separately for the two color channels and chips. β-values corresponding to low-confidence probes were set to missing. A 95% call rate threshold was applied on samples and CpG sites. Beta-mixture quantile normalization (BMIQ) was applied using the R package wateRmelon, version 1.0.3. Because KORA F4 samples were processed on 20 96-well plates in 9 batches, plate and batch effects were investigated using principal component analysis and eigenR2 analysis.

Depressive symptoms assessment

Depressive symptoms were assessed using the German version of the self-administered Patient Health Questionnaire (PHQ-9).[11] PHQ-9 represents a depression module that scores each of the nine Diagnostic and Statistical Manual of Mental Disorders, 4th Edition (DSM-IV) criteria for depression from zero to three. PHQ-9 scores of 5, 10, 15, and 20 represents mild, moderate, moderately severe and severe depression. It is used to monitor the severity of the depressive symptoms. PHQ scores higher or equal to 10 have been shown to have a sensitivity of 88% and a specificity of 88% for major depression. PHQ-9 is recommended by the American Psychiatric Association (APA) working group for DSM-V as instrument to evaluate the severity of major depressive disorder according to the new DSM-V criteria.

Covariates

The association between DNA methylation level as the outcome and depression as the independent variable was performed using linear regression models, adjusting for age, sex, smoking, antidepressants, white blood cell count, and technical covariates (analytic plate and chip position on plate).

Information on sociodemographic variables including sex, age, smoking status and use of antidepressants was collected by trained medical staff during a standardized interview. The estimated white blood cell proportions were obtained using the method by Houseman et al.[5]

Lothian Birth Cohorts of 1921 (LBC1921) and 1936 (LBC1936)

Description

The LBC1921 and LBC1936 represent relatively healthy older individuals, mostly living in the Lothian region of Scotland, born in 1921 and 1936, respectively. They have been involved in longitudinal assessments of psychological and medical traits from the age of 79 (LBC1921) or 70 (LBC1936) years.[12, 13] Here, we report the first wave, with the largest sample size: 433 (60.3% female) participants with a mean age of 79.1 ± 0.6 years (range = 77 - 80) in the LBC1921; 920 (49.5% female) participants with a mean age of 69.6±0.8 years (range = 67 - 71) in the LBC1936.

DNA methylation sample, measurement, normalization and quality control

514 whole blood samples in LBC1921 and 1 004 samples in LBC1936 were taken and DNA extracted using standard methods at, respectively, MRC Technology, Western General Hospital, Edinburgh and the Wellcome Trust Clinical Research Facility (WTCRF), Western General Hospital, Edinburgh. 485 512 methylation probes were typed at the WTCRF. The R minfi package was used to background correct raw intensity data and to generate methylation beta-values.[2] Probes with a low (< 95%) detection rate (at $p < 0.01$) were removed and manual quality control via inspection of the array control probe signals was used to locate and exclude low quality samples (e.g. samples with inadequate hybridisation, bisulfite conversion, nucleotide extension or staining signal). Samples with a low call rate (samples with < 450,000 probes detected at $p < 0.01$) were excluded based on the Illumina-recommended thresholds. Because the LBC samples had previously been genotyped (Illumina 610-Quadv1), genotypes derived from the 65 SNP control probes on the methylation array using the wateRmelon package[14] were compared to those on the genotyping array. Exclusions were made for samples with a low match of genotypes with SNP control probes, potentially indicative of sample contamination or mix-up (N = 9). Additionally, 8 subjects whose predicted sex, based on XY probes, did not align with reported sex were eliminated.

Depressive symptoms assessment

Depressive symptoms were assessed with the Hospital Anxiety Depression Scale[15], containing seven items each for anxiety and depression. The total score of the seven depression items was used. Self-reported smoking behavior (current, former and never) and medication use was also ascertained in the clinical testing session. The Anatomical Therapeutical Chemical-Code was used to code anti-depressants for current- and non-users.

Covariates

Age, sex, the first four principal components obtained from genotyping array data to index population substructure, B cells, monocytes, CD8T cells, CD4T cells, NK cells, position on BeadChip, methylation sample plate, BeadChip, hybridization date (and smoking for model 2). White blood cell counts were imputed from the methylation array using the Houseman method.[5]

Analysis

A linear regression of the HADS depression score and covariates on each individual methylation probe in the array was performed in R (v3.2.3).

The Rotterdam Study

Description

The Rotterdam Study (RS) is a prospective population based cohort study in a well-defined suburb in the city of Rotterdam, the Netherlands. The design and prospective of the Rotterdam Study has been described in details elsewhere.[16] For the current analysis we used data from individuals aged 45 years and older that participated in the third cohort of the Rotterdam Study. Samples were obtained in two stages. First stage included participants of the first visit of the third cohort (RS-III-1). Second stage was obtained as a part of the BBMRI-NL (Biobanking and Biomolecular Research Infrastructure Netherlands) – BIOS project and included participants of the fifth visit of the first cohort (RS-I-5), third visit of the second cohort (RS-II-3) and second visit of the third cohort (RS-III-2). The subsets from the two stages were analyzed individually and then meta-analyzed, together with other cohorts.

DNA methylation sample, measurement, normalization and quality control

In both subsets of in total 1 479 individuals a whole blood DNA methylation was quantified utilizing same methods. DNA was extracted from whole peripheral blood, stored in EDTA tubes, by standardized salting out methods. The genome-wide DNA-methylation levels were determined using the Illumina HumanMethylation 450K array (Illumina, Inc., San Diego, CA, USA).[17] In short, samples (500ng of DNA per sample) were first treated with bisulfite using the Zymo EZ-96 DNA-

methylation kit (Zymo Research, Irvine, CA, USA). Next, in accordance with the manufacturers' protocol the samples were hybridized to the arrays. The methylation percentage of a CpG site was reported as a β-value ranging between 0 for no methylation and 1 for full methylation.

Quality control (QC) of the samples was carried out using Illumina Genome Studio software (v2011.1, methylation module version 1.9.0; Illumina). During QC samples with incomplete bisulfite treatment, with a low detection rate (< 99%), or gender swaps, were excluded. Also, probes with a detection *p* value > 0.01 in > 1% samples, were filtered out. In total of 474 528 probes passed the QC in the first subset and 419 937 in the BIOS subset (excluding sex chromosome).Filtered β values were than normalized with DASEN implemented in the *wateRmelon* package of the R statistical software.[14]

Depressive symptoms assessment

Depressive symptoms were assessed with the Dutch version of the Centre for Epidemiologic Studies Depression scale (CES-D).[18] The CES-D scale was designed to assess presence and severity of self-reported depressive symptoms.[19] Participants were asked 20 questions that correspond with criterion based-symptoms associated with depression, and they could score from 0 - 60. The screening for depressive symptoms was performed during the home interview by trained research assistants.

Covariates

Age, sex, current smoking status and antidepressant medication were measured using standard cohort specific protocols at the time DNA samples were collected. During this interview participants reported on their smoking behavior (current, former and never) and psychotropic medication use. Moreover, to confirm the self-reported use of antidepressants prescription, the Anatomical Therapeutical Chemical-Code[20] data was collected from pharmacies linked records. Exposure to an antidepressant was defined as current user and non-user of any antidepressant as described in detail elsewhere.[20] White blood cells counts (monocytes, granulocytes and lymphocytes) were measured immediately at the research center using a standard hematology analyzer (Beckman Coulter, Pasadena, CA, USA).

Analysis

An epigenome-wide linear regression was performed per DNA methylation probe with CES-D depression score adjusting for covariates in R. Model 1 was adjusted for age, sex, smoking status, batch effects and cell composition, while Model 2 was additionally adjusted for antidepressant medication use.

Generation Scotland Study

Description

Generation Scotland: Scottish Family Health Study (GS:SFHS) is a population- and family-based cohort comprising ~24 000 individuals from the Scottish population. The cohort has been described in detail previously.[21] Two hundred and twenty-seven individuals (118 cases, 109 controls) were included in the current study. These individuals are of European descent and were aged 40 years or over at the time of recruitment to GS:SFHS. Antidepressant usage was determined by self-report at the time of recruitment. Blood samples for DNA extraction were obtained at the time of recruitment to GS:SFHS.

DNA methylation sample, measurement, normalization and quality control

Whole blood genomic DNA (500 ng) was treated with sodium bisulphite using the EZ-96 DNA Methylation Kit (Zymo Research, Irvine, California), according to the manufacturer's instructions. DNA methylation was assessed using the Infinium HumanMethylation450 BeadChip (Illumina Inc., San Diego, California), according to the manufacturers protocol. Samples were assigned to slides such that gender was counter-balanced. Methylation array processing was carried out at the Wellcome Trust Clinical Research Facility, The University of Edinburgh.

Raw intensity (.idat) files were read into R using the minfi package[2], which was used to perform initial quality control assessments by assessing signal from the array's internal control probes. All samples were deemed to have performed well based on the output of the control probes.

Probe and sample filtering was performed using the pfilter function in wateRmelon.[14] Probes were excluded if they had a detection p value ≥ 0.05 in more than 5% samples ($N = 1000$) or had a beadcount of less than 3 in more than 5% samples ($N = 466$). All samples satisfied an inclusion criterion of having no more than 5% probes with a detection p value ≥ 0.01. The final dataset comprised 484 145 probes measured in 227 individuals.

The raw data (β values) were normalized using the dasen function in wateRmelon.[14]

Assessment of depression

Diagnoses of depression were established by trained researchers who administered the screening questions of the Structured Clinical Interview for DSM-IV (SCID)[22] to all GS:SFHS participants. Those who screened positive were subsequently assessed using the mood disorder sub-sections of the SCID. Major depressive disorder (MDD) was diagnosed in those meeting the criteria for MDD whose symptoms could not be better explained by bipolar disorder, a general medical condition or substance abuse.

Covariates

The following variables were included as covariates: sex, age, current smoking status, family, estimated blood cell counts, batch (methylation slide ID), and methylation-based principal components. Antidepressant usage was additionally included as a covariate in a sensitivity analysis carried out to assess the contribution of antidepressant use to depression-associated methylation differences.

Smoking status was determined by self-report at the time of blood sample collection and was recorded as a binary variable (current smokers vs. non-smokers (never smokers and ex-smokers)). Antidepressant usage was also determined by self-report at the time of blood sample collection. No data on antidepressant usage was available for seven individuals.

Estimated proportions of six blood cell types (monocytes, granulocytes, CD4+ T-cells, CD8+ T-cells, B-cells and natural killer cells) were obtained using the estimateCellCounts function in minfi.[5, 23] In order to avoid collinearity, one cell type, natural killer cells, was excluded from the analyses.

Methylation-based principal components were identified using prcomp without scaling. PCA was performed on the matrix of residualised β-values obtained after residualising for the other covariates included in the model used to identify differentially methylated positions (principal components were estimated separately for Model 1 and Model 2). Principal components that individually accounted for at least 1% of the variance were included as covariates. For both models, this resulted in the inclusion of the first six principal components.

As our sample comprised related individuals (100 discordant sib pairs and 27 singletons), it was necessary to account for relatedness. This was achieved using the duplicateCorrelation function in limma[24, 25] to fit family as a random effect.

Atherosclerosis Risk In Communities (ARIC) Study

DNA methylation sample, measurement, normalization and quality control

DNA methylation analysis was conducted with the Infinium HumanMethylation450 BeadChip (HM450) array (Illumina Inc., San Diego, CA) on genomic DNA extracted from blood samples collected at ARIC Visit 2. Assay were performed on participants who had not restricted use of their DNA and for whom at least 1 µg of DNA and genome-wide genotyping data were available. Details of assay and QC procedures have been previously published.[27] Briefly, genomic DNA was treated with sodium bisulfite using the EZ-96 DNA methylation kit (Zymo Research Corporation, Irvine, CA) following the manufacturer's protocol. Bisulfite converted DNA was

amplified, enzymatically fragmented, purified and hybridized to the HM450 array in accordance with the manufacturer's directions. Methylation typing at 485 577 CpG sites was performed using GenomeStudio 2011.1 (Illumina Inc., San Diego, CA). Methylation level for each probe was derived as a beta value representing the fractional level of methylation at that location. Quality control analysis was performed using the wateRmelon R package.[14] Probe data were excluded if they had a low detection rate (< 95% at $p < 0.01$) and a high missing rate (greater than 1% across all samples). Sample data were excluded based on the following criteria: (1) greater than 5% missing values across all probes; (2) possible gender mismatch based on principal component analysis; (3) genotype mismatch based on 24 SNPs present on the HM450 array. Methylation values were normalized using the Beta MIxture Quantile dilation (BMIQ) method.[28] All ARIC Study participants have provided written informed consent for genomic studies and this work was conducted in compliance with the Helsinki Declaration. The Committee for the Protection of Human Subjects at the University of Texas Health Science Center at Houston has approved this research.

Depressive symptoms assessment

Depressive symptoms were assessed at visit 2 using the 21-item Maastricht Questionnaire.[29] Responses to the questionnaire are coded as yes = 2, don't know = 1, and no = 0. Two items, questions 9 and 14, are reversed coded (yes = 0, don't know = 1, no = 2). Responses are summed to obtain an overall vital exhaustion score, which ranges from 0 to 42, with higher scores representing more exhaustion. Cronbach alpha for internal consistency has been reported as 0.89.[29] Although designed to measure vital exhaustion, the 21-item Maastricht Questionnaire has been shown to correlate with measures of depressive symptoms[30] and was previously used to assess depressive symptoms.[31, 32]

Covariates

Age, sex, current smoking status and antidepressant medication were measured using standard cohort specific protocols at visit 2. Blood cell proportions were imputed using the Houseman method.[5] Genotype principal components were derived from genome-wide genotype data (Affymetrix 6.0).

Analysis

Analysis was performed according to a pre-specified analysis plan adjusting for biological and technical covariates (see attached spreadsheet for details).

Women's Health Initiative - Epigenetic Mechanisms of PM-Mediated CVD

Description

Women's Health Initiative - Epigenetic Mechanisms of PM-Mediated CVD (WHI-EMPC) is an ancillary study of epigenetic mechanisms underlying associations between ambient particulate matter (PM) air pollution and cardiovascular disease (CVD) in the Women's Health Initiative clinical trials (CT) cohort. It is funded by the National Institute of Environmental Health Sciences (R01-ES020836).[33]

The WHI-EMPC study population is a stratified, random sample of 2 200 WHI CT participants who were examined between 1993 and 2001; had available buffy coat, core analytes, electrocardiograms, and ambient concentrations of PM; but were not taking anti-arrhythmic medications at the time.

As such, WHI-EMPC is representative of the larger, multiethnic WHI CT population from which it was sampled: *N* = 68 132 participants aged 50 - 79 years who were randomized to hormone therapy, calcium / vitamin D supplementation, and / or dietary modification in 40 U.S. clinical centers at the baseline exam (1993 - 1998) and re-examined in the fasting state one, three, six, and nine years later.[34]

During participant visits, data on age, race/ethnicity, smoking status, depressive symptoms, and antidepressant medication use were obtained. Current analyses were in European Americans and involved information collected at the first visit with available DNA methylation (DNAm) data.

DNA Methylation

Genome-wide DNAm at CpG sites was measured using the Illumina 450K Infinium Methylation BeadChip, quantitatively represented by beta (the percentage of methylated cytosines over the sum of methylated and unmethylated cytosines), and quality controlled using the following filters: detection p values > 0.01 in > 10% of samples, detection p values > 0.01 in > 1% of probes, yielding values of beta at 484,220 sites. DNAm data was normalized using BMIQ[28], then stage- and plate-adjusted using ComBat.[10] Modeled epigenome-wide associations also adjusted for principal components for ancestral admixture, cell subtype proportions (CD8-T, CD4-T, B cell, natural killer, monocyte, and granulocyte)[5], and technical covariates including array, row, and column.

Depressive symptoms ascertainment

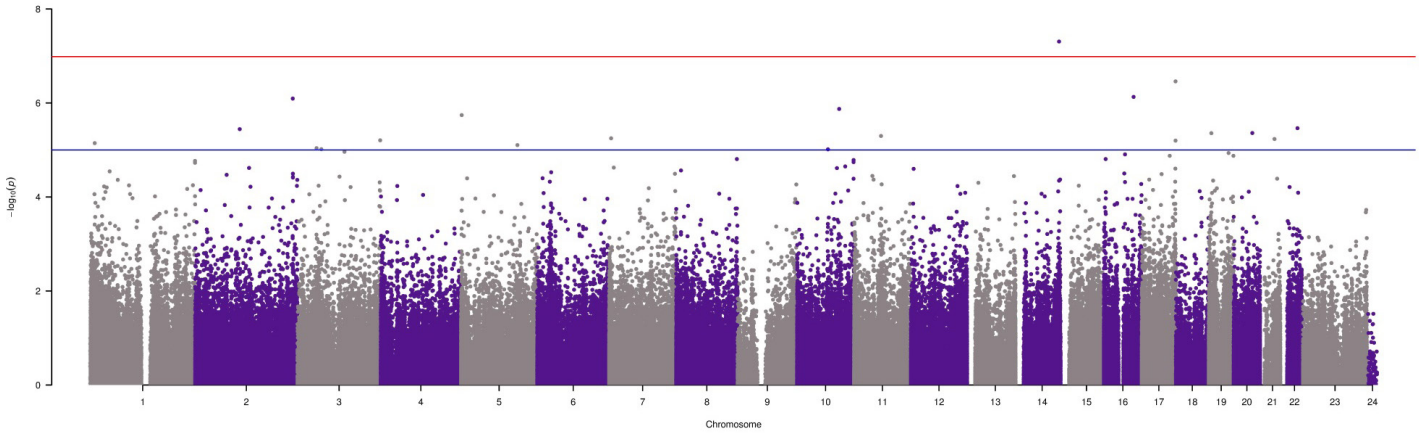
Depressive symptoms were obtained on the day of DNAm collection using a short screening instrument developed by Burnman et al.[35] that combines six items from the CES-D that highly correlate with the full instrument (r = 0.88)[36] with two items from the Diagnostic Interview Schedule (DIS). This CES-D/DIS screening instrument ranges from 0 to 1 with a higher score indicating a greater likelihood of depression.

REFERENCES FROM APPENDIX

1. Fried LP, Borhani NO, Enright P, Furberg CD, Gardin JM, Kronmal RA *et al.* The Cardiovascular Health Study: design and rationale. *Ann Epidemiol* 1991; 1(3): 263-276.
2. Aryee MJ, Jaffe AE, Corrada-Bravo H, Ladd-Acosta C, Feinberg AP, Hansen KD *et al.* Minfi: a flexible and comprehensive Bioconductor package for the analysis of Infinium DNA methylation microarrays. *Bioinformatics* 2014; 30(10): 1363-1369.
3. Maksimovic J, Gordon L, Oshlack A. SWAN: Subset-quantile within array normalization for illumina infinium HumanMethylation450 BeadChips. *Genome Biol* 2012; 13(6): R44.
4. Fortin JP, Labbe A, Lemire M, Zanke BW, Hudson TJ, Fertig EJ *et al.* Functional normalization of 450k methylation array data improves replication in large cancer studies. *Genome Biol* 2014; 15(12): 503.
5. Houseman EA, Accomando WP, Koestler DC, Christensen BC, Marsit CJ, Nelson HH *et al.* DNA methylation arrays as surrogate measures of cell mixture distribution. *BMC Bioinformatics* 2012; 13: 86.
6. Dawber TR, Meadors GF, Moore FE, Jr. Epidemiological approaches to heart disease: the Framingham Study. *Am J Public Health Nations Health* 1951; 41(3): 279-281.
7. Kannel WB, Feinleib M, McNamara PM, Garrison RJ, Castelli WP. An investigation of coronary heart disease in families. The Framingham offspring study. *Am J Epidemiol* 1979; 110(3): 281-290.
8. Chen YA, Lemire M, Choufani S, Butcher DT, Grafodatskaya D, Zanke BW *et al.* Discovery of cross-reactive probes and polymorphic CpGs in the Illumina Infinium HumanMethylation450 microarray. *Epigenetics* 2013; 8(2): 203-209.
9. Amin N, Jovanova O, Adams HH, Dehghan A, Kavousi M, Vernooij MW *et al.* Exome-sequencing in a large population-based study reveals a rare Asn396Ser variant in the LIPG gene associated with depressive symptoms. *Mol Psychiatry* 2016.
10. Johnson WE, Li C, Rabinovic A. Adjusting batch effects in microarray expression data using empirical Bayes methods. *Biostatistics* 2007; 8(1): 118-127.
11. Kroenke K, Spitzer RL, Williams JB. The PHQ-9: validity of a brief depression severity measure. *J Gen Intern Med* 2001; 16(9): 606-613.
12. Luciano M, Houlihan LM, Harris SE, Gow AJ, Hayward C, Starr JM *et al.* Association of existing and new candidate genes for anxiety, depression and personality traits in older people. *Behav Genet* 2010; 40(4): 518-532.
13. Deary IJ, Gow AJ, Pattie A, Starr JM. Cohort profile: the Lothian Birth Cohorts of 1921 and 1936. *Int J Epidemiol* 2012; 41(6): 1576-1584.
14. Pidsley R, CC YW, Volta M, Lunnon K, Mill J, Schalkwyk LC. A data-driven approach to preprocessing Illumina 450K methylation array data. *BMC Genomics* 2013; 14: 293.
15. Zigmond AS, Snaith RP. The hospital anxiety and depression scale. *Acta Psychiatr Scand* 1983; 67(6): 361-370.
16. Hofman A, Brusselle GG, Darwish Murad S, van Duijn CM, Franco OH, Goedegebure A *et al.* The Rotterdam Study: 2016 objectives and design update. *Eur J Epidemiol* 2015; 30(8): 661-708.
17. Sandoval J, Heyn H, Moran S, Serra-Musach J, Pujana MA, Bibikova M *et al.* Validation of a DNA methylation microarray for 450,000 CpG sites in the human genome. *Epigenetics* 2011; 6(6): 692-702.
18. Breslau N. Depressive symptoms, major depression, and generalized anxiety: a comparison of self-reports on CES-D and results from diagnostic interviews. *Psychiatry Res* 1985; 15(3): 219-229.
19. Hales DP, Dishman RK, Motl RW, Addy CL, Pfeiffer KA, Pate RR. Factorial validity and invariance of the center for epidemiologic studies depression (CES-D) scale in a sample of black and white adolescent girls. *Ethn Dis* 2006; 16(1): 1-8.

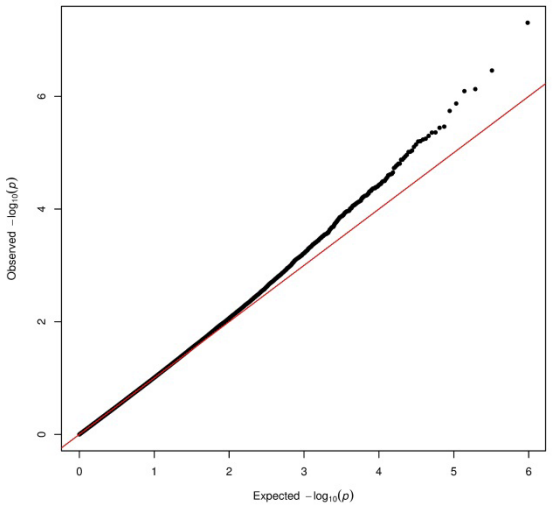
20. Noordam R, Aarts N, Verhamme KM, Sturkenboom MC, Stricker BH, Visser LE. Prescription and indication trends of antidepressant drugs in the Netherlands between 1996 and 2012: a dynamic population-based study. *Eur J Clin Pharmacol* 2015; 71(3): 369-375.
21. Smith BH, Campbell A, Linksted P, Fitzpatrick B, Jackson C, Kerr SM *et al.* Cohort Profile: Generation Scotland: Scottish Family Health Study (GS:SFHS). The study, its participants and their potential for genetic research on health and illness. *Int J Epidemiol* 2013; 42(3): 689-700.
22. First M, Gibbon, M., Spitzer, R., and Willaims, J. Structured clinical interview for DSM-IV-TR Axis Disorder (Research Version). SCID-I/P, 2002.
23. Jaffe AE, Irizarry RA. Accounting for cellular heterogeneity is critical in epigenome-wide association studies. *Genome Biol* 2014; 15(2): R31.
24. Smyth GK, Michaud J, Scott HS. Use of within-array replicate spots for assessing differential expression in microarray experiments. *Bioinformatics* 2005; 21(9): 2067-2075.
25. Ritchie ME, Phipson B, Wu D, Hu Y, Law CW, Shi W *et al.* limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res* 2015; 43(7): e47.
26. The Atherosclerosis Risk in Communities (ARIC) Study: design and objectives. The ARIC investigators. *Am J Epidemiol* 1989; 129(4): 687-702.
27. Demerath EW, Guan W, Grove ML, Aslibekyan S, Mendelson M, Zhou YH *et al.* Epigenome-wide association study (EWAS) of BMI, BMI change and waist circumference in African American adults identifies multiple replicated loci. *Hum Mol Genet* 2015; 24(15): 4464-4479.
28. Teschendorff AE, Marabita F, Lechner M, Bartlett T, Tegner J, Gomez-Cabrero D *et al.* A beta-mixture quantile normalization method for correcting probe design bias in Illumina Infinium 450 k DNA methylation data. *Bioinformatics* 2013; 29(2): 189-196.
29. Appels A, Hoppener P, Mulder P. A questionnaire to assess premonitory symptoms of myocardial infarction. *Int J Cardiol* 1987; 17(1): 15-24.
30. Kopp MS, Falger PR, Appels A, Szedmak S. Depressive symptomatology and vital exhaustion are differentially related to behavioral risk factors for coronary artery disease. *Psychosom Med* 1998; 60(6): 752-758.
31. Wojciechowski FL, Strik JJ, Falger P, Lousberg R, Honig A. The relationship between depressive and vital exhaustion symptomatology post-myocardial infarction. *Acta Psychiatr Scand* 2000; 102(5): 359-365.
32. Wattanakit K, Williams JE, Schreiner PJ, Hirsch AT, Folsom AR. Association of anger proneness, depression and low social support with peripheral arterial disease: the Atherosclerosis Risk in Communities Study. *Vasc Med* 2005; 10(3): 199-206.
33. Hidaka BH. Depression as a disease of modernity: explanations for increasing prevalence. *J Affect Disord* 2012; 140(3): 205-214.
34. Design of the Women's Health Initiative clinical trial and observational study. The Women's Health Initiative Study Group. *Control Clin Trials* 1998; 19(1): 61-109.
35. Burnam MA, Wells KB, Leake B, Landsverk J. Development of a brief screening instrument for detecting depressive disorders. *Med Care* 1988; 26(8): 775-789.
36. Wassertheil-Smoller S, Shumaker S, Ockene J, Talavera GA, Greenland P, Cochrane B *et al.* Depression and cardiovascular sequelae in postmenopausal women. The Women's Health Initiative (WHI). *Arch Intern Med* 2004; 164(3): 289-298.

Supplementary Figure 1: Genome-wide association plot of the discovery EWAS.



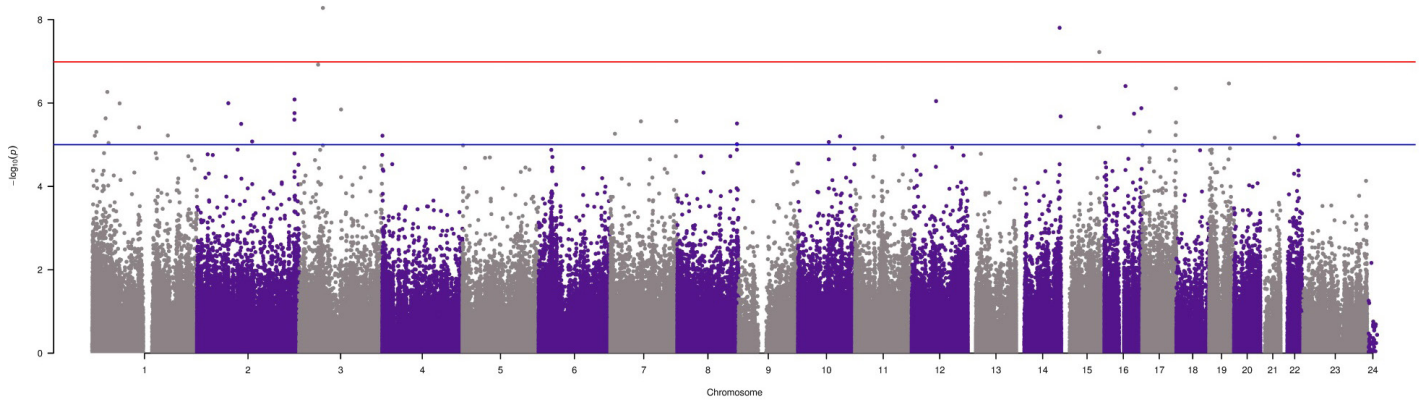
Horizontal axis depicts chromosomes and vertical axis depicts the negative logarithm of the association *p-value*. Each dot represents a CpG site. The solid red line indicates the genome-wide significance threshold and blue solid line indicates the suggestive threshold.

Supplementary Figure 2: Quantile-quantile plots for the discovery EWAS.



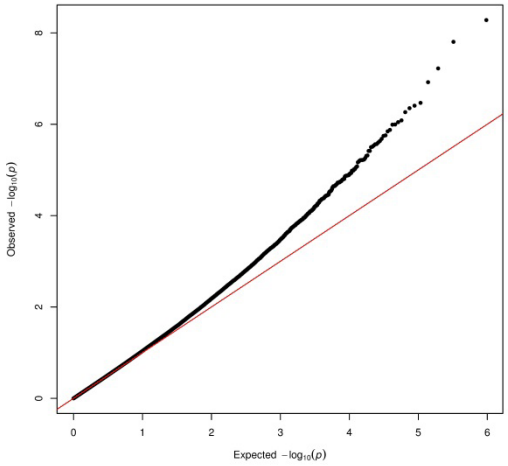
Horizontal axis depicts the expected negative logarithm of the association *p-value*, while the vertical axis shows the observed negative logarithm of the association *p-value* (λ = 1.030).

Supplementary Figure 3: Genome-wide association plot of the meta-analysis of discovery and replication EWAS.



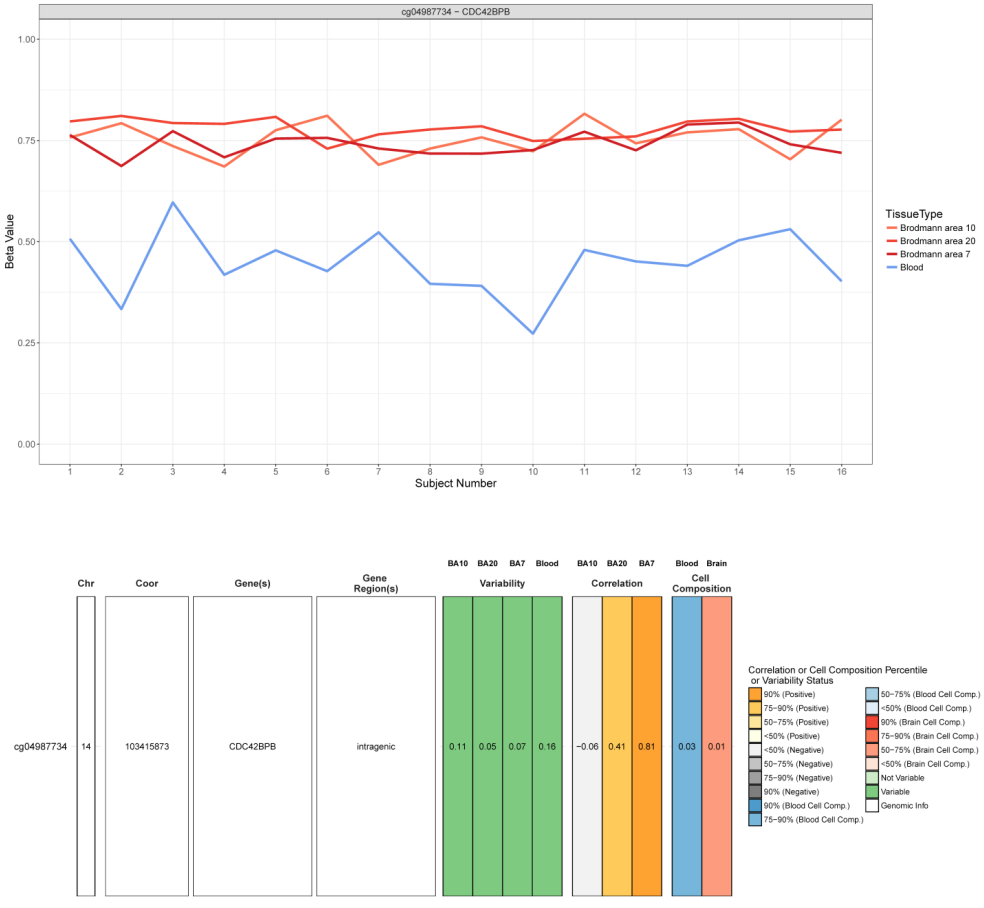
Horizontal axis depicts chromosomes and vertical axis depicts the negative logarithm of the association *p-value*. Each dot represents a CpG site. The solid red line indicates the genome-wide significance threshold and blue solid line indicates the suggestive threshold.

Supplementary Figure 4: Quantile-quantile plot of the meta-analysis of discovery and replication EWAS (λ = 1.088).

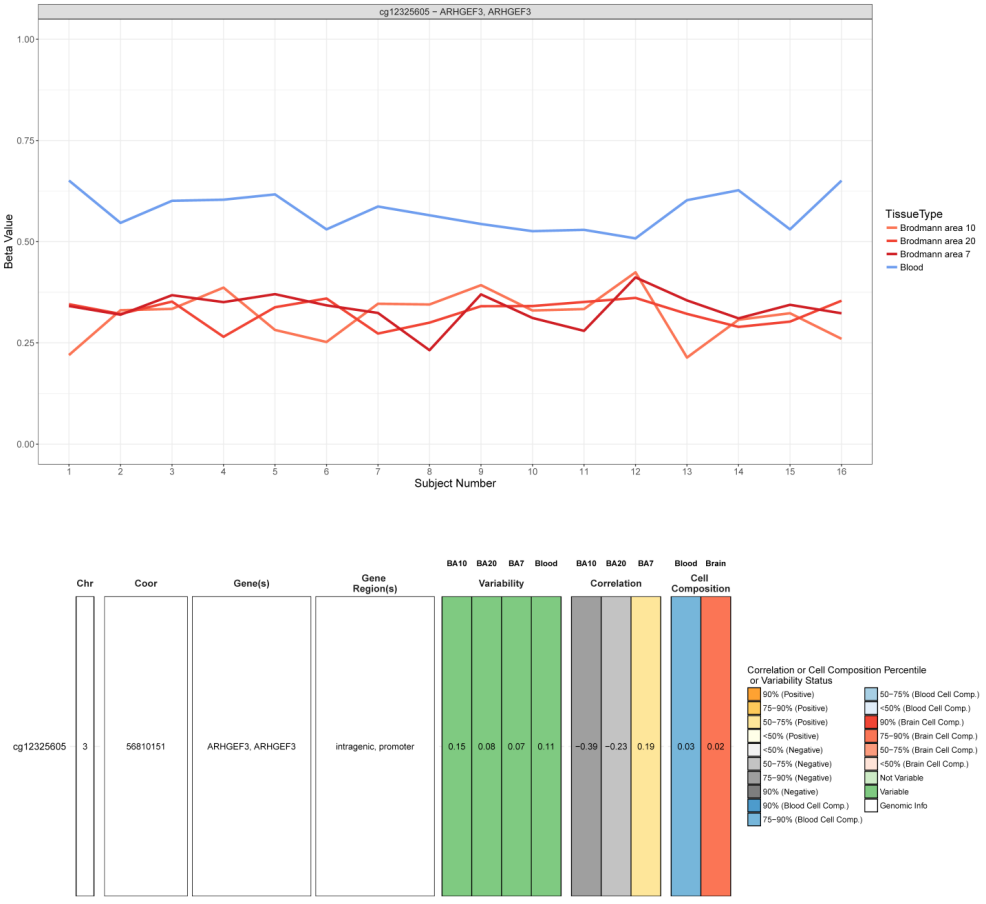


Horizontal axis depicts the expected negative logarithm of the association *p-value*, while the vertical axis shows the observed negative logarithm of the association *p-value*.

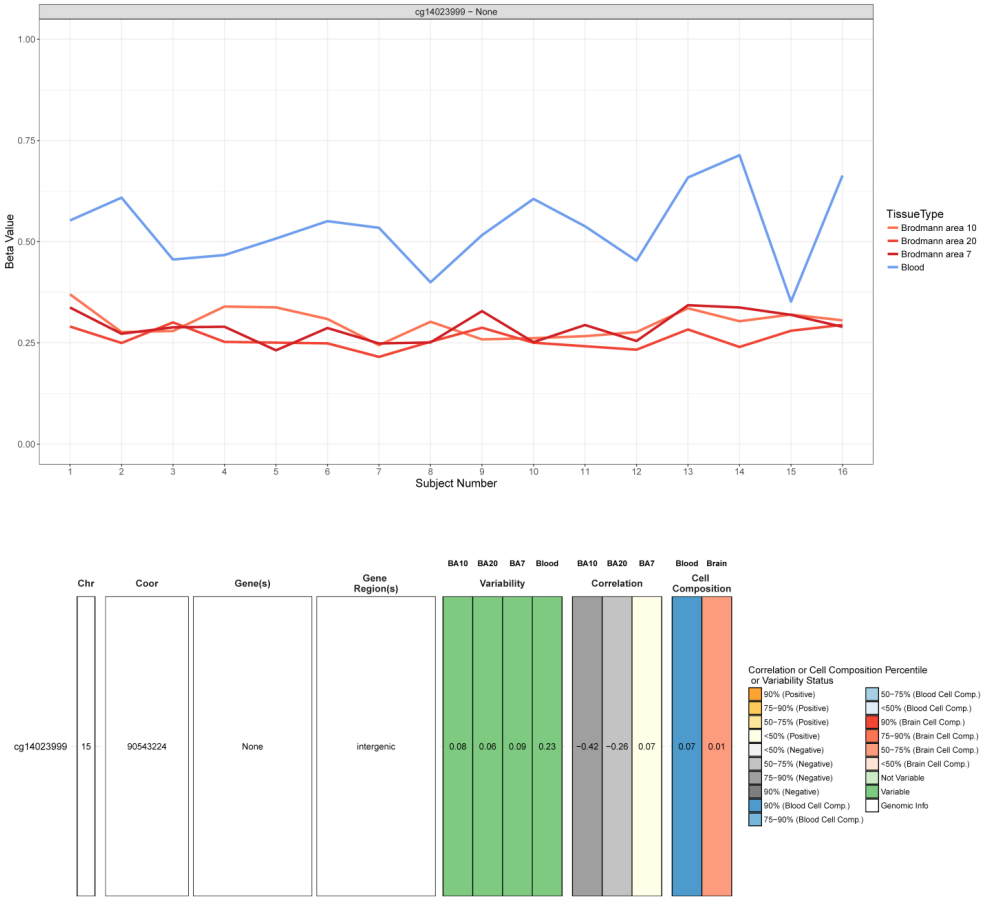
Supplementary Figure 5. Correlation between blood and brain methylation at cg04987734.



Supplementary Figure 6: Correlation between blood and brain methylation at cg12325605.



Supplementary Figure 7: Correlation between blood and brain methylation at cg14023999.



Supplementary Table 1. List of genes used in pathway analysis

HTR2B	ABHD15	ACSS3	ADCY7	AGRP	ALG5	ANKRD20B	APC2	ARHGEF12	ASAH2B
MARCH1	ABHD2	ACTA1	ADCY8	AGTRAP	ALG8	ANKRD22	APCDD1	ARHGEF16	ASAM
MARCH3	ABHD3	ACTA2	ADCY9	AGXT	ALG9	ANKRD23	APCDD1L	ARHGEF17	ASAP1
MARCH4	ABHD8	ACTC1	ADCYAP1R1	AGXT2	ALK	ANKRD24	APCS	ARHGEF18	ASAP2
MARCH6	ABI1	ACTG1	ADD1	AGXT2L2	ALKBH1	ANKRD28	APEH	ARHGEF19	ASB10
MARCH7	ABI3	ACTL6B	ADD2	AHCY	ALKBH4	ANKRD30A	APEX1	ARHGEF2	ASB13
MARCH8	ABL1	ACTL7B	ADD3	AHCYL1	ALKBH5	ANKRD31	APEX2	ARHGEF3	ASB14
SEPT4	ABL2	ACTL9	ADH4	AHCYL2	ALKBH6	ANKRD33B	APH1B	ARHGEF4	ASB18
SEPT5	ABLIM1	ACTN1	ADH6	AHDC1	ALKBH7	ANKRD34C	APITD1	ARHGEF6	ASB2
SEPT6	ABLIM2	ACTN2	ADH7	AHNAK	ALLC	ANKRD45	APLP2	ARHGEF7	ASB9
SEPT8	ABLIM3	ACTN3	ADHFE1	AHNAK2	ALOX12	ANKRD46	APOB48R	ARID1A	ASCC1
SEPT9	ABP1	ACTN4	ADIPOR1	AHRR	ALOX12B	ANKRD52	APOBEC3A	ARID1B	ASCC3
T12	ABR	ACTR1A	ADIPOR2	AIFM1	ALOX12P2	ANKRD53	APOBEC3B	ARID2	ASCL1
A1BG	ABTB2	ACTR1B	ADK	AIFM2	ALOX15	ANKRD54	APOBEC3F	ARID3A	ASCL2
A2BP1	ACAA1	ACTR2	ADM	AIFM3	ALOX15B	ANKRD57	APOBEC3G	ARID3C	ASF1B
A2M	ACAA2	ACTR3C	ADO	AIG1	ALOX5	ANKRD58	APOC1	ARID5B	ASFMR1
A2ML1	ACACB	ACTRT2	ADORA1	AIM1	ALOXE3	ANKRD6	APOD	ARIH1	ASGR1
A4GALT	ACAD10	ACVR1B	ADORA2A	AIM1L	ALPI	ANKRD9	APOL4	ARL1	ASGR2
A4GNT	ACAD11	ACVRL1	ADORA3	AIM2	ALPK3	ANKS1A	APOL6	ARL11	ASH2L
AAA1	ACADS	ACY1	ADPGK	AIMP1	ALPL	ANKS1B	APOLD1	ARL15	ASIP
AACS	ACAN	ACY3	ADPRH	AIP	ALS2CL	ANKS3	APOM	ARL16	ASPG
AADACL4	ACAP1	ADAD1	ADPRHL1	AIRE	ALS2CR11	ANKS6	APOOL	ARL17A	ASPH
AAK1	ACAP2	ADAL	ADRA1A	AJAP1	ALS2CR4	ANKZF1	APP	ARL3	ASPHD1
AANAT	ACAP3	ADAM11	ADRA1B	AK1	ALX1	ANLN	APPBP2	ARL4A	ASPHD2
AARS	ACBD3	ADAM12	ADRA2B	AK2	ALX3	ANO1	APPL2	ARL4C	ASPSCR1
AARS2	ACBD4	ADAM15	ADRA2C	AK3L1	ALX4	ANO2	AQP1	ARL4D	ASS1
AASDH	ACBD7	ADAM17	ADRB2	AK5	AMAC1L3	ANO4	AQP12A	ARL5C	ASTN2
AATF	ACCN1	ADAM19	ADRB3	AKAP1	AMBRA1	ANO7	AQP12B	ARL6	ASXL1
AATK	ACCN2	ADAM21	ADRBK1	AKAP12	AMD1	ANO8	AQP2	ARL6IP4	ASZ1
ABAT	ACCN3	ADAM21P1	ADRBK2	AKAP13	AMDHD1	ANO9	AQP5	ARL6IP6	ATAD1
ABCA1	ACCN4	ADAM32	ADRM1	AKAP2	AMDHD2	ANP32A	AQP6	ARMC1	ATAD3A
ABCA12	ACCN5	ADAM33	ADSSL1	AKAP6	AMH	ANP32B	AQP9	ARMC10	ATAD3C
ABCA13	ACE	ADAM5P	AEBP1	AKAP8	AMHR2	ANP32E	AR	ARMC2	ATAD5
ABCA17P	ACE2	ADAM7	AEBP2	AKD1	AMIGO1	ANPEP	ARAP1	ARMC8	ATE1
ABCA2	ACER3	ADAM8	AFAP1	AKIRIN1	AMIGO3	ANTXR1	ARAP2	ARMCX1	ATF2
ABCA3	ACHE	ADAMDEC1	AFAP1L1	AKNA	AMMECR1	ANTXR2	ARAP3	ARMCX2	ATF3
ABCA4	ACLY	ADAMTS1	AFAP1L2	AKR1A1	AMOTL1	ANXA11	ARF1	ARMCX3	ATF4
ABCA5	ACMSD	ADAMTS10	AFF1	AKR1B15	AMOTL2	ANXA4	ARF5	ARNT2	ATF5
ABCB1	ACN9	ADAMTS13	AFF2	AKR1C2	AMPD2	ANXA5	ARF6	ARNTL	ATF6
ABCB7	ACO1	ADAMTS15	AFF3	AKR1D1	AMPD3	ANXA6	ARFGAP2	ARNTL2	ATF6B
ABCB8	ACOT11	ADAMTS16	AFMID	AKR1E2	AMPH	ANXA7	ARFIP1	ARPC1A	ATF7
ABCB9	ACOT13	ADAMTS17	AFTPH	AKT1	AMT	ANXA9	ARFRP1	ARPC1B	ATG16L1
ABCC1	ACOT2	ADAMTS2	AGA	AKT1S1	AMZ1	AOAH	ARGLU1	ARPC3	ATG16L2
ABCC2	ACOT4	ADAMTS20	AGAP1	AKT2	ANAPC1	AOC3	ARHGAP1	ARPM1	ATG3
ABCC3	ACOT7	ADAMTS3	AGAP11	AKT3	ANAPC13	AOX1	ARHGAP10	ARPP-21	ATG4B
ABCC5	ACOX1	ADAMTS7	AGAP2	ALB	ANAPC5	AP1B1	ARHGAP12	ARR3	ATG4D
ABCC6	ACOX3	ADAMTS8	AGAP3	ALCAM	ANAPC7	AP1G1	ARHGAP15	ARRB1	ATG5
ABCC6P1	ACOXL	ADAMTS9	AGAP5	ALDH16A1	ANGPTL2	AP1G2	ARHGAP18	ARRB2	ATG7
ABCC6P2	ACP5	ADAMTSL1	AGAP6	ALDH18A1	ANK1	AP1S2	ARHGAP19	ARRDC2	ATG9B
ABCC8	ACP6	ADAMTSL3	AGAP8	ALDH1L1	ANK2	AP1S3	ARHGAP20	ARRDC5	ATHL1
ABCC9	ACPL2	ADAMTSL4	AGBL1	ALDH2	ANK3	AP2A1	ARHGAP22	ARSA	ATM
ABCD3	ACR	ADAP1	AGBL2	ALDH3A1	ANKAR	AP2A2	ARHGAP23	ARSB	ATN1
ABCD4	ACRBP	ADAP2	AGBL3	ALDH3B1	ANKDD1A	AP2B1	ARHGAP24	ARSD	ATOH1
ABCE1	ACRC	ADAR	AGBL4	ALDH3B2	ANKFN1	AP2M1	ARHGAP25	ARSG	ATOH8
ABCF1	ACRV1	ADARB1	AGBL5	ALDH4A1	ANKFY1	AP3B2	ARHGAP26	ARSI	ATOX1
ABCF2	ACSBG1	ADARB2	AGFG1	ALDH5A1	ANKH	AP3D1	ARHGAP27	ARSJ	ATP10A
ABCG1	ACSF2	ADAT1	AGFG2	ALDH7A1	ANKHD1	AP3M1	ARHGAP28	ART1	ATP10B
ABCG4	ACSF3	ADC	AGL	ALDOB	ANKLE2	AP3M2	ARHGAP29	ART3	ATP11A
ABCG5	ACSL1	ADCK5	AGPAT1	ALG1	ANKMY1	AP4E1	ARHGAP5	ART4	ATP11B
ABCG8	ACSL3	ADCY1	AGPAT2	ALG10B	ANKRD10	AP4S1	ARHGAP6	ARTN	ATP1A1
ABHD1	ACSL4	ADCY10	AGPAT3	ALG12	ANKRD11	APBA2	ARHGAP8	ARVCF	ATP1A3
ABHD10	ACSL5	ADCY2	AGPAT4	ALG13	ANKRD12	APBB1IP	ARHGAP9	ARX	ATP1A4
ABHD12	ACSL6	ADCY3	AGPAT5	ALG14	ANKRD13A	APBB2	ARHGDIA	AS3MT	ATP1B1
ABHD12B	ACSM1	ADCY4	AGPAT6	ALG1L2	ANKRD13B	APBB3	ARHGEF10	ASAH1	ATP1B2
ABHD14B	ACSM5	ADCY5	AGRN	ALG3	ANKRD13D	APC	ARHGEF10L	ASAH2	ATP2A2

ATP2A3
ATP2B1
ATP2B2
ATP2B3
ATP2B4
ATP2C1
ATP2C2
ATP4B
ATP5A1
ATP5B
ATP5E
ATP5G1
ATP5G2
ATP5H
ATP5J
ATP5J2
ATP5L
ATP5S
ATP5SL
ATP6V0A1
ATP6V0A2
ATP6V0A4
ATP6V0B
ATP6V0C
ATP6V0D1
ATP6V0D2
ATP6V0E2
ATP6V1B2
ATP6V1C1
ATP6V1C2
ATP6V1H
ATP7A
ATP8A2
ATP8B1
ATP8B2
ATP8B3
ATP8B4
ATP9B
ATPAF2
ATPBD4
ATPGD1
ATR
ATRIP
ATXN1
ATXN10
ATXN3
ATXN7
ATXN7L1
ATXN7L2
AURKB
AURKC
AUTS2
AVEN
AVP
AVPI1
AVPR1B
AVPR2
AWAT2
AXIN1
AXIN2
AXL
AZI1
AZI2
AZU1B2M
B3GALNT2
B3GALT4
B3GALTL
B3GAT1
B3GAT2
B3GNT2
B3GNT4
B3GNT7
B3GNT9
B3GNTL1
B4GALNT1
B4GALNT3
B4GALT1
B4GALT2
B4GALT3
B4GALT5
B9D1
B9D2
BAALC
BACE1
BACE2
BACH2
BAG2
BAG4
BAGE4
BAGE5
BAHCC1
BAHD1
BAI1
BAI3
BAIAP2
BAIAP2L1
BAIAP3
BAK1
BAMBI
BANF1
BANK1
BANP
BAP1
BARHL2
BARX2
BASP1
BAT1
BAT2
BAT2L1
BAT3
BAT4
BAT5
BAZ1A
BAZ1B
BAZ2B
BBC3
BBOX1
BBS5
BBS7
BBS9
BBX
BCAN
BCAP29
BCAP31
BCAR1
BCAR3
BCAS1
BCAS3
BCAS4
BCAT1
BCAT2
BCKDHA
BCKDHB
BCKDK
BCL10
BCL11A
BCL11B
BCL2
BCL2L1
BCL2L10
BCL2L14
BCL2L15
BCL2L2
BCL3
BCL6
BCL7A
BCL7B
BCL7C
BCL9
BCL9L
BCLAF1
BCOR
BCORL1
BCR
BCYRN1
BDH1
BDH2
BDKRB1
BDKRB2
BDNF
BDNFOS
BDP1
BEAN
BECN1
BEGAIN
BEND3
BEND4
BEND5
BEND6
BEST4
BET1
BEX4
BEX5
BEYLA
BFAR
BFSP2
BGN
BHLHB9
BHLHE23
BHLHE40
BHLHE41
BHMT
BHMT2
BICC1
BICD1
BICD2
BIK
BIN1
BIN2
BIN3
BIRC5
BIRC7
BIVM
BLCAP
BLK
BLMH
BLOC1S2
BLVRA
BMF
BMP1
BMP2K
BMP4
BMP6
BMP7
BMP8A
BMP8B
BMPER
BMPR2
BNC1
BNC2
BNIP1
BNIP2
BNIP3L
BOC
BOLA1
BOP1
BPGM
BPHL
BPIL1
BRAF
BRD1
BRD2
BRD3
BRD4
BRE
BRF1
BRF2
BRIP1
BRP44
BRP44L
BRPF1
BRPF3
BRS3
BRUNOL4
BRUNOL5
BRUNOL6
BRWD1
BSCL2
BSDC1
BSG
BSN
BSPRY
BST1
BST2
BSX
BTAF1
BTBD1
BTBD11
BTBD12
BTBD17
BTBD18
BTBD19
BTBD2
BTBD7
BTBD8
BTBD9
BTF3
BTG2
BTG3
BTG4
BTN2A1
BTN3A1
BTN3A3
BTNL2
BTNL9
BUB3
BUD13
BUD31
BVES
BZRAP1
BZW1
BZW2
C10orf105
C10orf107
C10orf11
C10orf114
C10orf116
C10orf119
C10orf122
C10orf125
C10orf128
C10orf129
C10orf140
C10orf2
C10orf26
C10orf27
C10orf4
C10orf46
C10orf47
C10orf54
C10orf55
C10orf57
C10orf58
C10orf62
C10orf68
C10orf71
C10orf72
C10orf75
C10orf76
C10orf79
C10orf81
C10orf82
C10orf90
C10orf91
C10orf95
C10orf99
C11orf1
C11orf17
C11orf31
C11orf34
C11orf41
C11orf45
C11orf46
C11orf49
C11orf54
C11orf57
C11orf58
C11orf63
C11orf66
C11orf67
C11orf71
C11orf73
C11orf75
C11orf80
C11orf82
C11orf84
C11orf85
C11orf87
C11orf88
C11orf9
C11orf90
C11orf91
C11orf93
C12orf11
C12orf12
C12orf23
C12orf26
C12orf27
C12orf34
C12orf35
C12orf4
C12orf42
C12orf43
C12orf48
C12orf49
C12orf5
C12orf51
C12orf54
C12orf56
C12orf60
C12orf67
C12orf72
C12orf75
C12orf76
C13orf15
C13orf16
C13orf27
C13orf36
C13orf37
C13orf38
C14orf1
C14orf104
C14orf106
C14orf118
C14orf119
C14orf126
C14orf132
C14orf135
C14orf143
C14orf148
C14orf159
C14orf162
C14orf167
C14orf178
C14orf179
C14orf181
C14orf183
C14orf184
C14orf19
C14orf2
C14orf21
C14orf23
C14orf34
C14orf37
C14orf4
C14orf42
C14orf43
C14orf45
C14orf49
C14orf50
C14orf64
C14orf68
C14orf70
C14orf72
C14orf73
C14orf93
C15orf17
C15orf2
C15orf23
C15orf27
C15orf28
C15orf29
C15orf33
C15orf38
C15orf40
C15orf41
C15orf42
C15orf43
C15orf53
C15orf58
C15orf60
C15orf61
C16orf11
C16orf13
C16orf35
C16orf38
C16orf42
C16orf45
C16orf46
C16orf5
C16orf52
C16orf62
C16orf65
C16orf67
C16orf68
C16orf7
C16orf70
C16orf72
C16orf73
C16orf74
C16orf75
C16orf81
C16orf86
C16orf87
C16orf90
C16orf92
C17orf101
C17orf102
C17orf104
C17orf108
C17orf42
C17orf46
C17orf49
C17orf51
C17orf57
C17orf59
C17orf60
C17orf62
C17orf64
C17orf66
C17orf68
C17orf70
C17orf74
C17orf75
C17orf76
C17orf77
C17orf81
C17orf82
C17orf85
C17orf87
C17orf88
C17orf89
C17orf90
C17orf91
C17orf93
C17orf96
C17orf97
C17orf98
C18orf1
C18orf10
C18orf18
C18orf2
C18orf20
C18orf26
C18orf32
C19orf10
C19orf12
C19orf18
C19orf21
C19orf23
C19orf24
C19orf25
C19orf28
C19orf29
C19orf34
C19orf36
C19orf40
C19orf41
C19orf43
C19orf45
C19orf46
C19orf48
C19orf50
C19orf51
C19orf52
C19orf53
C19orf6
C19orf60
C19orf63
C19orf71
C19orf77
C1GALT1C1
C1orf100
C1orf103
C1orf106
C1orf107
C1orf113
C1orf115
C1orf122
C1orf123
C1orf124
C1orf130
C1orf14
C1orf141
C1orf144
C1orf146
C1orf151
C1orf159
C1orf163
C1orf170
C1orf172
C1orf175
C1orf177
C1orf187
C1orf189
C1orf190
C1orf192
C1orf198
C1orf21
C1orf213
C1orf216
C1orf220
C1orf228
C1orf229
C1orf25
C1orf26
C1orf31
C1orf43
C1orf49
C1orf51
C1orf52
C1orf53
C1orf56
C1orf57
C1orf58
C1orf59
C1orf61
C1orf63
C1orf64
C1orf66
C1orf69
C1orf70
C1orf83
C1orf84
C1orf86
C1orf88
C1orf89
C1orf9
C1orf91
C1orf92
C1orf94
C1orf96
C1orf97
C1QBP
C1QL2
C1QL4
C1QTNF1
C1QTNF2
C1QTNF3
C1QTNF4
C1QTNF7
C1RL
C20orf103
C20orf108
C20orf111
C20orf112
C20orf117
C20orf12
C20orf123
C20orf135
C20orf160
C20orf165
C20orf166
C20orf177
C20orf195
C20orf196
C20orf199
C20orf24
C20orf26
C20orf3
C20orf30
C20orf43
C20orf46
C20orf54
C20orf56
C20orf70
C20orf72
C20orf85
C20orf95
C20orf96
C21orf119
C21orf122
C21orf129
C21orf130
C21orf29
C21orf33
C21orf34
C21orf59
C21orf62
C21orf63
C21orf66
C21orf67
C21orf7
C21orf70
C21orf71
C21orf81
C21orf91
C22orf15
C22orf24
C22orf25
C22orf26
C22orf32
C22orf34
C22orf42
C22orf43
C22orf45
C22orf46
C22orf9
C2CD2
C2CD2L
C2CD4A
C2CD4B
C2CD4C
C2orf18
C2orf28
C2orf29
C2orf34
C2orf39
C2orf42
C2orf48
C2orf50
C2orf54
C2orf55
C2orf57
C2orf61
C2orf63
C2orf65
C2orf68
C2orf7
C2orf70
C2orf71
C2orf73
C2orf74
C2orf76
C2orf78
C2orf81
C2orf86
C2orf89
C2orf90
C3orf15
C3orf17
C3orf20
C3orf21
C3orf23
C3orf25
C3orf26
C3orf30
C3orf31
C3orf38
C3orf39
C3orf48
C3orf50
C3orf52
C3orf54
C3orf57
C3orf59
C3orf74
C3orf75
C3orf77
C3P1
C4BPA
C4orf10
C4orf22
C4orf23
C4orf26
C4orf27
C4orf31
C4orf32
C4orf33
C4orf34
C4orf37
C4orf42
C4orf46
C4orf49
C4orf50
C5
C5orf13
C5orf15
C5orf20
C5orf27
C5orf30
C5orf32
C5orf33
C5orf35
C5orf36
C5orf37
C5orf38
C5orf39
C5orf4
C5orf42
C5orf45
C5orf49
C5orf53
C5orf56
C5orf58
C6
C6orf1
C6orf10
C6orf103
C6orf106
C6orf108
C6orf118
C6orf120
C6orf123
C6orf126
C6orf129
C6orf130
C6orf132
C6orf134
C6orf136
C6orf141
C6orf145
C6orf153
C6orf168
C6orf174
C6orf176
C6orf186
C6orf191

R1
R2
R3
R4
R5
R6
R7
R8
R9
R10
R11
R12
R13
R14
R15
R16
R17
R18
R19
R20
R21
R22
R23
R24
R25
R26
R27
R28
R29
R30
R31
R32
R33
R34
R35
R36
R37
R38
R39

C6orf195
C6orf204
C6orf208
C6orf221
C6orf222
C6orf226
C6orf227
C6orf27
C6orf35
C6orf48
C6orf52
C6orf57
C6orf64
C6orf70
C6orf72
C6orf94
C6orf97
C7orf10
C7orf13
C7orf16
C7orf20
C7orf23
C7orf25
C7orf27
C7orf29
C7orf30
C7orf41
C7orf42
C7orf43
C7orf47
C7orf49
C7orf50
C7orf51
C7orf55
C7orf57
C7orf58
C7orf60
C7orf63
C7orf68
C7orf70
C8A
C8orf12
C8orf34
C8orf42
C8orf51
C8orf59
C8orf71
C8orf73
C8orf75
C8orf84
C8orf85
C9orf100
C9orf119
C9orf128
C9orf152
C9orf153
C9orf163
C9orf167
C9orf171
C9orf24
C9orf25
C9orf3
C9orf47
C9orf50
C9orf57
C9orf64
C9orf68
C9orf69
C9orf75
C9orf86
C9orf89
C9orf93
C9orf96
CA1
CA10
CA12
CA13
CA3
CA4
CA5A
CA5B
CA6
CA7
CAB39
CAB39L
CABC1
CABIN1
CABLES1
CABLES2
CABP1
CABP2
CABP4
CABP5
CABP7
CACHD1
CACNA1A
CACNA1B
CACNA1C
CACNA1D
CACNA1E
CACNA1G
CACNA1H
CACNA1I
CACNA2D1
CACNA2D2
CACNA2D3
CACNA2D4
CACNB1
CACNB2
CACNB3
CACNB4
CACNG2
CACNG3
CACNG4
CACNG8
CADM1
CADM2
CADM3
CADPS
CADPS2
CALCA
CALCB
CALCOCO2
CALCR
CALD1
CALHM1
CALHM2
CALHM3
CALM1
CALM3
CALML4
CALML5
CALML6
CALN1
CALR
CALU
CALY
CAMK1D
CAMK2D
CAMK2G
CAMK4
CAMKK1
CAMKK2
CAMKV
CAMSAP1
CAMSAP1L1
CAMTA1
CAND1
CAND2
CANT1
CAP2
CAPG
CAPN1
CAPN10
CAPN12
CAPN13
CAPN14
CAPN2
CAPN3
CAPN5
CAPN8
CAPN9
CAPNS1
CAPRIN1
CAPRIN2
CAPS
CAPZA1
CAPZB
CARD11
CARD14
CARD8
CARKD
CARM1
CARS
CARS2
CARTPT
CASC1
CASC2
CASC3
CASC4
CASK
CASKIN1
CASKIN2
CASP10
CASP4
CASP5
CASP8AP2
CASQ2
CASR
CASS4
CAST
CASZ1
CATSPER2
CATSPER3
CATSPERB
CATSPERG
CAV1
CAV2
CBFA2T3
CBLC
CBLL1
CBLN3
CBR1
CBR4
CBS
CBWD1
CBWD2
CBX3
CBX4
CBX6
CBX7
CBX8
CBY1
CC2D2B
CCDC101
CCDC102A
CCDC102B
CCDC105
CCDC106
CCDC108
CCDC109A
CCDC109B
CCDC11
CCDC110
CCDC111
CCDC112
CCDC113
CCDC114
CCDC116
CCDC124
CCDC127
CCDC130
CCDC134
CCDC136
CCDC137
CCDC138
CCDC140
CCDC141
CCDC142
CCDC146
CCDC148
CCDC15
CCDC152
CCDC154
CCDC159
CCDC160
CCDC163P
CCDC18
CCDC19
CCDC22
CCDC23
CCDC24
CCDC27
CCDC28A
CCDC33
CCDC34
CCDC37
CCDC40
CCDC42
CCDC43
CCDC46
CCDC52
CCDC55
CCDC57
CCDC59
CCDC6
CCDC62
CCDC63
CCDC64
CCDC66
CCDC68
CCDC69
CCDC7
CCDC71
CCDC74B
CCDC75
CCDC81
CCDC85A
CCDC85C
CCDC87
CCDC88B
CCDC88C
CCDC92
CCDC93
CCDC94
CCDC96
CCDC97
CCHCR1
CCK
CCKBR
CCL1
CCL13
CCL17
CCL2
CCL22
CCL25
CCL26
CCL27
CCL28
CCM2
CCNA2
CCNC
CCND1
CCND2
CCND3
CCNE1
CCNF
CCNG1
CCNH
CCNI
CCNJ
CCNJL
CCNL1
CCNL2
CCNO
CCNY
CCR1
CCR10
CCR5
CCR6
CCR7
CCR8
CCR9
CCRL2
CCRN4L
CCS
CCT2
CD109
CD151
CD160
CD163
CD164L2
CD180
CD19
CD1B
CD1C
CD1D
CD1E
CD200
CD244
CD247
CD248
CD274
CD300C
CD300LG
CD302
CD36
CD38
CD3E
CD4
CD40
CD44
CD46
CD47
CD48
CD55
CD58
CD7
CD79B
CD80
CD81
CD82
CD86
CD8A
CD9
CD97
CD99L2
CDAN1
CDC14A
CDC2
CDC20B
CDC25C
CDC27
CDC42
CDC42BPB
CDC42BPG
CDC42EP2
CDC42EP3
CDC42EP5
CDC6
CDC7
CDC73
CDCA2
CDCA4
CDCA5
CDCA7
CDCA7L
CDCP2
CDGAP
CDH1
CDH11
CDH12
CDH13
CDH15
CDH17
CDH18
CDH22
CDH23
CDH3
CDH4
CDH5
CDH7
CDH9
CDK10
CDK11B
CDK12
CDK13
CDK14
CDK15
CDK16
CDK17
CDK19
CDK20
CDK2AP1
CDK5
CDK5R1
CDK5R2
CDK5RAP1
CDK6
CDK7
CDKL2
CDKN1B
CDKN1C
CDKN2AIP
CDKN2BAS
CDKN2C
CDKN3
CDO1
CDON
CDR2
CDR2L
CDRT15
CDRT15P
CDS2
CDSN
CDT1
CDX1
CDX2
CDYL
CDYL2
CEACAM21
CEACAM4
CEACAM6
CEBPA
CEBPE
CEBPG
CEBPZ
CECR1
CECR2
CECR4
CECR5
CELSR3
CEMP1
CENPA
CENPBD1
CENPF
CENPH
CENPK
CENPM
CENPP
CENPT
CENPV
CEP120
CEP164
CEP350
CEP55
CEP57
CEP68
CEP72
CEPT1
CERCAM
CERKL
CES7
CES8
CETN1
CETP
CFB
CFDP1
CFI
CFL2
CFLAR
CFP
CGB
CGB2
CGB8
CGGBP1
CGRRF1
CHADL
CHAF1A
CHCHD3
CHCHD5
CHCHD6
CHD1L
CHD2
CHD3
CHD5
CHD7
CHD8
CHDH
CHEK2
CHFR
CHGB
CHI3L1
CHI3L2
CHIA
CHIC2
CHID1
CHIT1
CHKB-CPT1B
CHL1
CHMP1A
CHMP2A
CHMP2B
CHMP4B
CHMP4C
CHMP6
CHODL
CHP2
CHPF2
CHRDL1
CHRM2
CHRM3
CHRM4
CHRM5
CHRNA10
CHRNA3
CHRNA5
CHRNA6
CHRNA7
CHRNB1
CHRNE
CHRNG
CHST1
CHST10
CHST11
CHST12
CHST13
CHST15
CHST2
CHST3
CHST4
CHST6
CHST7
CHST8
CHST9
CHSY1
CHSY3
CHUK
CIAO1
CIAPIN1
CIB4
CIC
CIDEA
CIDECP
CIITA
CILP
CILP2
CIR1
CISD3
CISH
CIT
CITED1
CITED2
CITED4
CIZ1
CKAP4
CKAP5
CKLF
CLASP1
CLASP2
CLC
CLCN1
CLCN4
CLCN5
CLCN6
CLCN7
CLCNKB
CLDN10
CLDN11
CLDN14
CLDN15
CLDN16
CLDN18
CLDN22
CLDN23
CLDN6
CLDN7
CLDND1
CLEC16A
CLEC17A
CLEC1A
CLEC2A
CLEC2B
CLEC2D
CLEC2L
CLEC3B
CLEC4F
CLEC4G
CLEC4GP1
CLEC9A
CLGN
CLIC1
CLIC2
CLIC3
CLIC5
CLIP1
CLIP2
CLIP3
CLK1
CLK2
CLK4
CLMN
CLN3
CLN5
CLN6
CLN8
CLNK
CLNS1A
CLP1
CLPB
CLPP
CLPS
CLPTM1
CLPTM1L
CLRN1
CLRN2
CLRN3
CLSPN
CLSTN1
CLTB
CLTC
CLU
CLUAP1
CLYBL
CMBL
CMC1
CMIP
CMKLR1
CMPK1
CMTM1
CMTM2
CMTM5
CNBD1
CNDP2
CNFN
CNGA3
CNGB1
CNGB3
CNIH
CNIH2
CNKSR2
CNKSR3
CNN1
CNN3
CNNM1
CNNM2
CNNM3
CNNM4
CNOT1
CNOT10
CNOT3
CNOT6
CNP
CNPY1
CNPY3
CNR1
CNR2
CNRIP1
CNST
CNTD2

CNTF
CNTFR
CNTN1
CNTN2
CNTN4
CNTN6
CNTNAP2
CNTNAP4
CNTNAP5
COBL
COBRA1
COCH
COG2
COG3
COG5
COG6
COG7
COG8
COIL
COL11A1
COL11A2
COL12A1
COL13A1
COL14A1
COL15A1
COL16A1
COL17A1
COL18A1
COL19A1
COL1A2
COL20A1
COL22A1
COL23A1
COL24A1
COL25A1
COL27A1
COL29A1
COL2A1
COL3A1
COL4A1
COL4A2
COL4A3BP
COL4A5
COL5A1
COL5A3
COL6A1
COL6A2
COL6A3
COL6A4P2
COL6A6
COL8A2
COL9A1
COL9A3
COLEC11
COLQ
COMMD10
COMMD5
COMMD7
COMMD8
COMP
COMT
COPB2
COPG
COPS3
COPS6
COPS7A
COPS8
COPZ1
COPZ2
COQ10A
COQ2
COQ5
COQ6
CORIN
CORO1B
CORO1C
CORO2A
CORO2B
CORO6
CORO7
CORT
COTL1
COX10
COX18
COX19
COX4NB
COX5A
COX7A2
COX7B2
COX8C
CP
CPA2
CPA4
CPA5
CPAMD8
CPEB1
CPEB2
CPLX1
CPLX2
CPLX3
CPN1
CPN2
CPNE2
CPNE4
CPNE5
CPNE6
CPNE7
CPNE9
CPO
CPPED1
CPSF1
CPSF3L
CPSF4
CPSF4L
CPSF6
CPT1A
CPT1C
CPT2
CPVL
CPXM2
CR1L
CR2
CRABP1
CRABP2
CRADD
CRAMP1L
CRB1
CRB2
CRB3
CRBN
CRCP
CREB1
CREB3
CREB3L1
CREB3L2
CREB3L3
CREB5
CREBBP
CREG1
CRELD1
CRELD2
CRH
CRHR1
CRHR2
CRIM1
CRIP1
CRIP3
CRK
CRLF1
CRMP1
CRNKL1
CROCC
CROCCL1
CROCCL2
CRTAM
CRTC1
CRY1
CRY2
CRYAB
CRYBA2
CRYBB2
CRYBB3
CRYGB
CRYGD
CRYGN
CRYL1
CRYM
CS
CSDAP1
CSDE1
CSE1L
CSF1
CSF2
CSF3
CSF3R
CSGALNACT1
CSH2
CSHL1
CSK
CSMD1
CSMD2
CSMD3
CSNK1A1
CSNK1D
CSNK1E
CSNK1G1
CSNK1G2
CSNK2A1
CSNK2A2
CSNK2B
CSPG4
CSPP1
CSRNP1
CSRP1
CSRP3
CST1
CST11
CST2
CST5
CST7
CST8
CST9L
CSTF2
CSTF3
CT45A6
CT62
CTAGE1
CTBP1
CTBP2
CTDP1
CTDSP1
CTDSPL
CTDSPL2
CTF1
CTGF
CTNNA2
CTNNAL1
CTNNBIP1
CTNND2
CTNS
CTPS2
CTRB2
CTRC
CTRL
CTSB
CTSD
CTSE
CTSG
CTSO
CTSS
CTSZ
CTTN
CTTNBP2
CTTNBP2NL
CTU1
CTXN2
CUBN
CUEDC1
CUEDC2
CUGBP2
CUL1
CUL3
CUL4A
CUL4B
CUL5
CUTA
CUX1
CUX2
CUZD1
CWC22
CWF19L2
CWH43
CX3CR1
CXADR
CXADRP3
CXCL1
CXCL12
CXCL13
CXCL14
CXCL16
CXCL17
CXCL2
CXCL5
CXCL6
CXCL9
CXCR1
CXCR3
CXCR4
CXCR6
CXCR7
CXorf1
CXorf36
CXorf40B
CXorf41
CXorf48
CXorf56
CXorf57
CXorf58
CXorf64
CXXC4
CXXC5
CYB561
CYB561D1
CYB561D2
CYB5B
CYB5R1
CYB5R3
CYBA
CYBB
CYBRD1
CYC1
CYCS
CYFIP1
CYFIP2
CYGB
CYHR1
CYP11A1
CYP11B1
CYP11B2
CYP17A1
CYP19A1
CYP1A1
CYP1A2
CYP1B1
CYP20A1
CYP24A1
CYP26A1
CYP26C1
CYP27A1
CYP2A13
CYP2A7
CYP2D7P1
CYP2F1
CYP2J2
CYP2S1
CYP2U1
CYP2W1
CYP39A1
CYP3A5
CYP46A1
CYP4A22
CYP4F11
CYP4F12
CYP4V2
CYP4X1
CYP7B1
CYS1
CYSLTR2
CYTH1
CYTH2
CYTH3
CYTIP
CYTSA
CYTSB
D2HGDH
DAAM2
DAB1
DAB2
DAB2IP
DACH1
DAG1
DAGLA
DAGLB
DAK
DALRD3
DAND5
DAOA
DAP
DAPK1
DAPK2
DAPK3
DAPP1
DAXX
DAZAP1
DAZL
DBC1
DBF4B
DBI
DBNDD1
DBNDD2
DBNL
DBT
DBX1
DBX2
DCAF12
DCAF12L2
DCAF13
DCAKD
DCBLD1
DCC
DCHS1
DCHS2
DCLRE1B
DCLRE1C
DCP1A
DCP1B
DCPS
DCST1
DCTD
DCTN1
DCTN2
DCTN3
DCTN6
DCTPP1
DCUN1D1
DCUN1D2
DCUN1D3
DCUN1D5
DCXR
DDAH1
DDAH2
DDB2
DDC
DDHD1
DDI2
DDIT4
DDIT4L
DDR1
DDRGK1
DDX11
DDX17
DDX18
DDX20
DDX23
DDX25
DDX26B
DDX31
DDX39
DDX3X
DDX4
DDX46
DDX47
DDX5
DDX50
DDX51
DDX55
DDX60
DEAF1
DECR1
DECR2
DEDD
DEF6
DEF8
DEFA10P
DEFA4
DEFB1
DEFB107A
DEFB110
DEFB119
DEFB122
DEFB124
DEFB125
DEFB126
DEFB133
DEFB136
DEGS1
DEGS2
DEM1
DENND1A
DENND1B
DENND2A
DENND2D
DENND3
DENND4B
DENND5A
DENND5B
DENR
DEPDC1
DEPDC1B
DEPDC6
DERA
DERL1
DES
DFFB
DFNA5
DFNB31
DGAT1
DGAT2
DGCR14
DGCR2
DGCR5
DGCR9
DGKA
DGKB
DGKD
DGKG
DGKH
DGKI
DGKQ
DGKZ
DHCR7
DHDDS
DHDH
DHRS12
DHRS2
DHRS3
DHRS4
DHRS4L1
DHRS7
DHRS7B
DHTKD1
DHX15
DHX16
DHX29
DHX30
DHX32
DHX34
DHX37
DHX38
DHX40P
DHX57
DHX58
DHX9
DIABLO
DIAPH1
DIAPH2
DIAPH3
DIDO1
DIMT1L
DIO3
DIP2C
DIRAS1
DIRAS3
DIRC1
DIRC2
DIRC3
DIS3L
DISP1
DISP2
DIXDC1
DKC1
DKFZP434H168
DKFZP686I15217
DKFZp686O24166
DKFZp761E198
DKK4
DLC1
DLEC1
DLEU7
DLG2
DLG3
DLG4
DLG5
DLGAP1
DLGAP2
DLGAP3
DLGAP4
DLGAP5
DLK1
DLL1
DLL3
DLL4
DLST
DLX1
DLX3
DLX4
DLX5
DLX6AS
DMAP1
DMBT1
DMBX1
DMD
DMKN
DMP1
DMPK
DMRT1
DMRT3
DMRTA2
DMXL1
DNAH10
DNAH11
DNAH14
DNAH17
DNAH2
DNAH5
DNAH7
DNAH9
DNAI2
DNAJA2
DNAJA3
DNAJA4
DNAJB12
DNAJB13
DNAJB2
DNAJB6
DNAJC1
DNAJC10
DNAJC11
DNAJC13
DNAJC14
DNAJC15
DNAJC16
DNAJC17
DNAJC18
DNAJC19
DNAJC2
DNAJC24
DNAJC3
DNAJC30
DNAJC4
DNAJC5
DNAJC5B
DNAJC6
DNAJC7
DNAJC8
DNAJC9
DNALI1
DNASE1L2
DNASE1L3
DNASE2
DND1
DNER
DNHD1
DNM1
DNM1P35
DNM2
DNM3
DNMBP
DNMT1
DNMT3A
DNMT3B
DNPEP
DNTT
DNTTIP2
DOC2A
DOC2B
DOCK1
DOCK10
DOCK2
DOCK3
DOCK4
DOCK5
DOCK9
DOK1
DOK2
DOK3
DOK5
DOK6
DOK7
DOLPP1
DOM3Z
DONSON
DOPEY1
DOPEY2
DOT1L
DPCR1
DPEP1
DPEP2
DPF3
DPH2
DPP10
DPP4
DPP6
DPP8
DPP9
DPPA3
DPPA4
DPT
DPY19L1
DPY19L2P4
DPYD
DPYSL2
DPYSL4
DPYSL5
DQX1
DR1
DRAM1
DRAM2
DRD1
DRD2
DRD4
DRG1
DRG2
DRP2
DSC3
DSCAM
DSCAML1
DSCC1
DSCR8
DSE
DSG4
DSN1
DSP
DSPP
DST
DSTN
DSTYK
DTHD1
DTNA
DTNB

R1
R2
R3
R4
R5
R6
R7
R8
R9
R10
R11
R12
R13
R14
R15
R16
R17
R18
R19
R20
R21
R22
R23
R24
R25
R26
R27
R28
R29
R30
R31
R32
R33
R34
R35
R36
R37
R38
R39

DTNBP1
DTX1
DTX4
DTYMK
DULLARD
DUOX1
DUS2L
DUS3L
DUSP1
DUSP10
DUSP11
DUSP16
DUSP19
DUSP2
DUSP22
DUSP26
DUSP27
DUSP28
DUSP3
DUSP4
DUSP5
DUSP5P
DUSP6
DUXA
DVL1
DYDC2
DYNC1H1
DYNC1I1
DYNC1LI2
DYNC2H1
DYNLL1
DYNLL2
DYNLRB2
DYNLT3
DYRK1B
DYRK2
DYRK3
DYSF
DZIP1L
E2F2
E2F7
E2F8
E4F1
EARS2
EBAG9
EBF1
EBF3
EBF4
EBP
ECD
ECE1
ECE2
ECHDC1
ECHDC2
ECHDC3
ECSCR
ECSIT
EDA
EDAR
EDC3
EDC4
EDEM2
EDEM3
EDF1
EDIL3
EDN1
EDN2
EDN3
EDNRB
EED
EEF1A2
EEF1D
EEF1DP3
EEF2
EEF2K
EEFSEC
EFCAB1
EFCAB2
EFCAB3
EFCAB4A
EFCAB4B
EFCAB5
EFCAB6
EFCAB8
EFEMP1
EFEMP2
EFHC1
EFHC2
EFHD1
EFHD2
EFNA2
EFNA3
EFNA4
EFNA5
EFNB1
EFNB3
EFR3A
EFR3B
EFS
EGF
EGFL6
EGFL7
EGFL8
EGFLAM
EGFR
EGLN1
EHBP1
EHBP1L1
EHD1
EHD2
EHD3
EHD4
EHF
EHMT1
EHMT2
EI24
EID3
EIF2AK2
EIF2AK3
EIF2AK4
EIF2B2
EIF2B4
EIF2B5
EIF2C1
EIF2C2
EIF2S3
EIF3A
EIF3B
EIF3D
EIF3E
EIF3G
EIF3I
EIF3K
EIF3L
EIF3M
EIF4A1
EIF4A3
EIF4E3
EIF4EBP1
EIF4EBP3
EIF4ENIF1
EIF4G1
EIF4G2
EIF4H
EIF5
EIF5A
EIF5A2
EIF5AL1
EIF5B
ELAC1
ELAC2
ELANE
ELAVL1
ELAVL3
ELAVL4
ELF1
ELF2
ELF4
ELF5
ELFN1
ELFN2
ELK3
ELK4
ELL
ELL2
ELMO1
ELMO2
ELMO3
ELMOD1
ELMOD2
ELMOD3
ELOF1
ELOVL1
ELOVL3
ELOVL4
ELOVL5
ELOVL6
ELP2P
ELP4
ELSPBP1
EMB
EME1
EMID1
EMID2
EMILIN2
EML2
EML6
EMP1
EMP2
EMP3
EMR2
EMX1
EMX2
EMX2OS
EN1
EN2
ENAH
ENC1
ENG
ENGASE
ENHO
ENOSF1
ENOX1
ENOX2
ENPP1
ENPP3
ENPP4
ENPP5
ENPP6
ENPP7
ENTPD1
ENTPD2
ENTPD3
ENTPD6
ENTPD7
ENTPD8
EOMES
EP300
EP400
EP400NL
EPAS1
EPB41
EPB41L1
EPB41L2
EPB41L3
EPB41L4A
EPB42
EPB49
EPC2
EPCAM
EPDR1
EPHA1
EPHA10
EPHA4
EPHA5
EPHA8
EPHB2
EPHB3
EPHB6
EPHX1
EPHX3
EPM2A
EPN1
EPN2
EPOR
EPPK1
EPRS
EPS15L1
EPS8L1
EPS8L2
EPSTI1
EPX
ERAL1
ERAS
ERBB2
ERBB2IP
ERBB3
ERBB4
ERC1
ERC2
ERCC1
ERCC2
ERCC4
ERCC6
ERCC8
ERGIC1
ERGIC2
ERH
ERI2
ERI3
ERICH1
ERLIN1
ERMAP
ERMN
ERN1
ERN2
ERRFI1
ESAM
ESPN
ESPNL
ESPNP
ESR1
ESR2
ESRP1
ESRP2
ESRRB
ESRRG
ESX1
ESYT2
ETF1
ETFA
ETNK2
ETS1
ETV3
ETV4
ETV5
ETV6
EVC2
EVI5
EVI5L
EVL
EVPL
EVX1
EVX2
EWSR1
EXD2
EXD3
EXO1
EXOC2
EXOC3
EXOC3L
EXOC3L2
EXOC4
EXOC8
EXOSC10
EXOSC7
EXT1
EXT2
EXTL3
EYA4
EYS
EZH1
EZR
F10
F11R
F12
F2RL1
F2RL3
F7
F8
F9
FAAH
FABP1
FABP2
FABP3
FABP6
FABP9
FADS2
FAH
FAHD2A
FAHD2B
FAIM
FAIM3
FAM100B
FAM101B
FAM102B
FAM103A1
FAM105A
FAM107B
FAM109B
FAM110A
FAM111B
FAM113B
FAM116A
FAM117A
FAM118A
FAM120A
FAM122B
FAM122C
FAM123B
FAM123C
FAM124A
FAM125B
FAM126A
FAM126B
FAM127B
FAM128A
FAM129A
FAM129B
FAM129C
FAM131A
FAM131C
FAM132A
FAM134B
FAM136A
FAM13A
FAM149A
FAM151B
FAM153B
FAM155A
FAM155B
FAM159A
FAM159B
FAM160A1
FAM160A2
FAM160B1
FAM160B2
FAM162A
FAM163A
FAM164C
FAM166B
FAM167A
FAM167B
FAM168A
FAM168B
FAM169B
FAM170A
FAM170B
FAM171A1
FAM171A2
FAM172A
FAM173B
FAM176A
FAM177B
FAM178B
FAM179A
FAM180B
FAM181A
FAM181B
FAM184A
FAM184B
FAM185A
FAM186A
FAM188B
FAM189A1
FAM18A
FAM190A
FAM190B
FAM192A
FAM193B
FAM195A
FAM195B
FAM196B
FAM19A1
FAM19A2
FAM19A3
FAM19A4
FAM19A5
FAM20A
FAM20B
FAM20C
FAM21A
FAM22D
FAM25A
FAM26D
FAM27C
FAM27L
FAM32A
FAM36A
FAM38A
FAM3B
FAM3C
FAM3D
FAM40A
FAM41C
FAM43A
FAM43B
FAM45B
FAM46A
FAM46B
FAM46C
FAM47A
FAM47E
FAM48A
FAM49A
FAM50B
FAM53B
FAM53C
FAM54A
FAM55C
FAM55D
FAM57B
FAM58A
FAM58B
FAM59A
FAM5B
FAM60A
FAM63A
FAM63B
FAM64A
FAM65A
FAM65B
FAM65C
FAM66B
FAM69A
FAM69C
FAM71E2
FAM73B
FAM76A
FAM78A
FAM81A
FAM82A1
FAM82A2
FAM82B
FAM83A
FAM83B
FAM83H
FAM84B
FAM89A
FAM91A1
FAM92A1
FAM96A
FAM96B
FAM98C
FAM9A
FANCA
FANCC
FANCD2
FANCI
FANK1
FAP
FAR2
FARP1
FARS2
FARSA
FARSB
FAS
FASN
FASTK
FASTKD1
FAT1
FAT3
FATE1
FBLIM1
FBLL1
FBLN1
FBLN2
FBLN7
FBN1
FBN2
FBN3
FBP1
FBRS
FBRSL1
FBXL12
FBXL15
FBXL16
FBXL17
FBXL18
FBXL19
FBXL2
FBXL20
FBXL21
FBXL3
FBXL6
FBXL7
FBXL8
FBXO11
FBXO16
FBXO17
FBXO21
FBXO24
FBXO25
FBXO28
FBXO3
FBXO31
FBXO32
FBXO34
FBXO4
FBXO41
FBXO42
FBXO44
FBXO45
FBXO46
FBXO6
FBXO8
FBXO9
FBXW11
FBXW7
FBXW8
FBXW9
FCAMR
FCAR
FCER2
FCGBP
FCGR2B
FCGR3A
FCGRT
FCHO1
FCHSD1
FCHSD2
FCN2
FCRL2
FCRL4
FCRLA
FDFT1
FDX1
FDXACB1
FDXR
FECH
FEM1A
FEM1B
FEN1
FER
FER1L5
FERMT2
FERMT3
FES
FEV
FEZ1
FEZ2
FEZF2
FGB
FGD1
FGD4
FGD6
FGF1
FGF11
FGF12
FGF13
FGF14
FGF17
FGF20
FGF22
FGF23
FGF3
FGF4
FGF8
FGF9
FGFBP1
FGFR1
FGFR1OP
FGFR2
FGFR3
FGFR4
FGFRL1
FGGY
FGR
FHAD1
FHDC1
FHIT
FHL1
FHL2
FHL3
FHL5
FHOD1
FHOD3
FIBIN
FICD
FIGLA
FIGN
FIS1
FITM1
FIZ1
FKBP10
FKBP14
FKBP15
FKBP1A
FKBP5
FKBP9L
FKBPL
FLAD1
FLCN
FLI1
FLII
FLJ10357
FLJ10661
FLJ11235
FLJ12825
FLJ13197
FLJ22536
FLJ23834
FLJ25006
FLJ26850
FLJ30058
FLJ31306
FLJ32063
FLJ32810
FLJ33360
FLJ35024
FLJ35390
FLJ36031
FLJ37453
FLJ39582
FLJ39609
FLJ39653
FLJ40330
FLJ40434
FLJ41350
FLJ41941

FLJ42709
FLJ43663
FLJ43950
FLJ44606
FLJ44817
FLJ45079
FLJ45244
FLJ45340
FLJ45983
FLJ90757
FLNA
FLNB
FLNC
FLOT1
FLOT2
FLRT2
FLRT3
FLT1
FLT3
FLT3LG
FLT4
FLVCR1
FLVCR2
FLYWCH1
FLYWCH2
FMN2
FMNL1
FMNL2
FMO1
FMO3
FMR1NB
FN3KRP
FNBP1
FNDC1
FNDC3A
FNDC3B
FNDC4
FNDC7
FNIP1
FNIP2
FNTA
FOLH1
FOLH1B
FOLR2
FOLR3
FOS
FOSB
FOSL1
FOSL2
FOXA3
FOXB2
FOXC1
FOXC2
FOXD3
FOXE3
FOXF1
FOXF2
FOXG1
FOXH1
FOXI2
FOXI3
FOXJ2
FOXJ3
FOXK1
FOXK2
FOXM1
FOXN3
FOXN4
FOXO1
FOXO3
FOXP1
FOXP2
FOXP3
FOXP4
FOXR1
FPR1
FPR2
FRAS1
FRAT1
FRAT2
FREM2
FREM3
FREQ
FRK
FRMD1
FRMD4A
FRMD4B
FRMD5
FRMPD1
FRMPD2
FRMPD4
FRS2
FRS3
FRY
FRZB
FSCB
FSCN2
FSD1
FSTL1
FSTL4
FTH1
FTHL17
FTL
FTMT
FTSJ1
FTSJD1
FUBP3
FUCA2
FUK
FUNDC2
FUS
FUT10
FUT11
FUT3
FUT4
FUT5
FUT6
FUT8
FUZ
FXN
FXR1
FXYD1
FXYD4
FXYD5
FXYD6
FYCO1
FYN
FZD2
FZD3
FZD4
FZD5
FZD7
FZD8
FZD9
FZR1
G0S2
G2E3
G3BP1
G3BP2
G6PC
G6PC3
GAA
GAB4
GABARAP
GABBR1
GABBR2
GABPB2
GABRA1
GABRA2
GABRA3
GABRA5
GABRA6
GABRB1
GABRB2
GABRB3
GABRD
GABRG2
GABRG3
GABRP
GABRQ
GABRR1
GABRR2
GAD1
GADD45A
GADD45GIP1
GAGE10
GAGE2A
GAK
GAL
GAL3ST4
GALE
GALK1
GALK2
GALNS
GALNT10
GALNT12
GALNT14
GALNT2
GALNT5
GALNT7
GALNT9
GALNTL2
GALNTL4
GALNTL5
GALNTL6
GALP
GANAB
GAP43
GAPVD1
GAR1
GARNL3
GARS
GAS1
GAS2L2
GAS7
GAST
GATA1
GATA2
GATA3
GATA4
GATA5
GATA6
GATAD2A
GATAD2B
GATM
GATS
GATSL1
GBA
GBE1
GBF1
GBGT1
GBP4
GBP7
GBX1
GC
GCA
GCAT
GCDH
GCET2
GCG
GCK
GCLC
GCLM
GCM1
GCM2
GCN1L1
GCNT2
GCNT7
GDAP1
GDAP1L1
GDAP2
GDE1
GDEP
GDF1
GDF2
GDF5
GDF6
GDF7
GDF9
GDI1
GDNF
GDPD1
GDPD5
GEFT
GEM
GEN1
GFAP
GFI1
GFOD2
GFPT1
GFPT2
GFRA1
GFRA2
GFRA3
GFRA4
GGA2
GGCX
GGN
GGNBP2
GGPS1
GGT1
GGT7
GGT8P
GHRL
GHRLOS
GIGYF1
GIMAP1
GIMAP4
GIMAP8
GIN1
GINS1
GIPC1
GIPC3
GIPR
GIT1
GIT2
GJA3
GJA4
GJA5
GJB5
GJB6
GJC1
GJD3
GJD4
GK5
GLA
GLB1L
GLB1L3
GLDC
GLDN
GLG1
GLI2
GLI3
GLIS1
GLIS2
GLIS3
GLOD4
GLOD5
GLP2R
GLRA1
GLRA3
GLRX2
GLRX3
GLRX5
GLS2
GLT1D1
GLT25D1
GLT25D2
GLT8D1
GLTP
GLTPD1
GLTPD2
GLTSCR1
GLUD2
GLUL
GLYATL3
GLYCTK
GLYR1
GM2A
GMCL1
GMCL1L
GMDS
GMEB1
GMEB2
GMFB
GML
GMPR
GMPR2
GMPS
GNA11
GNA12
GNAL
GNAO1
GNAQ
GNAS
GNAT1
GNAT2
GNAZ
GNB1
GNB1L
GNB2
GNB2L1
GNB3
GNB5
GNG12
GNG2
GNG4
GNG7
GNL1
GNL2
GNL3
GNL3L
GNLY
GNMT
GNPDA1
GNPTAB
GNRHR
GOLGA2
GOLGA2L1
GOLGA3
GOLGA4
GOLGA6A
GOLGA7B
GOLGA8B
GOLIM4
GOLM1
GOLPH3
GOLSYN
GOLT1A
GOPC
GORASP1
GOSR2
GOT1
GP1BB
GP2
GP5
GP9
GPATCH4
GPBAR1
GPBP1L1
GPC2
GPC3
GPC4
GPC6
GPD1
GPD1L
GPD2
GPHN
GPI
GPIHBP1
GPKOW
GPLD1
GPM6A
GPN1
GPN2
GPNMB
GPR101
GPR109A
GPR111
GPR114
GPR120
GPR123
GPR124
GPR132
GPR133
GPR135
GPR137
GPR137B
GPR142
GPR143
GPR155
GPR156
GPR157
GPR158
GPR160
GPR161
GPR172A
GPR173
GPR176
GPR177
GPR182
GPR183
GPR19
GPR25
GPR27
GPR3
GPR31
GPR35
GPR39
GPR4
GPR45
GPR50
GPR56
GPR62
GPR63
GPR68
GPR77
GPR78
GPR85
GPR88
GPR89B
GPR98
GPRASP2
GPRC5B
GPRC5C
GPRIN2
GPSM1
GPSM2
GPSM3
GPT
GPT2
GPX1
GPX6
GPX7
GRAMD1A
GRAMD1B
GRAMD3
GRAMD4
GRAPL
GRASP
GRB10
GRB2
GRB7
GREM2
GRHL1
GRHL2
GRHL3
GRIA2
GRID1
GRID2
GRID2IP
GRIK1
GRIK3
GRIK4
GRIN1
GRIN2A
GRIN2B
GRIN2C
GRIN2D
GRIN3A
GRINA
GRIP1
GRIP2
GRK1
GRK4
GRK5
GRK6
GRK7
GRM1
GRM4
GRM6
GRM7
GRM8
GRP
GRPR
GRRP1
GRSF1
GRTP1
GRWD1
GRXCR2
GSC
GSC2
GSDMB
GSDMD
GSG1
GSG1L
GSK3A
GSN
GSPT1
GSR
GSTA4
GSTCD
GSTM3
GSTO2
GSX1
GTDC1
GTF2A1
GTF2A2
GTF2B
GTF2E2
GTF2F1
GTF2F2
GTF2H3
GTF2H4
GTF2H5
GTF2IRD2B
GTF3A
GTF3C1
GTF3C2
GTPBP1
GTPBP10
GTPBP3
GTPBP8
GUCA2B
GUCY1A2
GUCY1B2
GUCY1B3
GUCY2C
GUCY2D
GUCY2E
GUF1
GUK1
GULP1
GUSB
GUSBL1
GXYLT1
GXYLT2
GYG1
GYLTL1B
GYPA
GYPC
GYS1
GZMM
H19
H1F0
H1FNT
H1FX
H2AFJ
H2AFY2
H3F3B
H6PD
HABP2
HADH
HAGH
HAL
HAMP
HAND1
HAND2
HAP1
HAPLN1
HAPLN2
HAPLN3
HAPLN4
HAR1B
HARS2
HAS1
HAS2AS
HAS3
HAT1
HAUS2
HAUS3
HAUS4
HAUS5
HAX1
HBA1
HBB
HBM
HBP1
HBS1L
HCCA2
HCCS
HCFC1
HCFC1R1
HCG18
HCG22
HCG27
HCG2P7
HCG4
HCG9
HCK
HCLS1
HCN1
HCN3

R1
R2
R3
R4
R5
R6
R7
R8
R9
R10
R11
R12
R13
R14
R15
R16
R17
R18
R19
R20
R21
R22
R23
R24
R25
R26
R27
R28
R29
R30
R31
R32
R33
R34
R35
R36
R37
R38
R39

HCN4
HCRT
HCRTR2
HDAC1
HDAC11
HDAC2
HDAC3
HDAC4
HDAC6
HDAC7
HDAC9
HDC
HDGF
HDGF2
HDGFL1
HDGFRP3
HDLBP
HDX
HEATR1
HEATR2
HEATR7A
HECA
HECW1
HECW2
HEG1
HELLS
HELT
HELZ
HEMK1
HEPACAM
HEPH
HEPHL1
HERC2
HERC3
HERC5
HES1
HES3
HES7
HEXDC
HEXIM1
HEY2
HEYL
HFM1
HGC6.3
HGS
HGSNAT
HHAT
HHEX
HHIPL1
HIAT1
HIBADH
HIC1
HIC2
HIF1A
HIGD1A
HIGD1B
HINFP
HINT1
HINT2
HIP1
HIP1R
HIPK1
HIPK2
HIPK3
HIPK4
HIRIP3
HIST1H1A
HIST1H1D
HIST1H2AI
HIST1H2AM
HIST1H2BC
HIST1H2BD
HIST1H2BE
HIST1H2BH
HIST1H2BK
HIST1H2BN
HIST1H3A
HIST1H3D
HIST1H3E
HIST1H3G
HIST1H3H
HIST1H4D
HIST1H4G
HIST1H4L
HIST2H2AB
HIST2H2AC
HIST2H2BF
HIVEP1
HIVEP2
HIVEP3
HK1
HK2
HK3
HKDC1
HLA-A
HLA-B
HLA-DMB
HLA-DOA
HLA-DOB
HLA-DPA1
HLA-DPB1
HLA-DPB2
HLA-DQA2
HLA-DQB1
HLA-DQB2
HLA-DRA
HLA-E
HLA-F
HLA-G
HLA-H
HLA-J
HLA-L
HLTF
HLX
HM13
HMBOX1
HMBS
HMCN1
HMG20A
HMGA2
HMGB1
HMGB2
HMGB3
HMGB3L1
HMGCLL1
HMGN4
HMHA1
HMHB1
HMOX2
HMP19
HMSD
HMX2
HNF1B
HNF4A
HNMT
HNRNPA1
HNRNPA1L2
HNRNPA2B1
HNRNPA3
HNRNPAB
HNRNPC
HNRNPD
HNRNPF
HNRNPH1
HNRNPK
HNRNPU
HNRNPUL1
HNRPA1L-2
HOMER3
HOOK1
HOOK3
HOTAIR
HOXA1
HOXA10
HOXA11AS
HOXA13
HOXA2
HOXA3
HOXA4
HOXA7
HOXB13
HOXB2
HOXB3
HOXB7
HOXB8
HOXC10
HOXC11
HOXC12
HOXC13
HOXC4
HOXC9
HOXD1
HOXD10
HOXD13
HOXD3
HOXD4
HOXD8
HOXD9
HPCA
HPCAL4
HPDL
HPRT1
HPS3
HPSE2
HPYR1
HRASLS5
HRC
HRH1
HRH3
HRK
HRNBP3
HS2ST1
HS3ST1
HS3ST2
HS3ST3A1
HS3ST3B1
HS3ST4
HS3ST5
HS6ST1
HS6ST2
HS6ST3
HSBP1
HSBP1L1
HSD11B1
HSD11B1L
HSD11B2
HSD17B1
HSD17B10
HSD17B13
HSD17B2
HSD17B4
HSD3B2
HSD3B7
HSF1
HSF4
HSF5
HSFYP1
HSP90AA1
HSP90AB1
HSPA12B
HSPA13
HSPA1A
HSPA1B
HSPA1L
HSPA2
HSPA4
HSPA6
HSPA7
HSPA8
HSPB1
HSPB2
HSPB6
HSPB8
HSPBAP1
HSPBP1
HSPC072
HSPC159
HSPD1
HSPE1
HSPG2
HTATIP2
HTATSF1
HTR1E
HTR1F
HTR2A
HTR2B
HTR2C
HTR3A
HTR3C
HTR3D
HTR5A
HTR6
HTR7P
HTRA1
HTRA3
HTT
HULC
HUWE1
HYAL1
HYAL2
HYAL3
HYDIN
HYLS1
HYMAI
IAH1
IARS2
IBSP
IBTK
ICA1
ICA1L
ICAM3
ICAM5
ICK
ICOS
ID1
ID4
IDH1
IDH3G
IDO2
IDS
IDUA
IER2
IER3
IER5
IER5L
IFFO1
IFFO2
IFI27
IFI27L1
IFI27L2
IFI44L
IFIH1
IFIT1
IFIT5
IFLTD1
IFNA8
IFNAR1
IFRD2
IFT122
IFT140
IFT74
IFT88
IGDCC3
IGDCC4
IGF1R
IGF2BP1
IGF2BP2
IGF2BP3
IGFBP2
IGFBP3
IGFBP4
IGFBP5
IGFBP6
IGFBP7
IGFL2
IGFN1
IGHMBP2
IGLL3
IGLON5
IGSF10
IGSF11
IGSF21
IGSF22
IGSF3
IGSF8
IGSF9
IGSF9B
IHH
IKBIP
IKBKE
IKBKG
IKZF1
IKZF5
IL10RB
IL11
IL12A
IL12B
IL12RB2
IL13
IL13RA1
IL15
IL15RA
IL16
IL17B
IL17C
IL17F
IL17RA
IL17RC
IL17RD
IL17RE
IL18BP
IL19
IL1A
IL1B
IL1R2
IL1RAP
IL1RAPL1
IL1RL1
IL1RL2
IL20
IL20RA
IL21R
IL23A
IL27
IL27RA
IL28RA
IL2RB
IL31
IL4
IL4I1
IL4R
IL5RA
IL6
IL6R
IL7
IL9
ILDR1
ILDR2
ILF2
ILF3
ILKAP
IMMP2L
IMP3
IMPA2
IMPG1
IMPG2
INADL
INCA1
INF2
ING1
ING5
INHBC
INMT
INO80
INO80D
INO80E
INPP1
INPP4A
INPP4B
INPP5A
INPP5B
INPP5D
INPP5F
INPP5K
INPPL1
INSC
INSIG1
INS-IGF2
INSL3
INSL6
INSM1
INSR
INTS1
INTS12
INTS3
INTS4
INTS4L1
INTS5
INTS6
INTS7
INTS8
IP6K1
IP6K3
IPCEF1
IPMK
IPO11
IPO13
IPO4
IPO9
IPW
IQCA1
IQCB1
IQCC
IQCD
IQCE
IQCF6
IQCG
IQCJ
IQGAP2
IQSEC1
IQSEC2
IQSEC3
IQUB
IRAK1BP1
IRAK2
IRAK3
IREB2
IRF2
IRF2BP1
IRF2BP2
IRF3
IRF4
IRF5
IRF6
IRF7
IRF8
IRS1
IRS2
IRS4
IRX1
IRX2
IRX3
IRX4
IRX5
ISG20
ISL1
ISL2
ISLR2
ISM1
ISPD
ITFG3
ITGA1
ITGA2
ITGA2B
ITGA5
ITGA7
ITGA8
ITGA9
ITGAE
ITGAL
ITGAM
ITGB1
ITGB1BP2
ITGB1BP3
ITGB2
ITGB3
ITGB4
ITGB6
ITGB7
ITGB8
ITIH5
ITIH5L
ITK
ITLN1
ITM2A
ITM2B
ITM2C
ITPK1
ITPKA
ITPKB
ITPR1
ITPR2
ITPR3
ITPRIP
ITSN1
ITSN2
IVD
IYD
JAG1
JAG2
JAK1
JAK2
JAK3
JAKMIP1
JAKMIP2
JAKMIP3
JAM3
JARID2
JAZF1
JDP2
JMJD1C
JMJD5
JMJD7-PLA2G4B
JOSD1
JPH2
JPH3
JRK
JRKL
JUB
JUN
JUP
KAAG1
KALRN
KANK2
KANK3
KAT2A
KAT2B
KAT5
KATNAL1
KATNAL2
KAZALD1
KBTBD10
KBTBD11
KBTBD3
KBTBD5
KCNA10
KCNA2
KCNA3
KCNA4
KCNA7
KCNAB2
KCNB1
KCNC1
KCNC2
KCNC3
KCNC4
KCND3
KCNE1
KCNE1L
KCNF1
KCNG2
KCNG3
KCNG4
KCNH1
KCNH2
KCNH3
KCNH4
KCNH6
KCNIP1
KCNIP2
KCNIP3
KCNIP4
KCNJ1
KCNJ10
KCNJ11
KCNJ12
KCNJ14
KCNJ15
KCNJ16
KCNJ2
KCNJ5
KCNJ6
KCNJ8
KCNJ9
KCNK1
KCNK12
KCNK13
KCNK16
KCNK17
KCNK2
KCNK3
KCNK4
KCNK7
KCNK9
KCNMA1
KCNMB2
KCNMB3
KCNMB4
KCNN1
KCNN2
KCNN3
KCNN4
KCNQ1
KCNQ1OT1
KCNQ2
KCNQ3
KCNQ4
KCNQ5
KCNRG
KCNS1
KCNS3
KCNT1
KCP
KCTD1
KCTD11
KCTD12
KCTD13
KCTD15
KCTD18
KCTD2
KCTD5
KCTD6
KCTD8
KDM1B
KDM2A
KDM2B
KDM4A
KDM4B
KDM6A
KDM6B
KDSR
KEAP1
KEL
KHDC1
KHDC1L
KHDRBS2
KHK
KHNYN
KHSRP
KIAA0020
KIAA0040
KIAA0090
KIAA0101
KIAA0146
KIAA0182
KIAA0195
KIAA0196
KIAA0226
KIAA0247
KIAA0284
KIAA0319L
KIAA0355
KIAA0391
KIAA0406
KIAA0415
KIAA0430
KIAA0467
KIAA0494
KIAA0495
KIAA0528
KIAA0556
KIAA0562
KIAA0564
KIAA0586
KIAA0652
KIAA0664
KIAA0892
KIAA0895
KIAA0895L

KIAA0907
KIAA0913
KIAA0922
KIAA1009
KIAA1012
KIAA1024
KIAA1026
KIAA1147
KIAA1161
KIAA1191
KIAA1199
KIAA1211
KIAA1217
KIAA1244
KIAA1257
KIAA1274
KIAA1279
KIAA1310
KIAA1324
KIAA1324L
KIAA1377
KIAA1407
KIAA1409
KIAA1429
KIAA1467
KIAA1468
KIAA1486
KIAA1529
KIAA1530
KIAA1543
KIAA1549
KIAA1598
KIAA1614
KIAA1632
KIAA1688
KIAA1712
KIAA1737
KIAA1751
KIAA1841
KIAA1875
KIAA1949
KIAA1984
KIAA2013
KIAA2018
KIDINS220
KIF11
KIF13A
KIF13B
KIF14
KIF15
KIF16B
KIF18B
KIF19
KIF1A
KIF1B
KIF1C
KIF20B
KIF21B
KIF25
KIF26A
KIF26B
KIF3A
KIF3C
KIF4A
KIF5A
KIF5C
KIF6
KIF7
KIFAP3
KIFC1
KIFC3
KIR3DL2
KIR3DL3
KIRREL
KIRREL2
KIRREL3
KISS1
KIT
KITLG
KLB
KLC2
KLC4
KLF1
KLF10
KLF12
KLF13
KLF14
KLF15
KLF16
KLF2
KLF3
KLF6
KLF7
KLF9
KLHDC1
KLHDC10
KLHDC3
KLHDC7B
KLHDC8A
KLHDC9
KLHL10
KLHL13
KLHL14
KLHL15
KLHL17
KLHL2
KLHL21
KLHL23
KLHL25
KLHL26
KLHL29
KLHL30
KLHL32
KLHL33
KLHL34
KLHL35
KLHL7
KLHL8
KLK10
KLK15
KLKBL4
KLRAQ1
KLRG2
KMO
KNDC1
KNG1
KPNA2
KPNA4
KPNA6
KPNA7
KPNB1
KPTN
KRAS
KRBA1
KRBA2
KRCC1
KREMEN1
KRR1
KRT2
KRT23
KRT24
KRT26
KRT28
KRT3
KRT31
KRT33A
KRT39
KRT6A
KRT7
KRT72
KRT75
KRT79
KRT81
KRT86
KRTAP10-3
KRTAP12-4
KRTAP1-3
KRTAP13-1
KRTAP15-1
KRTAP19-1
KRTAP19-5
KRTAP20-1
KRTAP20-2
KRTAP24-1
KRTAP4-1
KRTAP5-9
KRTAP6-3
KRTAP9-2
KRTDAP
KSR1
KSR2
KTN1
KY
L1CAM
L1TD1
L3MBTL
LACTB
LAG3
LAIR1
LAMA3
LAMA4
LAMA5
LAMB1
LAMB2
LAMB3
LAMB4
LAMC1
LAMC2
LAMC3
LAMP2
LAMP3
LANCL2
LANCL3
LAPTM4A
LAPTM4B
LAPTM5
LARP1
LARP4B
LARS2
LASP1
LASS1
LASS2
LASS4
LASS5
LASS6
LAT
LAT2
LATS1
LATS2
LAYN
LBP
LBX1
LBX2
LBXCOR1
LCAT
LCE1E
LCE3E
LCK
LCLAT1
LCMT1
LCN10
LCN15
LCN8
LCOR
LCP1
LCTL
LDB1
LDB2
LDHA
LDHAL6B
LDHB
LDHC
LDLR
LDLRAD2
LDLRAD3
LDLRAP1
LDOC1L
LEAP2
LECT1
LEF1
LEFTY1
LEFTY2
LEMD3
LENG1
LENG8
LEO1
LEP
LEPR
LEPRE1
LEPREL1
LEPROTL1
LETM1
LETM2
LETMD1
LEUTX
LFNG
LGALS12
LGALS13
LGALS3
LGALS7
LGALS7B
LGALS8
LGALS9C
LGI2
LGI3
LGI4
LGMN
LGR4
LGR5
LGR6
LGTN
LHB
LHFP
LHFPL2
LHFPL3
LHFPL5
LHPP
LHX1
LHX2
LHX4
LHX6
LHX8
LHX9
LIFR
LIG4
LILRA2
LILRA4
LILRA6
LILRB1
LILRB3
LILRP2
LIMA1
LIMCH1
LIMD1
LIMD2
LIMK2
LIMS2
LIN28
LIN37
LIN7A
LINGO1
LINGO3
LINS1
LIPA
LIPE
LIPH
LITAF
LIX1
LIX1L
LLGL2
LMAN2
LMBR1L
LMBRD1
LMF1
LMNA
LMO2
LMO3
LMO4
LMO7
LMOD2
LMTK2
LMTK3
LMX1A
LMX1B
LNP1
LNPEP
LNX1
LOC100101938
LOC100125556
LOC100128164
LOC100128239
LOC100128292
LOC100128542
LOC100128640
LOC100128675
LOC100128731
LOC100128788
LOC100128811
LOC100128977
LOC100129550
LOC100129637
LOC100129716
LOC100130093
LOC100130274
LOC100130331
LOC100130581
LOC100130691
LOC100130776
LOC100130872
LOC100130872-SPON2
LOC100130987
LOC100132215
LOC100132963
LOC100133091
LOC100133161
LOC100133308
LOC100133612
LOC100133893
LOC100133957
LOC100133991
LOC100134259
LOC100134368
LOC100134713
LOC100134868
LOC100144603
LOC100144604
LOC100169752
LOC100188947
LOC100188949
LOC100190939
LOC100192378
LOC100268168
LOC100271715
LOC100271836
LOC100272217
LOC100286844
LOC100302401
LOC100302652
LOC100329109
LOC113230
LOC116437
LOC121952
LOC127841
LOC134466
LOC139201
LOC144776
LOC145474
LOC145663
LOC145814
LOC145837
LOC145845
LOC146880
LOC148696
LOC150185
LOC150527
LOC150568
LOC150622
LOC151174
LOC151534
LOC152217
LOC153328
LOC158376
LOC221710
LOC222699
LOC253724
LOC254559
LOC256880
LOC257358
LOC283050
LOC283999
LOC284009
LOC284688
LOC284749
LOC284788
LOC284798
LOC284805
LOC285033
LOC285205
LOC285375
LOC285401
LOC285419
LOC285456
LOC285501
LOC285548
LOC285550
LOC285593
LOC285692
LOC285780
LOC285796
LOC285830
LOC285954
LOC286016
LOC286238
LOC338799
LOC339290
LOC339524
LOC339674
LOC340074
LOC340094
LOC344595
LOC344967
LOC375196
LOC387646
LOC388428
LOC388789
LOC388796
LOC389332
LOC389333
LOC389493
LOC389634
LOC390594
LOC390858
LOC391322
LOC399959
LOC400657
LOC400696
LOC400759
LOC400794
LOC400804
LOC400927
LOC400931
LOC400940
LOC401010
LOC401097
LOC401127
LOC401463
LOC404266
LOC407835

LOC440173
LOC440461
LOC440839
LOC441046
LOC441204
LOC441208
LOC442308
LOC492303
LOC494141
LOC541473
LOC554203
LOC619207
LOC641367
LOC642006
LOC642587
LOC642597
LOC643719
LOC643837
LOC644538
LOC644649
LOC645323
LOC645961
LOC646498
LOC646627
LOC646762
LOC646982
LOC647288
LOC647979
LOC648691
LOC650368
LOC650623
LOC651250
LOC652276
LOC653113
LOC727677
LOC728264
LOC728276
LOC728392
LOC728448
LOC728554
LOC728613
LOC728743
LOC728855
LOC729080
LOC729082
LOC729375
LOC729384
LOC729991-MEF2B
LOC730668
LOC731275
LOC731779
LOC731789
LOC84740
LOC84931
LOC90246
LOC91316
LOC91948
LOH12CR1
LONP1
LONRF1
LONRF3
LOX
LOXHD1
LOXL1
LOXL2
LOXL3
LPA
LPAR1
LPAR3
LPAR4
LPAR5
LPAR6
LPCAT1
LPGAT1
LPHN1
LPHN2
LPHN3
LPIN1
LPIN2
LPL
LPO
LPP
LPPR2
LPPR3
LPPR4
LPPR5
LRAT
LRCH1
LRCH3
LRDD
LRFN5
LRG1
LRIG1
LRIG2
LRIG3
LRIT2
LRP1
LRP11
LRP3
LRP4
LRP5
LRP5L
LRP6
LRP8
LRPAP1
LRPPRC
LRRC1
LRRC10
LRRC10B
LRRC14B
LRRC15
LRRC16A
LRRC18
LRRC20
LRRC23
LRRC24
LRRC27
LRRC28
LRRC29
LRRC3
LRRC30
LRRC33
LRRC36
LRRC37A3
LRRC37B
LRRC37B2
LRRC41
LRRC43
LRRC47
LRRC49
LRRC4B
LRRC6
LRRC61
LRRC66
LRRC69
LRRC7
LRRC8A
LRRC8B
LRRC8D
LRRC8E
LRRFIP1
LRRFIP2
LRRIQ3
LRRIQ4
LRRK1
LRRN1
LRRN2
LRRN3
LRRN4
LRRN4CL
LRSAM1
LRTM1
LRTM2
LSAMP
LSM12
LSM14B
LSM2
LSM5
LSM6
LSP1
LST1
LTA
LTB4R
LTB4R2
LTBP1
LTBP4
LTF
LTK
LUC7L
LUC7L2
LUC7L3
LUZP1
LUZP2
LUZP6
LVRN
LXN
LY6D
LY6E
LY6G5C
LY6G6E
LY75
LY86
LY9
LYAR
LYN
LYPD5
LYPD6B
LYPLA2
LYRM2
LYRM4
LYRM5
LYRM7
LYSMD2
LYSMD3
LYST
LYZL2
LZTFL1
LZTR1
LZTS2
M6PR
MAB21L1
MACF1
MACROD1
MACROD2
MAD1L1
MAD2L1BP
MADD
MAEA
MAEL
MAFB
MAFF
MAFG
MAFK
MAG
MAGEA5
MAGEA6
MAGEB10
MAGEB18
MAGEB2
MAGEB3
MAGEB4
MAGEB6
MAGEC1
MAGEC2
MAGEC3
MAGED1
MAGED2
MAGEE1
MAGEE2
MAGEH1
MAGI1
MAGI2
MAGI3
MAGIX
Magmas
MAGOH
MAGT1
MAL
MAL2
MALAT1
MALL
MAMDC4
MAML1
MAML2
MAML3
MAMSTR
MAN1A2
MAN1C1
MAN2A2
MAN2B2
MANBA
MANBAL
MANEAL
MANF
MAOB
MAP1A
MAP1D
MAP1LC3B2
MAP1LC3C
MAP1S
MAP2
MAP2K1
MAP2K2
MAP2K3
MAP2K4
MAP2K6
MAP2K7
MAP3K10
MAP3K11
MAP3K15
MAP3K3
MAP3K4
MAP3K5
MAP3K6
MAP3K7
MAP3K7IP2
MAP3K8
MAP3K9
MAP4
MAP4K1
MAP4K2
MAP4K3
MAP4K4
MAP4K5
MAP6
MAP7D1
MAPK10
MAPK11
MAPK12
MAPK15
MAPK4
MAPK7
MAPK8IP2
MAPK8IP3
MAPK9
MAPKAP1
MAPKAPK3
MAPKAPK5
MAPKBP1
MAPRE1
MAPRE2
MAPRE3
MAPT
MARK1
MARK2
MARK3
MARS
MARS2
MARVELD2
MARVELD3
MAS1L
MASP1
MAST1
MAST2
MAST3
MAST4
MAT2B
MATK
MATN1
MATN2
MATN3
MATN4
MATR3
MAVS
MAX
MAZ
MBD2
MBD3
MBD6
MBIP
MBLAC1
MBNL2
MBOAT2
MBOAT7
MBP
MC1R
MC3R
MC4R
MCART6
MCAT
MCC
MCCC1
MCF2
MCF2L
MCF2L2
MCHR2
MCM2
MCM3
MCM5
MCM7
MCM8
MCM9
MCOLN1
MCPH1
MCTP1
MCTS1
MDC1
MDFI
MDGA1
MDGA2
MDH2
MDK
MDM1
MDP1
ME3
MECOM
MECR
MED1
MED11
MED12L
MED13L
MED14
MED15
MED19
MED20
MED21
MED23
MED24
MED28
MED30
MED9
MEF2C
MEF2D
MEFV
MEG3
MEG8
MEGF10
MEGF11
MEGF6
MEGF8
mei-01
MEIG1
MEIS2
MELK
MEMO1
MEN1
MEOX2
MEPCE
MERTK
MESDC1
MESP1
MESP2
MEST
MESTIT1
MET
METAP2
METRN
METRNL
METT10D
METT5D1
METTL10
METTL14
METTL2A
METTL2B
METTL5
METTL7B
METTL8
METTL9
MEX3A
MEX3B
MEX3C
MFAP1
MFAP2
MFAP3L
MFAP4
MFHAS1
MFI2
MFN1
MFN2
MFSD1
MFSD11
MFSD2B
MFSD3
MFSD4
MFSD5
MFSD6
MFSD7
MFSD9
MGA
MGAT3
MGAT4B
MGAT4C
MGAT5
MGC14436
MGC15885
MGC16121
MGC16275
MGC16384
MGC23284
MGC26597
MGC27382
MGC42105
MGC45800
MGC57346
MGC70857
MGC87042
MGEA5
MGLL
MGMT
MGRN1
MGST1
MIA
MIA3
MIB2
MICA
MICAL1
MICAL2
MICAL3
MICALCL
MICALL2
MICB
MID1IP1
MIER2
MIF4GD
MIMT1
MINK1
MINPP1
MIPEP
MIR100
MIR1182
MIR1185-2
MIR1205
MIR1229
MIR1237
MIR124-2
MIR124-3
MIR1256
MIR1259
MIR125A
MIR1266
MIR127
MIR128-2
MIR1289-2
MIR1306
MIR132
MIR138-1
MIR141
MIR143
MIR1468
MIR152
MIR153-2
MIR1539
MIR182
MIR1826
MIR183
MIR1908
MIR1909
MIR191
MIR1912
MIR1913
MIR192
MIR193A
MIR193B
MIR194-2
MIR196A2
MIR196B
MIR19A
MIR2052
MIR21
MIR2110
MIR219-2
MIR30A
MIR320D1
MIR328
MIR329-2
MIR330
MIR33A
MIR340
MIR345
MIR365-1
MIR365-2
MIR376B
MIR377
MIR384
MIR410
MIR423
MIR487A
MIR492
MIR493
MIR494
MIR495
MIR505
MIR507
MIR508
MIR510
MIR514-2
MIR515-2
MIR516A1
MIR516B2
MIR517C
MIR519B
MIR519E
MIR520B
MIR525
MIR548A2
MIR548F3
MIR548F5
MIR548H4
MIR548N
MIR574
MIR589
MIR596
MIR598
MIR600
MIR611
MIR645
MIR654
MIR665
MIR670
MIR675
MIR708
MIR711
MIR7-3
MIR760
MIR888
MIR922
MIR9-3
MIRLET7C
MIRLET7G
MIRLET7I
MITF
MIXL1
MKL1
MKL2
MKLN1
MKNK2
MKRN2
MKRN3
MKX
MLC1
MLEC
MLF1IP
MLH1
MLL
MLL3
MLL4
MLL5
MLLT1
MLLT3
MLLT4
MLNR
MLX
MLXIP
MLXIPL
MLYCD
MMACHC
MMADHC
MMD
MME
MMEL1
MMGT1
MMP1
MMP14
MMP15
MMP17
MMP23A
MMP25
MMP28
MMP3
MMP9
MMRN2
MND1
MNDA
MNT
MOBKL1A
MOBKL1B
MOBKL2A
MOBKL2B
MOBKL2C
MOBKL3
MOG
MOGAT3
MOGS
MON1A
MON1B
MORC4
MORF4
MORN1
MOSC1
MOSC2
MOV10
MOXD1
MOXD2
MPDU1
MPEG1
MPG
MPHOSPH8
MPHOSPH9
MPL
MPND
MPO
MPP1
MPP2
MPP6
MPPE1
MPPED1
MPRIP
MRAP
MRAP2
MRAS
MRC2
MRE11A
MREG
MRGPRD
MRGPRG
MRM1
MRO
MRPL13
MRPL15
MRPL16

MRPL19
MRPL20
MRPL21
MRPL22
MRPL23
MRPL32
MRPL33
MRPL38
MRPL48
MRPL52
MRPS11
MRPS12
MRPS14
MRPS16
MRPS17
MRPS18A
MRPS18B
MRPS22
MRPS23
MRPS26
MRPS27
MRPS33
MRPS35
MRPS36
MRS2
MRTO4
MRVI1
MS4A13
MS4A2
MS4A3
MS4A6A
MSC
MSH2
MSH3
MSH6
MSI1
MSI2
MSL2
MSL3
MSL3L2
MSLN
MSN
MSRA
MSRB3
MST1R
MST4
MSX1
MT1DP
MT1E
MT1F
MT1G
MT1H
MT1X
MT2A
MT4
MTA1
MTCH1
MTCH2
MTERF
MTERFD2
MTERFD3
MTF1
MTFR1
MTG1
MTHFD1L
MTHFR
MTHFS
MTHFSD
MTIF3
MTM1
MTMR10
MTMR12
MTMR15
MTMR4
MTMR6
MTMR8
MTMR9
MTMR9L
MTNR1A
MTNR1B
MTP18
MTSS1
MTSS1L
MTUS1
MTUS2
MTVR2
MTX1
MTX2
MTX3
MUC12
MUC15
MUC2
MUC21
MUC4
MUC5B
MUC6
MUC7
MUL1
MUM1
MUM1L1
MUPCDH
MURC
MUS81
MUSTN1
MVD
MXD4
MXRA5
MXRA7
MXRA8
MYADM
MYADML
MYB
MYBBP1A
MYBL2
MYBPC1
MYBPC2
MYBPC3
MYBPH
MYCBP2
MYCBPAP
MYCL1
MYCN
MYEOV2
MYF5
MYH10
MYH11
MYH14
MYH15
MYH16
MYH3
MYH4
MYH6
MYH7
MYH7B
MYL1
MYL12A
MYL2
MYL3
MYL5
MYL6
MYL6B
MYL9
MYLIP
MYLK
MYLK4
MYNN
MYO10
MYO15A
MYO15B
MYO16
MYO18A
MYO18B
MYO1B
MYO1C
MYO1D
MYO1E
MYO1F
MYO1G
MYO1H
MYO3B
MYO5B
MYO5C
MYO7A
MYO7B
MYO9B
MYOCD
MYOD1
MYOF
MYOG
MYOM2
MYOM3
MYOT
MYOZ2
MYOZ3
MYPN
MYPOP
MYST2
MYST4
MYT1
MYT1L
MZF1
N4BP2
N4BP2L1
N4BP2L2
N4BP3
N6AMT2
NAA15
NAA16
NAA20
NAA30
NAA40
NAA50
NAAA
NAALADL1
NAALADL2
NAB1
NACAD
NACAP1
NACC1
NACC2
NADK
NADSYN1
NAE1
NAF1
NAGA
NAGPA
NAGS
NAMPT
NANOS3
NAP1L1
NAP1L4
NAPA
NAPEPLD
NARF
NARFL
NARS2
NAT1
NAT10
NAT14
NAT15
NAT2
NAT8L
NAV1
NAV2
NAV3
NBEAL1
NBEAL2
NBL1
NBLA00301
NBN
NBPF14
NBPF3
NCAM1
NCAN
NCAPD2
NCAPD3
NCAPG2
NCAPH
NCBP2
NCCRP1
NCDN
NCF2
NCK1
NCK2
NCKAP5
NCKAP5L
NCL
NCLN
NCOA1
NCOA2
NCOA4
NCOA6
NCOR1
NCOR2
NCR1
NCRNA00081
NCRNA00087
NCRNA00092
NCRNA00094
NCRNA00095
NCRNA00110
NCRNA00111
NCRNA00114
NCRNA00115
NCRNA00119
NCRNA00120
NCRNA00160
NCRNA00162
NCRNA00164
NCRNA00167
NCRNA00171
NCRNA00173
NCRNA00174
NCRNA00175
NCRNA00176
NCRNA00181
NCRNA00188
NCRNA00200
NCRNA00207
NCSTN
NDE1
NDEL1
NDFIP1
NDN
NDRG1
NDRG2
NDRG4
NDST1
NDST2
NDUFA13
NDUFA4L2
NDUFA5
NDUFA6
NDUFA7
NDUFA9
NDUFAF2
NDUFB1
NDUFB10
NDUFB11
NDUFB3
NDUFB5
NDUFB8
NDUFC1
NDUFC2
NDUFS2
NDUFS3
NDUFS4
NDUFS6
NDUFS7
NDUFV1
NDUFV2
NEBL
NECAB1
NECAB2
NECAP1
NEDD4
NEDD4L
NEDD9
NEFL
NEFM
NEGR1
NEIL1
NEIL2
NEIL3
NEK1
NEK10
NEK11
NEK5
NEK6
NELF
NELL1
NELL2
NEO1
NEU1
NEU2
NEU4
NEURL
NEURL1B
NEURL2
NEURL3
NEURL4
NEUROD1
NEUROD2
NEUROD6
NEUROG1
NEUROG2
NEXN
NF1
NF2
NFAM1
NFASC
NFAT5
NFATC1
NFATC2
NFATC4
NFE2L2
NFE2L3
NFIA
NFIC
NFIL3
NFIX
NFKB1
NFKB2
NFKBIA
NFKBIE
NFKBIL1
NFKBIL2
NFKBIZ
NFRKB
NFU1
NFXL1
NFYA
NFYB
NFYC
NGDN
NGEF
NGF
NGFR
NGFRAP1
NHEDC1
NHEDC2
NHEJ1
NHLH1
NHLH2
NHLRC3
NHLRC4
NHP2
NHP2L1
NHS
NHSL1
NHSL2
NID1
NID2
NIF3L1
NIN
NINJ2
NIPAL2
NIPAL3
NIPAL4
NIPBL
NIPSNAP1
NIPSNAP3A
NISCH
NIT1
NIT2
NKAIN2
NKAIN3
NKAIN4
NKAP
NKD1
NKD2
NKPD1
NKTR
NKX1-2
NKX2-1
NKX2-2
NKX2-4
NKX2-5
NKX3-1
NKX6-1
NKX6-2
NKX6-3
NLGN1
NLGN2
NLGN3
NLGN4X
NLRC3
NLRC5
NLRP1
NLRP11
NLRP12
NLRP3
NLRP6
NLRP7
NMB
NMD3
NME1
NME2
NME5
NME7
NMI
NMNAT1
NMNAT2
NMRAL1
NMT1
NMT2
NMUR1
NNAT
NNMT
NOD1
NOD2
NODAL
NOL11
NOL12
NOL3
NOL4
NOL7
NOM1
NOMO1
NONO
NOP14
NOP16
NOP56
NOP58
NOS1
NOS1AP
NOS2
NOS3
NOSTRIN
NOTCH1
NOTCH2
NOTCH3
NOTCH4
NOTO
NOTUM
NOV
NOVA1
NOXA1
NPAS1
NPAS2
NPAS3
NPAS4
NPBWR2
NPC1L1
NPEPPS
NPFF
NPFFR1
NPFFR2
NPHP1
NPHP4
NPHS2
NPL
NPLOC4
NPM1
NPM3
NPNT
NPPA
NPPB
NPPC
NPR1
NPTN
NPTX1
NPTX2
NPY
NPY1R
NPY2R
NPY5R
NQO1
NR0B2
NR1D1
NR1D2
NR1H2
NR1H3
NR1I2
NR1I3
NR2C1
NR2E1
NR2F1
NR2F6
NR3C1
NR4A1
NR4A3
NR5A1
NR5A2
NR6A1
NRAP
NRARP
NRBF2
NRBP1
NRBP2
NRCAM
NRD1
NRF1
NRG1
NRG2
NRG3

NRG4
NRGN
NRIP2
NRN1
NRP2
NRSN1
NRSN2
NRXN1
NRXN2
NRXN3
NSA2
NSD1
NSFL1C
NSL1
NSMCE2
NSMCE4A
NSUN2
NSUN4
NSUN6
NSUN7
NT5DC1
NT5DC2
NTAN1
NTF3
NTF4
NTHL1
NTM
NTN1
NTN4
NTN5
NTNG1
NTRK1
NTRK3
NTS
NTSR1
NTSR2
NUAK1
NUAK2
NUB1
NUBP1
NUBP2
NUBPL
NUCB2
NUCKS1
NUDC
NUDCD1
NUDCD3
NUDT10
NUDT16
NUDT16P
NUDT18
NUDT19
NUDT3
NUDT6
NUDT9P1
NUF2
NUMA1
NUMB
NUMBL
NUP107
NUP133
NUP188
NUP205
NUP210
NUP210L
NUP43
NUP50
NUP54
NUP62
NUP62CL
NUP85
NUP93
NUPL1
NUPL2
NUPR1
NUS1
NUSAP1
NWD1
NXF5
NXN
NXNL1
NXPH1
NXPH2
NXPH3
NXPH4
NXT1
NYNRIN
OAS1
OAS2
OAZ1
OAZ2
OAZ3
OBFC1
OBFC2A
OBSCN
OBSL1
OCA2
OCIAD2
OCLN
OCM
OCM2
OCRL
ODC1
ODF2L
ODF3L2
ODZ2
ODZ3
ODZ4
OGDH
OGDHL
OGFR
OGFRL1
OGT
OLA1
OLFM1
OLFM3
OLFM4
OLFML1
OLFML2B
OLFML3
OLIG1
OLIG2
OMA1
ONECUT1
ONECUT2
ONECUT3
OPA1
OPA3
OPALIN
OPCML
OPLAH
OPN1LW
OPN3
OPN5
OPRK1
OPRL1
OPTC
OPTN
OR10A3
OR10A5
OR10A7
OR10AG1
OR10G4
OR10G8
OR10H1
OR10H4
OR10H5
OR10J1
OR10V1
OR11L1
OR12D2
OR12D3
OR14J1
OR1A1
OR1A2
OR2B2
OR2C1
OR2C3
OR2G2
OR2H1
OR2J2
OR2L13
OR2L1P
OR2M1P
OR2T3
OR2T34
OR2T6
OR2W1
OR2Y1
OR4A47
OR4K2
OR4N3P
OR51T1
OR52A4
OR52B4
OR52B6
OR52J3
OR52K2
OR52W1
OR56A1
OR56A5
OR5A1
OR5AC2
OR5C1
OR5E1P
OR5H1
OR5L2
OR5T3
OR6B1
OR6B3
OR6K2
OR6T1
OR7C1
OR7E37P
OR7G1
OR7G2
OR7G3
OR8A1
OR8G2
OR8H1
ORAI2
ORC3L
ORC4L
ORC6L
ORMDL2
ORMDL3
OSBP2
OSBPL10
OSBPL1A
OSBPL2
OSBPL3
OSBPL5
OSBPL8
OSCAR
OSCP1
OSGIN1
OSR1
OSR2
OSTalpha
OSTBETA
OSTC
OSTCL
OSTM1
OTOA
OTOF
OTOL1
OTOP1
OTOP2
OTOP3
OTP
OTUB1
OTUD3
OTUD4
OTUD5
OTUD7B
OTX1
OTX2
OTX2OS1
OVCA2
OVCH2
OVOL1
OXA1L
OXCT1
OXER1
OXGR1
OXR1
OXSR1
OXT
P2RX1
P2RX3
P2RX4
P2RX5
P2RX6
P2RX7
P2RY1
P2RY6
P4HA1
P4HA2
P4HB
P4HTM
PA2G4
PA2G4P4
PABPC1
PABPC4
PABPC4L
PABPN1
PACRG
PACS1
PACS2
PACSIN1
PACSIN2
PADI1
PADI3
PADI6
PAEP
PAF1
PAFAH1B1
PAFAH1B2
PAG1
PAGE2
PAGE2B
PAGE3
PAICS
PAIP1
PAIP2
PAIP2B
PAK1
PAK2
PAK3
PAK4
PAK6
PALB2
PALLD
PALM
PALM2
PAMR1
PAN2
PANK3
PANK4
PANX1
PANX2
PAOX
PAPD4
PAPD5
PAPL
PAPLN
PAPOLA
PAPOLB
PAPOLG
PAPPA
PAPSS2
PAQR5
PARD3
PARD3B
PARD6A
PARD6B
PARD6G
PARK7
PARM1
PARN
PARP1
PARP10
PARP14
PARP16
PARP6
PARP9
PARS2
PART1
PARVA
PARVB
PARVG
PASD1
PATL2
PAWR
PAX1
PAX2
PAX3
PAX4
PAX5
PAX6
PAX7
PAX8
PAX9
PBRM1
PBX1
PBX2
PBX3
PBX4
PBXIP1
PC
PCBD2
PCBP1
PCBP2
PCBP3
PCBP4
PCCA
PCDH1
PCDH10
PCDH15
PCDH19
PCDH21
PCDH24
PCDH7
PCDH8
PCDH9
PCDHA2
PCDHA6
PCDHA7
PCDHAC1
PCDHB10
PCDHB11
PCDHB15
PCDHB17
PCDHB18
PCDHB5
PCDHGA1
PCDHGA2
PCDHGA4
PCF11
PCGF1
PCGF2
PCGF3
PCGF5
PCIF1
PCLO
PCM1
PCMT1
PCMTD2
PCNA
PCNT
PCNXL2
PCNXL3
PCOTH
PCP2
PCSK1
PCSK1N
PCSK5
PCSK6
PCSK7
PCSK9
PCYT1A
PCYT1B
PCYT2
PDCD1
PDCD11
PDCD5
PDCD6
PDCL2
PDE10A
PDE11A
PDE12
PDE1A
PDE1B
PDE1C
PDE2A
PDE3A
PDE3B
PDE4A
PDE4B
PDE4C
PDE4D
PDE4DIP
PDE5A
PDE6A
PDE6B
PDE6G
PDE7B
PDE8A
PDE8B
PDGFA
PDGFB
PDGFD
PDGFRA
PDGFRL
PDHA2
PDHB
PDHX
PDIA2
PDIA3
PDIA4
PDIA6
PDK2
PDK3
PDK4
PDLIM3
PDP1
PDPK1
PDPN
PDS5A
PDS5B
PDSS2
PDX1
PDXDC1
PDXK
PDYN
PDZD2
PDZD4
PDZD7
PDZRN3
PDZRN4
PEAR1
PEBP4
PECR
PEF1
PEG10
PEG3
PELI1
PELI2
PELI3
PELP1
PEMT
PENK
PEPD
PER1
PER2
PERP
PES1
PET112L
PEX10
PEX11G
PEX14
PEX19
PEX5
PF4V1
PFAS
PFDN1
PFDN4
PFKFB1
PFKFB2
PFKFB3
PFKFB4
PFKL
PFKM
PFKP
PFN2
PFN3
PGAM5
PGAP2
PGAP3
PGBD4
PGBD5
PGCP
PGGT1B
PGK1
PGK2
PGLS
PGLYRP1
PGLYRP2
PGM1
PGM5
PGRMC2
PGS1
PHACTR1
PHACTR2
PHACTR3
PHACTR4
PHB
PHB2
PHC3
PHEX
PHF1
PHF10
PHF11
PHF12
PHF13
PHF14
PHF15
PHF17
PHF21A
PHF21B
PHF23
PHF6
PHGDH
PHGR1
PHIP
PHKA1
PHKA2
PHKB
PHLDA2
PHLDB1
PHLDB2
PHLDB3
PHLPP1
PHOX2A
PHOX2B
PHRF1
PHTF2
PHYH
PHYHIP
PHYHIPL
PI15
PI4K2A
PI4KA
PI4KAP1
PI4KB
PIAS1
PIAS3
PIAS4
PICK1
PID1
PIF1
PIGA
PIGC
PIGG
PIGM
PIGO
PIGQ
PIGR
PIGU
PIGV
PIGW
PIGX
PIGY
PIGZ
PIH1D2
PIK3C2B
PIK3C2G
PIK3CA
PIK3CB
PIK3CD
PIK3CG
PIK3IP1
PIK3R1
PIK3R2
PIK3R3
PIK3R5
PIK3R6
PIKFYVE
PILRA
PILRB
PIM1
PIN1
PINX1
PIP4K2A
PIP5K1A
PIP5K1B
PIP5K1C
PIP5KL1
PIPOX
PISD
PITPNA
PITPNC1
PITPNM1
PITPNM2
PITPNM3
PITRM1
PITX1
PITX2
PITX3
PIWIL2
PJA2
PKD1
PKD1L1
PKD1L3
PKD2L2
PKHD1
PKHD1L1
PKLR
PKM2
PKN3
PKNOX2
PKP1
PKP3
PKP4
PL-5283
PLA2G12A
PLA2G15
PLA2G2A
PLA2G2E
PLA2G4C
PLA2G4D
PLA2G4E
PLA2G6
PLA2G7
PLAC1
PLAC2
PLAC8
PLAGL1
PLAT
PLAUR
PLBD1
PLBD2
PLCB4
PLCD1
PLCD3
PLCD4
PLCE1
PLCH1
PLCH2
PLCL1
PLCL2
PLCXD2
PLCZ1
PLD1
PLD2
PLD4
PLD5
PLD6
PLDN
PLEC1
PLEK
PLEK2
PLEKHA1
PLEKHA2
PLEKHA5
PLEKHA7
PLEKHA9
PLEKHF1
PLEKHG1
PLEKHG3

PLEKHG4
PLEKHG4B
PLEKHG5
PLEKHG7
PLEKHH1
PLEKHH2
PLEKHH3
PLEKHM1
PLEKHN1
PLEKHO1
PLEKHO2
PLIN1
PLIN2
PLIN3
PLK1S1
PLK2
PLK5P
PLLP
PLOD2
PLP2
PLS3
PLTP
PLXDC1
PLXNA1
PLXNA3
PLXNA4
PLXNB1
PLXNB2
PLXNB3
PLXNC1
PLXND1
PMEPA1
PMF1
PMFBP1
PML
PMM1
PMP22
PMPCB
PMS2CL
PMS2L11
PMS2L3
PMS2L5
PNCK
PNKD
PNLDC1
PNLIPRP1
PNLIPRP2
PNMA1
PNMA2
PNMA5
PNMAL1
PNMAL2
PNN
PNOC
PNP
PNPLA2
PNPLA6
PNPLA7
PNRC1
PNRC2
PODN
PODNL1
PODXL
POF1B
POFUT1
POGZ
POLA1
POLA2
POLD1
POLDIP2
POLDIP3
POLE
POLE2
POLE3
POLE4
POLG
POLG2
POLI
POLK
POLL
POLN
POLQ
POLR1A
POLR1B
POLR1E
POLR2B
POLR2E
POLR2H
POLR2I
POLR2J
POLR3A
POLR3B
POLR3E
POLR3F
POLR3GL
POLR3K
POLRMT
POLS
POM121
POM121C
POM121L12
POM121L2
POM121L9P
POMC
POMGNT1
POMP
POMT2
POMZP3
POP1
POP4
POP7
POR
PORCN
POTEE
POTEF
POU2F1
POU2F2
POU2F3
POU3F2
POU3F3
POU3F4
POU4F1
POU6F2
PP14571
PPA1
PPA2
PPAP2A
PPAP2B
PPAP2C
PPAPDC1A
PPAPDC1B
PPARA
PPARD
PPARG
PPARGC1B
PPAT
PPBPL1
PPFIA1
PPFIA3
PPFIA4
PPFIBP1
PPFIBP2
PPHLN1
PPIC
PPID
PPIH
PPIL1
PPIL2
PPIL4
PPIL5
PPL
PPM1E
PPM1F
PPM1H
PPM1K
PPM1L
PPM1M
PPME1
PPP1CA
PPP1CC
PPP1R10
PPP1R11
PPP1R12B
PPP1R12C
PPP1R13L
PPP1R14A
PPP1R14B
PPP1R16B
PPP1R2P1
PPP1R2P3
PPP1R3B
PPP1R3C
PPP1R3E
PPP1R3F
PPP1R3G
PPP1R7
PPP1R8
PPP1R9A
PPP2CA
PPP2R2A
PPP2R2B
PPP2R2C
PPP2R2D
PPP2R3A
PPP2R5A
PPP2R5C
PPP2R5D
PPP2R5E
PPP3CB
PPP3CC
PPP3R1
PPP4C
PPP4R1L
PPPDE1
PPRC1
PPT1
PPT2
PPTC7
PPY
PPY2
PQLC1
PQLC2
PRAM1
PRAME
PRAMEF20
PRB3
PRB4
PRCC
PRCD
PRCP
PRDM1
PRDM10
PRDM11
PRDM13
PRDM15
PRDM16
PRDM2
PRDM4
PRDM6
PRDM7
PRDM8
PRDX1
PRDX2
PRDX5
PREB
PRELP
PREP
PREPL
PREX1
PREX2
PRG2
PRG4
PRH1
PRHOXNB
PRIC285
PRICKLE1
PRICKLE2
PRICKLE3
PRICKLE4
PRIM2
PRKAB1
PRKACA
PRKAG2
PRKAR1B
PRKAR2A
PRKAR2B
PRKCA
PRKCB
PRKCD
PRKCE
PRKCH
PRKCI
PRKCQ
PRKCZ
PRKD2
PRKDC
PRKG1
PRKRA
PRKRIP1
PRKRIR
PRKX
PRLH
PRLHR
PRM1
PRM2
PRM3
PRMT1
PRMT10
PRMT5
PRMT7
PRMT8
PRND
PRNP
PRNT
PROC
PROCA1
PROCR
PROKR1
PROM2
PROX1
PRPF18
PRPF31
PRPF38B
PRPF39
PRPF8
PRPS1L1
PRPS2
PRPSAP1
PRPSAP2
PRR12
PRR18
PRR24
PRR25
PRR3
PRR4
PRRC1
PRRG3
PRRT1
PRRT2
PRRT3
PRRX1
PRRX2
PRSS12
PRSS22
PRSS23
PRSS27
PRSS36
PRSS37
PRSS38
PRSS8
PRTN3
PRX
PSAP
PSAPL1
PSCA
PSD2
PSD3
PSD4
PSEN1
PSEN2
PSENEN
PSG2
PSG8
PSIP1
PSKH1
PSKH2
PSMA4
PSMA5
PSMA6
PSMA7
PSMB10
PSMB11
PSMB4
PSMB5
PSMB6
PSMB7
PSMB8
PSMB9
PSMC1
PSMC2
PSMC3IP
PSMC5
PSMD1
PSMD12
PSMD13
PSMD14
PSMD2
PSMD3
PSMD5
PSMD6
PSMD8
PSMD9
PSME1
PSME2
PSME3
PSMG2
PSMG3
PSORS1C1
PSORS1C3
PSPN
PSRC1
PSTPIP1
PSTPIP2
PTBP1
PTCD1
PTCD2
PTCH1
PTCH2
PTCHD1
PTCHD2
PTCRA
PTDSS2
PTEN
PTF1A
PTGDR
PTGDS
PTGER1
PTGER2
PTGER3
PTGES
PTGS2
PTH
PTH1R
PTK2B
PTK6
PTK7
PTN
PTOV1
PTP4A1
PTP4A3
PTPDC1
PTPLA
PTPLB
PTPN1
PTPN11
PTPN12
PTPN14
PTPN18
PTPN3
PTPN5
PTPN6
PTPN7
PTPRA
PTPRB
PTPRC
PTPRCAP
PTPRE
PTPRF
PTPRG
PTPRJ
PTPRK
PTPRM
PTPRN
PTPRN2
PTPRO
PTPRQ
PTPRR
PTPRS
PTPRU
PTRF
PTRH1
PTS
PUF60
PUM1
PURG
PUS1
PUS10
PUS7
PUSL1
PVALB
PVR
PVRL1
PVRL3
PVRL4
PVT1
PWP2
PWWP2B
PXDN
PXDNL
PXK
PXMP2
PXMP3
PXMP4
PXN
PYCARD
PYCR1
PYCR2
PYCRL
PYGM
PYY
PYY2
QARS
QPCT
QPRT
QRICH1
QRICH2
QRSL1
QSOX1
QSOX2
QTRT1
R3HCC1
R3HDM1
R3HDML
RAB11A
RAB11FIP1
RAB11FIP3
RAB11FIP4
RAB11FIP5
RAB12
RAB15
RAB19
RAB1A
RAB1B
RAB20
RAB21
RAB22A
RAB23
RAB27B
RAB28
RAB2B
RAB30
RAB31
RAB35
RAB37
RAB39
RAB39B
RAB3C
RAB3D
RAB3GAP1
RAB3GAP2
RAB3IP
RAB40AL
RAB40B
RAB40C
RAB42
RAB43
RAB4A
RAB4B
RAB6A
RAB6B
RAB7A
RAB8B
RAB9B
RABAC1
RABEP1
RABGEF1
RABGGTA
RABL2A
RAC1
RAC3
RACGAP1
RAD17
RAD18
RAD21
RAD51AP2
RAD51L1
RAD52
RAD54L
RAD9A
RAD9B
RADIL
RAE1
RAET1E
RAET1G
RAF1
RAG1AP1
RAGE
RAI1
RAI14
RALA
RALB
RALBP1
RALGAPA2
RALGDS
RALGPS2
RALYL
RAMP1
RAMP3
RANBP17
RANBP3
RANBP9
RANGAP1
RANGRF
RAP1A
RAP1B
RAP1GAP
RAP1GAP2
RAP1GDS1
RAPGEF3
RAPGEF4
RAPGEF5
RAPGEF6
RAPGEFL1
RAPSN
RARA
RARB
RARG
RARRES2
RASA1
RASA3
RASAL1
RASAL2
RASAL3
RASD1
RASD2
RASGEF1A
RASGEF1B
RASGEF1C
RASGRF1
RASGRF2
RASGRP1
RASGRP2
RASL10A
RASL10B
RASL11B
RASSF1
RASSF2
RASSF3
RASSF4
RASSF5
RASSF6
RASSF7
RAVER1
RAX
RB1
RBBP6
RBCK1
RBL1
RBM10
RBM12B
RBM14
RBM15
RBM15B
RBM16
RBM17
RBM18
RBM19
RBM20
RBM22
RBM24
RBM25
RBM26
RBM28

RBM33
RBM34
RBM38
RBM39
RBM4
RBM41
RBM42
RBM44
RBM47
RBM4B
RBM5
RBM6
RBM9
RBMS1
RBMS2
RBMS3
RBMX
RBMXL1
RBMXL3
RBP1
RBP2
RBP4
RBP7
RBPJ
RBPMS
RBPMS2
RBX1
RC3H2
RCAN1
RCC2
RCCD1
RCHY1
RCN1
RCN2
RCOR1
RCOR2
RCSD1
RCVRN
RD3
RDBP
RDH10
RDH12
RDH14
RDX
RECQL4
RECQL5
REEP1
REEP2
REEP3
REEP4
REEP5
REEP6
REG1A
REG3G
REL
RELA
RELL1
RELN
RELT
REM1
REM2
RENBP
REP15
REPIN1
REPS1
RER1
RERE
RERG
RESP18
REST
RET
RETN
REV1
REXO1
REXO1L2P
REXO2
RFC1
RFC4
RFC5
RFESD
RFPL3S
RFPL4A
RFPL4B
RFTN2
RFWD2
RFWD3
RFX1
RFX4
RFX5
RFX7
RG9MTD2
RGAG1
RGL1
RGL2
RGL3
RGMA
RGN
RGNEF
RGS10
RGS12
RGS13
RGS14
RGS18
RGS19
RGS2
RGS20
RGS3
RGS5
RGS7
RGS9
RHBDD2
RHBDD3
RHBDF1
RHBDF2
RHBDL2
RHBG
RHCG
RHEB
RHOB
RHOBTB2
RHOBTB3
RHOD
RHOG
RHOT1
RHOU
RHOV
RHPN1
RHPN2
RIBC1
RIC8A
RICH2
RICS
RICTOR
RIF1
RILPL1
RIMBP2
RIMBP3
RIMKLA
RIMKLB
RIMS1
RIMS2
RIMS4
RIN1
RIN2
RIN3
RING1
RINL
RIPK1
RIPK3
RIPK4
RIPPLY2
RIT2
RLF
RMND5A
RNASE10
RNASE4
RNASE9
RNASEH1
RNASEK
RNF113A
RNF125
RNF126P1
RNF128
RNF13
RNF135
RNF138
RNF139
RNF14
RNF144A
RNF144B
RNF145
RNF149
RNF165
RNF166
RNF169
RNF17
RNF170
RNF182
RNF183
RNF185
RNF186
RNF19B
RNF2
RNF208
RNF212
RNF213
RNF216
RNF220
RNF222
RNF24
RNF25
RNF31
RNF34
RNF39
RNF4
RNF41
RNF44
RNF5
RNF6
RNFT1
RNFT2
RNGTT
RNLS
RNMTL1
RNPEPL1
RNU11
RNU5E
RNU6ATAC
ROBLD3
ROBO1
ROBO2
ROBO3
ROCK1
ROCK2
ROPN1B
ROR1
ROR2
RORA
RORC
ROS1
RP1
RP1L1
RP2
RPA1
RPA2
RPAIN
RPAP3
RPE
RPE65
RPF1
RPGR
RPGRIP1
RPH3A
RPH3AL
RPIA
RPL10
RPL10A
RPL13
RPL13AP17
RPL13AP20
RPL13AP5
RPL14
RPL17
RPL18AP3
RPL19
RPL19P12
RPL21
RPL22
RPL22L1
RPL23A
RPL23AP82
RPL23P8
RPL24
RPL26
RPL28
RPL29
RPL3
RPL36AL
RPL38
RPL39
RPL41
RPL5
RPLP0
RPLP1
RPN1
RPN2
RPP14
RPP21
RPP40
RPRD2
RPRM
RPS10
RPS11
RPS15
RPS18
RPS19
RPS19BP1
RPS23
RPS24
RPS29
RPS5
RPS6KA1
RPS6KA2
RPS6KA3
RPS6KA4
RPS6KB2
RPS6KC1
RPS6KL1
RPS8
RPS9
RPSA
RPSAP58
RPTOR
RPUSD2
RRAD
RRBP1
RREB1
RRM1
RRN3P1
RRN3P3
RRP1
RRP12
RRP7B
RSAD1
RSAD2
RSL24D1
RSPH1
RSPH10B
RSPH9
RSPO1
RSPO2
RSPO3
RSRC2
RSU1
RTBDN
RTDR1
RTF1
RTKN
RTN1
RTN4
RTN4IP1
RTN4R
RTN4RL2
RTP1
RTP2
RUFY1
RUFY4
RUNDC1
RUNDC3A
RUNX1
RUNX1T1
RUNX3
RUSC1
RUSC2
RWDD1
RXFP3
RXFP4
RXRA
RXRB
RYK
RYR1
RYR2
RYR3
S100A13
S100A2
S100A3
S100A4
S100A5
S100A6
S100A8
S100P
S100PBP
S1PR1
S1PR2
S1PR5
SAC3D1
SACS
SAFB
SAFB2
SAG
SALL1
SALL2
SAMD12
SAMD13
SAMD14
SAMD4A
SAMD5
SAMD9
SAMM50
SAP130
SAP18
SAP25
SAP30L
SAPS2
SAR1B
SARDH
SARM1
SARNP
SARS
SARS2
SASH1
SASH3
SAT1
SAT2
SATB1
SATB2
SAV1
SBF1
SBF2
SBK1
SBK2
SBNO2
SC4MOL
SCAF1
SCAI
SCAMP1
SCAMP3
SCAMP4
SCAND1
SCAND3
SCAP
SCAPER
SCARB1
SCARB2
SCARF1
SCARF2
SCARNA11
SCCPDH
SCD5
SCFD2
SCG2
SCG5
SCGB2A1
SCGB2A2
SCGN
SCHIP1
SCMH1
SCML1
SCML2
SCN11A
SCN1B
SCN2B
SCN3B
SCN4A
SCN4B
SCN5A
SCNM1
SCNN1B
SCOC
SCPEP1
SCRIB
SCRN1
SCRT1
SCRT2
SCTR
SCUBE1
SCUBE2
SDAD1
SDC1
SDC2
SDCBP
SDCBP2
SDCCAG10
SDCCAG8
SDF2
SDF2L1
SDF4
SDHA
SDHAP1
SDHAP2
SDK1
SDK2
SDR16C6
SDR9C7
SEC1
SEC13
SEC14L1
SEC14L3
SEC14L4
SEC14L5
SEC16A
SEC22B
SEC23A
SEC23IP
SEC24B
SEC24C
SEC24D
SEC31A
SEC31B
SEC61A1
SEC61A2
SEC61G
SEC63
SECISBP2L
SECTM1
SEL1L
SEL1L2
SEL1L3
SELI
SELK
SELO
SELP
SELPLG
SEMA3A
SEMA3B
SEMA3C
SEMA3G
SEMA4B
SEMA4C
SEMA4F
SEMA4G
SEMA5A
SEMA5B
SEMA6A
SEMA6B
SEMA6C
SEMA6D
SEMA7A
SEMG1
SENP1
SENP2
SENP3
SEPHS2
SEPN1
SEPP1
SEPSECS
SEPX1
SERAC1
SERBP1
SERF2
SERGEF
SERINC2
SERINC3
SERINC4
SERP2
SERPINA1
SERPINA11
SERPINA12
SERPINA13
SERPINA3
SERPINA4
SERPINA7
SERPINB11
SERPIND1
SERPINE1
SERPINF1
SERPINF2
SERPING1
SERPINH1
SERPINI1
SERPINI2
SERTAD2
SERTAD3
SESN2
SESTD1
SET
SETBP1
SETD1A
SETD1B
SETD2
SETD3
SETD4
SETD7
SETDB2
SETMAR
SEZ6
SEZ6L2
SF1
SF3A1
SF3A2
SF3B1
SF3B4
SFMBT1
SFMBT2
SFN
SFRP1
SFRP2
SFRP4
SFRP5
SFRS12
SFRS13A
SFRS13B
SFRS18
SFRS2B
SFRS3
SFRS4
SFRS5
SFRS6
SFRS8
SFT2D1
SFT2D2
SFTA2
SFTA3
SFTPA1
SFXN2
SFXN3
SGCB
SGCD
SGCE
SGCZ
SGEF
SGIP1
SGK1
SGK269
SGK3
SGMS1
SGOL1
SGPP1
SGPP2
SGSH
SGSM1
SGSM2
SGTA
SGTB
SH2B1
SH2B2
SH2B3
SH2D1A
SH2D2A
SH2D3C
SH2D4B
SH2D6
SH2D7
SH3BGRL
SH3BGRL2
SH3BP2
SH3BP4
SH3BP5
SH3BP5L
SH3D20
SH3GL1
SH3GLB2
SH3KBP1
SH3PXD2A
SH3PXD2B
SH3RF1
SH3RF2
SH3RF3
SH3TC1
SH3TC2
SH3YL1
SHANK1
SHANK2
SHANK3
SHARPIN
SHC1
SHC2
SHC3
SHC4
SHD
SHF
SHH
SHISA3
SHISA4
SHISA6
SHISA7
SHISA9
SHMT1
SHMT2
SHOC2
SHOX2
SHQ1
SHROOM1
SHROOM2
SHROOM3
SHROOM4
SIAE
SIDT1
SIGIRR
SIGLEC1
SIGLEC10
SIGLEC12
SIGLEC6
SIGLEC9
SIGMAR1
SIK3
SIL1
SIM1
SIM2
SIN3A
SIN3B
SIPA1
SIPA1L1
SIPA1L2
SIRPA
SIRPB2
SIRPG
SIRT2
SIRT5
SIRT6

SIRT7
SIT1
SIX3
SIX4
SIX6
SKA2
SKAP2
SKI
SKIL
SKINTL
SKIV2L
SKP1
SKP2
SLA2
SLAIN1
SLAIN2
SLAMF1
SLAMF8
SLAMF9
SLBP
SLC10A2
SLC10A4
SLC10A6
SLC10A7
SLC11A1
SLC11A2
SLC12A1
SLC12A2
SLC12A4
SLC12A5
SLC12A6
SLC12A7
SLC12A8
SLC13A4
SLC15A1
SLC15A3
SLC15A4
SLC16A10
SLC16A13
SLC16A2
SLC16A3
SLC16A4
SLC16A5
SLC16A6
SLC16A7
SLC16A8
SLC16A9
SLC17A5
SLC17A6
SLC17A7
SLC18A1
SLC18A2
SLC19A1
SLC19A3
SLC1A2
SLC1A4
SLC1A6
SLC1A7
SLC20A2
SLC22A1
SLC22A11
SLC22A12
SLC22A13
SLC22A16
SLC22A17
SLC22A18
SLC22A18AS
SLC22A20
SLC22A23
SLC22A24
SLC22A4
SLC22A6
SLC22A8
SLC22A9
SLC23A1
SLC23A3
SLC24A1
SLC24A4
SLC25A10
SLC25A13
SLC25A15
SLC25A16
SLC25A17
SLC25A20
SLC25A21
SLC25A22
SLC25A23
SLC25A25
SLC25A26
SLC25A29
SLC25A30
SLC25A31
SLC25A33
SLC25A35
SLC25A37
SLC25A39
SLC25A42
SLC25A44
SLC25A45
SLC25A46
SLC26A10
SLC26A11
SLC26A2
SLC26A4
SLC26A5
SLC26A8
SLC27A1
SLC27A3
SLC28A3
SLC29A1
SLC29A2
SLC29A3
SLC29A4
SLC2A1
SLC2A12
SLC2A13
SLC2A14
SLC2A2
SLC2A3
SLC2A5
SLC2A6
SLC2A7
SLC2A9
SLC30A1
SLC30A2
SLC30A4
SLC30A8
SLC33A1
SLC34A2
SLC35A1
SLC35B1
SLC35B3
SLC35C1
SLC35C2
SLC35D1
SLC35E3
SLC35E4
SLC35F1
SLC35F3
SLC36A1
SLC36A2
SLC36A4
SLC37A1
SLC37A4
SLC38A1
SLC38A10
SLC38A3
SLC38A4
SLC38A5
SLC38A6
SLC38A7
SLC38A8
SLC38A9
SLC39A14
SLC39A4
SLC39A6
SLC39A7
SLC39A8
SLC3A1
SLC40A1
SLC41A3
SLC43A1
SLC43A2
SLC44A2
SLC44A4
SLC44A5
SLC45A1
SLC45A4
SLC46A1
SLC46A2
SLC46A3
SLC47A1
SLC4A1
SLC4A2
SLC4A3
SLC4A4
SLC4A5
SLC4A8
SLC5A1
SLC5A10
SLC5A11
SLC5A5
SLC5A7
SLC5A8
SLC5A9
SLC6A1
SLC6A12
SLC6A15
SLC6A16
SLC6A19
SLC6A20
SLC6A3
SLC6A4
SLC6A5
SLC6A9
SLC7A13
SLC7A2
SLC7A5
SLC7A5P2
SLC7A7
SLC7A8
SLC8A1
SLC8A2
SLC8A3
SLC9A1
SLC9A11
SLC9A2
SLC9A3
SLC9A3R1
SLC9A3R2
SLC9A5
SLC9A6
SLC9A7
SLC9A8
SLC9A9
SLCO1A2
SLCO1B3
SLCO2A1
SLCO3A1
SLCO4A1
SLCO5A1
SLCO6A1
SLFN11
SLFN12L
SLFN13
SLFN14
SLIT1
SLIT2
SLIT3
SLITRK1
SLITRK3
SLITRK4
SLITRK6
SLMO1
SLTM
SMAD3
SMAD4
SMAD5
SMAD6
SMAD7
SMAD9
SMAGP
SMAP2
SMARCA1
SMARCA2
SMARCA4
SMARCAL1
SMARCC1
SMARCD3
SMARCE1
SMC2
SMC3
SMC5
SMCHD1
SMCR5
SMCR7L
SMEK1
SMEK2
SMEK3P
SMG6
SMG7
SMOC1
SMOC2
SMOX
SMPD1
SMPD3
SMPDL3B
SMPX
SMR3A
SMTN
SMTNL2
SMYD1
SMYD3
SMYD4
SNAI1
SNAP25
SNAP29
SNAP47
SNAP91
SNAPC1
SNAPC2
SNAPC3
SNAPC5
SNAPIN
SNAR-D
SND1
SNED1
SNHG3-RCC1
SNHG4
SNHG5
SNHG6
SNHG8
SNORA21
SNORA26
SNORA27
SNORA30
SNORA38
SNORA39
SNORA69
SNORD105
SNORD113-2
SNORD113-6
SNORD114-1
SNORD114-25
SNORD114-30
SNORD115-11
SNORD115-13
SNORD115-14
SNORD115-15
SNORD115-35
SNORD115-4
SNORD116-15
SNORD116-4
SNORD119
SNORD12
SNORD18A
SNORD2
SNORD24
SNORD27
SNORD30
SNORD43
SNORD45C
SNORD56B
SNORD63
SNORD87
SNORD89
SNORD93
SNPH
SNRNP200
SNRNP35
SNRNP40
SNRNP70
SNRPB
SNRPD1
SNRPD2
SNRPD3
SNRPF
SNRPN
SNTB2
SNTG1
SNTG2
SNUPN
SNX14
SNX15
SNX16
SNX18
SNX19
SNX2
SNX22
SNX24
SNX26
SNX27
SNX29
SNX31
SNX32
SNX33
SNX4
SNX6
SNX8
SNX9
SOAT1
SOAT2
SOBP
SOCS1
SOCS2
SOCS3
SOCS4
SOCS5
SOD3
SOHLH2
SOLH
SON
SORBS1
SORBS2
SORBS3
SORCS2
SORCS3
SORL1
SORT1
SOS1
SOS2
SOX1
SOX10
SOX12
SOX13
SOX14
SOX15
SOX17
SOX2OT
SOX3
SOX4
SOX5
SOX6
SP110
SP140
SP2
SP6
SP9
SPAG1
SPAG11A
SPAG11B
SPAG16
SPAG4
SPAG4L
SPAG6
SPAG7
SPANXN5
SPARC
SPAST
SPATA1
SPATA13
SPATA18
SPATA19
SPATA20
SPATA3
SPATA5L1
SPATC1
SPATS1
SPATS2
SPATS2L
SPC24
SPCS2
SPDEF
SPDYE4
SPEG
SPEM1
SPEN
SPG21
SPG7
SPHK1
SPI1
SPIN1
SPINK1
SPINK2
SPINK5L3
SPINK6
SPINK8
SPINT1
SPINT2
SPIRE2
SPN
SPNS1
SPNS2
SPNS3
SPO11
SPOCK1
SPOCK2
SPON2
SPP2
SPPL2A
SPPL2B
SPPL3
SPRED1
SPRED2
SPRED3
SPRR2B
SPRR2C
SPRR4
SPRY1
SPRY4
SPRYD3
SPRYD4
SPSB1
SPSB2
SPSB3
SPSB4
SPTA1
SPTB
SPTBN1
SPTBN2
SPTBN4
SPTBN5
SPTLC1
SPTLC2
SPTY2D1
SQLE
SQRDL
SR140
SRA1
SRBD1
SRC
SRCAP
SRCIN1
SRCRB4D
SRD5A1
SRD5A1P1
SRD5A2
SREBF1
SRF
SRFBP1
SRGAP1
SRGAP3
SRI
SRL
SRM
SRMS
SRP14
SRPK1
SRPK2
SRPX
SRPX2
SRRM1
SRRM3
SRRM4
SRRT
SS18L1
SS18L2
SSBP1
SSBP3
SSH1
SSH2
SSH3
SSPN
SSPO
SSR1
SSR2
SST
SSTR2
SSTR3
SSU72
SSX3
SSX7
ST14
ST18
ST3GAL1
ST3GAL2
ST3GAL3
ST3GAL5
ST3GAL6
ST5
ST6GAL1
ST6GALNAC2
ST6GALNAC4
ST7OT3
ST8SIA2
ST8SIA5

ST8SIA6
STAB1
STAC2
STAG1
STAG2
STAG3L2
STAM2
STAMBP
STAP1
STAR
STARD10
STARD3
STARD4
STARD7
STARD9
STAT1
STAT2
STAT4
STAT5A
STATH
STAU1
STC2
STEAP1
STEAP2
STEAP3
STEAP4
STIL
STIM1
STIM2
STK11
STK16
STK19
STK25
STK3
STK31
STK32A
STK32B
STK32C
STK35
STK38
STK39
STK4
STK40
STMN1
STMN2
STMN3
STOML1
STOML2
STON1-GTF2A1L
STON2
STOX1
STOX2
STRA13
STRA6
STRA8
STRAP
STRN
STRN3
STRN4
STT3A
STT3B
STX11
STX12
STX18
STX1A
STX2
STX5
STXBP2
STXBP4
STXBP5L
STXBP6
STYK1
STYXL1
SUB1
SUCLG1
SUCLG2
SUFU
SUGT1P1
SULF1
SULF2
SULT1A3
SULT1C4
SUMF1
SUMF2
SUMO2
SUMO3
SUNC1
SUOX
SUPT3H
SUPT6H
SUSD1
SUSD3
SUV39H1
SUV420H1
SUV420H2
SUZ12
SV2B
SV2C
SVIL
SVOP
SVOPL
SYCE1
SYCE1L
SYCE2
SYCN
SYCP3
SYF2
SYN1
SYN2
SYN3
SYNE1
SYNE2
SYNGAP1
SYNGR1
SYNGR3
SYNJ1
SYNJ2
SYNPO
SYNPO2L
SYNPR
SYNRG
SYP
SYPL1
SYPL2
SYS1
SYT1
SYT10
SYT11
SYT13
SYT14
SYT15
SYT17
SYT2
SYT3
SYT6
SYT7
SYT8
SYT9
SYTL1
SYTL2
SYTL3
SYTL4
SYVN1
T
TAAR3
TAAR9
TAC1
TACC1
TACC2
TACC3
TACR1
TACR3
TACSTD2
TADA1
TADA2A
TADA2B
TADA3
TAF10
TAF12
TAF15
TAF1B
TAF1C
TAF2
TAF4
TAF4B
TAF5
TAF7
TAF7L
TAF8
TAF9
TAF9B
TAGAP
TAGLN
TAL1
TANC1
TANC2
TANK
TAOK1
TAOK2
TAOK3
TAP1
TAP2
TAPBP
TARBP2
TARS2
TAS1R1
TAS1R2
TAS1R3
TAS2R38
TAS2R8
TASP1
TATDN1
TATDN3
TBC1D1
TBC1D10B
TBC1D10C
TBC1D12
TBC1D13
TBC1D14
TBC1D15
TBC1D16
TBC1D2
TBC1D21
TBC1D22A
TBC1D22B
TBC1D23
TBC1D24
TBC1D26
TBC1D28
TBC1D2B
TBC1D4
TBC1D8
TBC1D9
TBC1D9B
TBCB
TBCC
TBCCD1
TBCD
TBCK
TBKBP1
TBL1X
TBL1XR1
TBL2
TBP
TBPL1
TBR1
TBRG4
TBX1
TBX10
TBX15
TBX18
TBX2
TBX20
TBX21
TBX3
TBX4
TBX5
TBXA2R
TBXAS1
TC2N
TCAP
TCEA1
TCEA3
TCEAL2
TCEAL5
TCEAL6
TCEB3
TCEB3B
TCERG1
TCERG1L
TCF12
TCF15
TCF19
TCF21
TCF23
TCF25
TCF3
TCF7
TCF7L1
TCF7L2
TCFL5
TCIRG1
TCOF1
TCP1
TCP10L2
TCP11L1
TCP11L2
TCTE1
TCTE3
TCTEX1D4
TCTN1
TDG
TDH
TDRD10
TDRD12
TDRD5
TDRD6
TDRD9
TEAD1
TEAD3
TEC
TECPR1
TECPR2
TECR
TEF
TEK
TEKT1
TEKT2
TEKT3
TEKT4
TELO2
TENC1
TEPP
TERC
TERF2
TERT
TES
TESC
TESK1
TESK2
TESSP1
TET1
TEX101
TEX13B
TEX14
TEX15
TEX19
TEX2
TEX261
TEX264
TEX9
TF
TFAM
TFAP2A
TFAP2B
TFAP2D
TFAP2E
TFAP4
TFCP2L1
TFDP1
TFDP3
TFE3
TFEB
TFF1
TFF2
TFG
TFIP11
TFPI2
TFPT
TFR2
TG
TGDS
TGFA
TGFB2
TGFBR2
TGFBR3
TGFBRAP1
TGIF1
TGM2
TGM5
TGM6
TGM7
TGOLN2
TGS1
TH
TH1L
THADA
THAP11
THAP3
THAP4
THAP5
THAP7
THAP9
THBD
THBS1
THBS2
THBS4
THEM5
THNSL1
THNSL2
THOC1
THOC4
THOC5
THOP1
THRA
THRAP3
THRB
THSD1
THSD1P
THSD4
THTPA
THUMPD1
THUMPD2
THY1
THYN1
TIAL1
TIAM1
TIAM2
TICAM1
TIE1
TIFAB
TIGD3
TIGIT
TIMELESS
TIMM17A
TIMM17B
TIMM44
TIMM50
TIMM8A
TIMM8B
TIMM9
TIMP2
TIMP4
TINAGL1
TINF2
TIPARP
TIPIN
TIPRL
TIRAP
TJP1
TJP2
TK1
TK2
TKTL2
TLCD2
TLE1
TLE2
TLE3
TLE4
TLL1
TLL2
TLR1
TLR2
TLR4
TLR5
TLR6
TLR9
TLX1
TLX2
TLX3
TM2D1
TM2D2
TM4SF5
TMBIM4
TMC1
TMC6
TMC8
TMCC1
TMCC3
TMCO1
TMCO3
TMCO6
TMCO7
TMED1
TMED10
TMED2
TMED4
TMED6
TMED7-TICAM2
TMEM100
TMEM101
TMEM102
TMEM106A
TMEM108
TMEM109
TMEM110
TMEM114
TMEM115
TMEM116
TMEM117
TMEM120A
TMEM120B
TMEM121
TMEM123
TMEM127
TMEM129
TMEM130
TMEM132A
TMEM132B
TMEM132C
TMEM132D
TMEM132E
TMEM134
TMEM135
TMEM138
TMEM139
TMEM140
TMEM146
TMEM14A
TMEM14C
TMEM150A
TMEM151A
TMEM151B
TMEM155
TMEM156
TMEM159
TMEM161A
TMEM163
TMEM165
TMEM167B
TMEM168
TMEM170B
TMEM171
TMEM173
TMEM174
TMEM175
TMEM178
TMEM179
TMEM18
TMEM181
TMEM183A
TMEM183B
TMEM185A
TMEM187
TMEM188
TMEM19
TMEM191A
TMEM192
TMEM194A
TMEM194B
TMEM195
TMEM196
TMEM2
TMEM200A
TMEM200B
TMEM200C
TMEM201
TMEM211
TMEM212
TMEM216
TMEM219
TMEM22
TMEM222
TMEM229B
TMEM231
TMEM233
TMEM25
TMEM39A
TMEM40
TMEM45A
TMEM48
TMEM49
TMEM5
TMEM51
TMEM53
TMEM57
TMEM59L
TMEM62
TMEM63A
TMEM63B
TMEM64
TMEM71
TMEM79
TMEM82
TMEM85
TMEM86B
TMEM88
TMEM89
TMEM8A
TMEM91
TMEM92
TMEM98
TMEM9B
TMF1
TMIGD2
TMLHE
TMOD1
TMOD3
TMOD4
TMPO
TMPRSS11D
TMPRSS11F
TMPRSS13
TMPRSS3
TMPRSS4
TMPRSS6
TMPRSS9
TMSB15A
TMSB15B
TMTC1
TMTC4
TMUB1
TMUB2
TMX1
TMX2
TMX4
TNC
TNF
TNFAIP1
TNFAIP3
TNFAIP8
TNFAIP8L2
TNFAIP8L3
TNFRSF10B
TNFRSF10D
TNFRSF11B
TNFRSF13C
TNFRSF18
TNFRSF1A
TNFRSF1B
TNFRSF25
TNFRSF8
TNFSF10
TNFSF13
TNFSF13B
TNFSF18
TNIK
TNIP1
TNIP2
TNK1
TNK2
TNKS
TNKS1BP1
TNKS2
TNN
TNNC2
TNNT3
TNP1
TNPO1
TNPO2
TNPO3
TNR
TNRC18
TNRC4
TNRC6B
TNRC6C
TNS1
TNS3
TNS4
TNXB
TOB1
TOLLIP
TOM1L2
TOMM20
TOMM20L
TOMM22
TOMM34
TOMM40
TOMM7
TOP1
TOP1MT
TOP2A
TOP3A
TOP3B
TOPBP1
TOPORS
TOR1AIP1
TOR1AIP2
TOR1B
TOR3A
TOX
TOX2
TP53
TP53AIP1
TP53BP1
TP53I11
TP53I13
TP53I3
TP53INP2
TP53RK
TP53TG5
TP63
TP73
TPBG
TPCN2
TPD52
TPD52L2
TPH2
TPK1
TPM3
TPM4
TPO
TPP1
TPPP
TPPP2
TPPP3
TPRA1
TPRKB
TPRXL
TPST1
TPTE
TPX2
TRADD
TRAF1
TRAF3
TRAF3IP1
TRAF3IP2
TRAF5

TRAF7
TRAIP
TRAK1
TRAM1L1
TRAM2
TRANK1
TRAP1
TRAPPC2
TRAPPC2L
TRAPPC3
TRAPPC4
TRAPPC5
TRAPPC9
TRDMT1
TREH
TREM2
TREML1
TREML2
TREML2P
TRERF1
TRH
TRHDE
TRHR
TRIB1
TRIB2
TRIB3
TRIL
TRIM10
TRIM13
TRIM15
TRIM17
TRIM2
TRIM24
TRIM25
TRIM26
TRIM27
TRIM28
TRIM29
TRIM3
TRIM31
TRIM33
TRIM35
TRIM36
TRIM37
TRIM38
TRIM39
TRIM40
TRIM41
TRIM42
TRIM44
TRIM46
TRIM47
TRIM50
TRIM52
TRIM53
TRIM54
TRIM56
TRIM59
TRIM62
TRIM63
TRIM65
TRIM66
TRIM6-
TRIM34
TRIM7
TRIM71
TRIM72
TRIM75
TRIM8
TRIM9
TRIML1
TRIO
TRIOBP
TRIP10
TRIP11
TRIP13
TRIP4
TRMT11
TRMT2B
TRMT61B
TRMU
TRNAU1AP
TRNP1
TROAP
TRPA1
TRPC3
TRPC4
TRPC4AP
TRPC7
TRPM2
TRPM4
TRPM5
TRPM8
TRPS1
TRPV3
TRPV4
TRRAP
TSC2
TSC22D1
TSC22D3
TSC22D4
TSEN2
TSEN54
TSG1
TSGA10IP
TSGA14
TSHR
TSHZ2
TSKS
TSKU
TSLP
TSNARE1
TSNAX-
DISC1
TSPAN1
TSPAN12
TSPAN13
TSPAN14
TSPAN17
TSPAN2
TSPAN31
TSPAN33
TSPAN4
TSPAN5
TSPAN6
TSPAN7
TSPAN9
TSPO
TSPYL3
TSPYL4
TSPYL5
TSPYL6
TSR2
TSSC1
TSSC4
TSSK6
TST
TSTA3
TSTD1
TTBK1
TTBK2
TTC12
TTC13
TTC15
TTC21B
TTC23
TTC25
TTC26
TTC28
TTC29
TTC3
TTC30A
TTC30B
TTC33
TTC35
TTC36
TTC38
TTC39A
TTC39C
TTC4
TTC7A
TTC7B
TTC8
TTF2
TTK
TTLL1
TTLL10
TTLL12
TTLL3
TTLL5
TTLL7
TTLL8
TTN
TTR
TTTY18
TTYH1
TTYH2
TTYH3
TUB
TUBA1B
TUBA1C
TUBA3D
TUBA3E
TUBA8
TUBB
TUBB2A
TUBB2C
TUBB3
TUBB4Q
TUBB6
TUBB8
TUBBP5
TUBE1
TUBGCP2
TUBGCP5
TUBGCP6
TUFT1
TUG1
TULP2
TULP3
TUSC3
TUSC4
TWF1
TWF2
TWISTNB
TXLNA
TXN
TXN2
TXNDC11
TXNDC17
TXNIP
TXNL1
TXNL4A
TXNRD1
TXNRD2
TYK2
TYMS
TYRO3
TYROBP
TYRP1
TYSND1
U2AF1
U2AF2
UACA
UAP1L1
UBA1
UBA3
UBA7
UBAC1
UBAC2
UBAP1
UBAP2L
UBASH3B
UBE2B
UBE2C
UBE2D1
UBE2E1
UBE2E2
UBE2E3
UBE2F
UBE2G1
UBE2G2
UBE2I
UBE2L3
UBE2M
UBE2MP1
UBE2O
UBE2Q2
UBE2R2
UBE2U
UBE2V1
UBE2Z
UBE3B
UBE3C
UBE4B
UBFD1
UBIAD1
UBL4B
UBN1
UBQLN2
UBQLNL
UBR1
UBR2
UBR5
UBTD1
UBTF
UBXN1
UBXN10
UBXN11
UBXN4
UCHL3
UCK2
UCN3
UCP1
UCRC
UEVLD
UGCG
UGDH
UGP2
UGT1A10
UGT2B15
UGT3A1
UHMK1
UHRF1
UHRF1BP1L
ULBP1
ULBP2
ULK1
ULK2
ULK4
UMODL1
UNC13A
UNC13B
UNC13D
UNC45B
UNC5A
UNC5B
UNC5CL
UNC5D
UNC80
UNC84A
UNC93A
UNC93B1
UNCX
UNG
UNK
UNKL
UPF1
UPF3A
UPF3B
UPK1A
UPK1B
UPK3A
UPP2
UQCC
UQCRC2
UQCRFS1
UQCRHL
URB1
URB2
URGCP
UROD
USF1
USF2
USO1
USP1
USP10
USP12
USP13
USP14
USP16
USP2
USP20
USP21
USP22
USP25
USP3
USP31
USP32
USP33
USP34
USP35
USP36
USP4
USP40
USP42
USP43
USP44
USP46
USP49
USP5
USP51
USP53
USP54
USP6NL
USP7
UST
UTP14A
UTP14C
UTP6
UTRN
UXS1
UXT
VAC14
VAMP2
VAMP3
VAMP4
VANGL1
VANGL2
VAPA
VARS
VARS2
VASH1
VASH2
VASP
VAT1
VAV1
VAV2
VAV3
VAX1
VAX2
VCAN
VCX2
VCX3B
VDAC1
VDAC3
VEGFA
VENTX
VENTXP1
VENTXP7
VEPH1
VGF
VGLL2
VGLL3
VGLL4
VIL1
VILL
VIP
VIPR1
VIPR2
VKORC1
VMO1
VN1R4
VOPP1
VPRBP
VPS13A
VPS13B
VPS13D
VPS16
VPS24
VPS29
VPS36
VPS37A
VPS37B
VPS37C
VPS39
VPS41
VPS45
VPS4A
VPS52
VPS53
VPS54
VPS72
VSIG8
VSNL1
VSTM2L
VSX2
VTCN1
VTI1A
VTI1B
VTN
VTRNA1-1
VTRNA1-3
VWA1
VWA2
VWA3A
VWA3B
VWA5B2
VWC2L
VWF
WAC
WASF1
WASF3
WBP2
WBP2NL
WBP4
WBSCR16
WBSCR17
WBSCR22
WDFY1
WDFY2
WDFY3
WDFY4
WDR1
WDR12
WDR13
WDR17
WDR20
WDR24
WDR25
WDR26
WDR27
WDR31
WDR33
WDR35
WDR37
WDR45
WDR45L
WDR46
WDR47
WDR52
WDR53
WDR5B
WDR60
WDR64
WDR65
WDR67
WDR69
WDR70
WDR73
WDR74
WDR8
WDR81
WDR86
WDR89
WDR93
WDTC1
WEE1
WFDC1
WFDC3
WFDC5
WFDC9
WFIKKN2
WFS1
WHAMM
WHAMML1
WHSC1L1
WHSC2
WIBG
WIPF3
WIPI1
WIPI2
WIZ
WNK1
WNK2
WNK3
WNT1
WNT10A
WNT10B
WNT11
WNT16
WNT2
WNT2B
WNT3
WNT4
WNT5A
WNT5B
WNT6
WNT7A
WNT7B
WNT8B
WNT9A
WNT9B
WRB
WRN
WRNIP1
WSB2
WSCD1
WSCD2
WT1
WTAP
WWC1
WWC2
WWOX
WWP1
WWP2
XAB2
XAF1
XAGE3
XIRP1
XIRP2
XK
XKR4
XKR5
XKR6
XKR7
XKR8
XKR9
XKRX
XPC
XPNPEP1
XPNPEP3
XPO6
XPOT
XRCC2
XRCC4
XRCC5
XRCC6
XRRA1
XYLT1
YAF2
YARS2
YBX2
YIF1A
YIPF1
YIPF3
YIPF4
YIPF5
YIPF7
YJEFN3
YOD1
YPEL1
YPEL2
YPEL4
YPEL5
YSK4
YTHDC1
YTHDF3
YWHAB
YWHAE
YWHAG
YWHAQ
YWHAZ
ZACN
ZADH2
ZAK
ZAR1
ZBED2
ZBED4
ZBED5
ZBTB11
ZBTB12
ZBTB16
ZBTB17
ZBTB2
ZBTB20
ZBTB22
ZBTB24
ZBTB25
ZBTB3
ZBTB38
ZBTB4
ZBTB40
ZBTB44
ZBTB46
ZBTB47
ZBTB48
ZBTB5
ZBTB7A
ZBTB7B
ZBTB8B
ZBTB9
ZC3H11A
ZC3H12A
ZC3H12C
ZC3H12D
ZC3H14
ZC3H15
ZC3H18
ZC3H3
ZC3H4
ZC3HAV1L
ZC3HC1
ZC4H2
ZCCHC12
ZCCHC14
ZCCHC24
ZCCHC6
ZCCHC8
ZCWPW1
ZDHHC1
ZDHHC11
ZDHHC14
ZDHHC15
ZDHHC16
ZDHHC17
ZDHHC19
ZDHHC2
ZDHHC21
ZDHHC3
ZDHHC4
ZDHHC5
ZDHHC7
ZDHHC8
ZDHHC9
ZEB2
ZFAND2A
ZFC3H1
ZFHX3
ZFHX4
ZFP1
ZFP3
ZFP30
ZFP36
ZFP36L1
ZFP41
ZFP62
ZFP64
ZFP91
ZFP92
ZFPM1
ZFPM2
ZFR
ZFYVE1
ZFYVE16
ZFYVE20
ZFYVE21
ZFYVE27
ZFYVE28
ZFYVE9

ZGLP1
ZGPAT
ZHX2
ZHX3
ZIC1
ZIC3
ZIC4
ZIC5
ZKSCAN1
ZKSCAN5
ZMIZ1
ZMIZ2
ZMPSTE24
ZMYM1
ZMYM2
ZMYM4
ZMYND11
ZMYND19
ZMYND8
ZNF10
ZNF117
ZNF12
ZNF121
ZNF136
ZNF138
ZNF14
ZNF143
ZNF146
ZNF148
ZNF16
ZNF167
ZNF169
ZNF174
ZNF181
ZNF182
ZNF185
ZNF187
ZNF192
ZNF193
ZNF2
ZNF20
ZNF200
ZNF204P
ZNF205
ZNF211
ZNF212
ZNF214
ZNF215
ZNF219
ZNF22
ZNF226
ZNF229
ZNF23
ZNF230
ZNF232
ZNF235
ZNF236
ZNF238
ZNF239
ZNF24
ZNF248
ZNF25
ZNF250
ZNF251
ZNF252
ZNF256
ZNF26
ZNF260
ZNF263
ZNF266
ZNF267
ZNF268
ZNF274
ZNF276
ZNF280B
ZNF280C
ZNF281
ZNF282
ZNF283
ZNF286B
ZNF295
ZNF296
ZNF3
ZNF311
ZNF317
ZNF318
ZNF322A
ZNF330
ZNF331
ZNF333
ZNF334
ZNF337
ZNF341
ZNF345
ZNF347
ZNF35
ZNF354C
ZNF362
ZNF382
ZNF385B
ZNF395
ZNF397
ZNF398
ZNF404
ZNF415
ZNF416
ZNF420
ZNF423
ZNF425
ZNF426
ZNF429
ZNF430
ZNF431
ZNF438
ZNF446
ZNF449
ZNF451
ZNF460
ZNF462
ZNF467
ZNF469
ZNF48
ZNF484
ZNF485
ZNF487
ZNF488
ZNF496
ZNF497
ZNF498
ZNF500
ZNF506
ZNF513
ZNF516
ZNF518B
ZNF524
ZNF525
ZNF527
ZNF528
ZNF530
ZNF532
ZNF536
ZNF540
ZNF542
ZNF543
ZNF547
ZNF551
ZNF552
ZNF566
ZNF568
ZNF57
ZNF574
ZNF575
ZNF576
ZNF578
ZNF580
ZNF583
ZNF584
ZNF585A
ZNF586
ZNF589
ZNF593
ZNF595
ZNF596
ZNF597
ZNF598
ZNF599
ZNF606
ZNF607
ZNF608
ZNF610
ZNF611
ZNF613
ZNF614
ZNF618
ZNF619
ZNF620
ZNF621
ZNF622
ZNF624
ZNF625
ZNF628
ZNF629
ZNF630
ZNF639
ZNF641
ZNF643
ZNF644
ZNF646
ZNF652
ZNF660
ZNF662
ZNF665
ZNF671
ZNF672
ZNF674
ZNF678
ZNF679
ZNF681
ZNF687
ZNF688
ZNF692
ZNF696
ZNF697
ZNF7
ZNF702P
ZNF703
ZNF704
ZNF709
ZNF710
ZNF711
ZNF713
ZNF714
ZNF718
ZNF721
ZNF737
ZNF740
ZNF749
ZNF75D
ZNF76
ZNF763
ZNF764
ZNF768
ZNF77
ZNF770
ZNF771
ZNF774
ZNF775
ZNF777
ZNF781
ZNF783
ZNF784
ZNF786
ZNF787
ZNF788
ZNF793
ZNF8
ZNF800
ZNF805
ZNF808
ZNF813
ZNF815
ZNF816A
ZNF823
ZNF827
ZNF828
ZNF829
ZNF83
ZNF830
ZNF831
ZNF835
ZNF836
ZNF837
ZNF843
ZNF844
ZNF845
ZNF85
ZNF860
ZNF862
ZNF879
ZNFX1
ZNHIT2
ZNHIT3
ZNHIT6
ZNRD1
ZNRF1
ZNRF2
ZNRF3
ZNRF4
ZP2
ZP3
ZPBP
ZPBP2
ZPLD1
ZSCAN10
ZSCAN12L1
ZSCAN20
ZSCAN21
ZSCAN22
ZSCAN23
ZSCAN4
ZSCAN5A
ZSCAN5B
ZSWIM2
ZSWIM4
ZSWIM5
ZSWIM6
ZSWIM7
ZW10
ZWILCH
ZXDC
ZYG11A
ZZEF1
ZZZ3

Supplementary Table 2. Top DNA methylation sites associated with depressive symptoms in the discovery EWAS: M2 is model additionally adjusted for antidepressant medication

CpG site ID	Chr	Location	Gene symbol	Direction M1$	*p* value M1	Direction M2$	*p* value M2
cg04987734	14	103415873	*CDC42BPB*	++--+-+++	4.93×10^{-8}	++--+-+++	3.94×10^{-8}
cg07012687	17	80195180	*SLC16A3*	+++-+++++	3.47×10^{-7}	+++-+++++	4.51×10^{-7}
cg08796240	16	70733832	*VAC14*	+++++++++	7.43×10^{-7}	++-++++++	4.06×10^{-6}
cg06096336	2	231989800	*PSMD1;HTR2B*	++-++-+++	8.06×10^{-7}	+++++-+++	2.11×10^{-6}
cg16745930	10	100220809	*HPSE2*	---------	1.34×10^{-6}	---------	2.19×10^{-6}
cg09849319	5	1494983	*LPCATI1*	+++-+-+++	1.81×10^{-6}	+++-+-+++	6.14×10^{-6}
cg17237086	22	40814966	*MKL1*	+++++-+++	3.44×10^{-6}	++-++-+++	1.30×10^{-5}
cg03985718	2	105924245	*TGFBRAP1*	+++-+-+++	3.61×10^{-6}	+++-+-+++	1.28×10^{-5}
cg21098005	20	44538605	*PLTP*	+++++++++	4.36×10^{-6}	+++++++++	7.86×10^{-6}
cg16466652	19	6271960	*MLLT1*	+++-+++++	4.39×10^{-6}	+++-+++++	7.08×10^{-6}
cg07884764	11	64107517	*CCDC88B*	+++-+++++	5.03×10^{-6}	+++-+++++	1.50×10^{-5}
cg01541347	7	4729920	*FOXK1*	++++++++-	5.64×10^{-6}	++++++++-	6.66×10^{-6}
cg02341197	21	34185927	*C21or62*	+++++-+++	5.84×10^{-6}	+++++++++	8.57×10^{-6}
cg01947751	3	196728969	-	+++-+++++	6.23×10^{-6}	+++-+++++	5.93×10^{-6}
cg13747876	17	80195402	*SLC16A3*	+++++-+++	6.32×10^{-6}	+++++++++	1.31×10^{-5}
cg12764201	1	105101123	*CORT;APITD1*	+++++-+++	7.15×10^{-6}	+++++-+++	1.33×10^{-5}
cg08295111	5	133866097	*PHF15*	+++-+++++	7.87×10^{-6}	+++-+++++	1.73×10^{-5}
cg18030453	3	45506216	*LARS2*	+++-+-+++	9.16×10^{-6}	+++-+-+++	2.44×10^{-5}
cg12325605	3	56810151	*ARHGEF3*	+++++++++	9.62×10^{-6}	++-++++++	1.82×10^{-5}
cg23282441	10	73533927	*C10orf54;CDH23*	++--+++++	9.69×10^{-6}	++-++++++	5.30×10^{-6}

$ Order of cohorts in the direction column: Discovery meta-analysis: CHS, FS, GS, HBCS, KORA, LBC1921, LBC1936, RS_BIOS, RS_III; Replication meta-analysis: WHI-EMPC, ARIC; Discovery and replication samples meta-analysis: CHS, FS, GS, HBCS, KORA, LBC1921, LBC1936, RS_BIOS, RS_III, WHI-EMPC, and ARIC; + hyper-methylation and – hypo-methylation.

Supplementary Table 3. Top DNA methylation sites associated with depressive symptoms in the meta-EWAS (*N* = 11 256)

MarkerName	Weight	Zscore	*P* value	Direction$	HetPVal	CHR	MAPINFO	UCSC_RefGene
cg12325605	11256	5.839	5.24×10-09	++++++++++	0.903	3	56810151	*ARHGEF3*
cg04987734	11256	5.653	1.57×10-08	++--+-++++-	0.004874	14	103415873	*CDC42BPB*
cg14023999	11256	5.419	5.99×10-08	+++++++++-++	0.1514	15	90543224	
cg18030453	11253	5.293	1.20×10-07	+++-+-++++-	0.4738	3	45506216	*LARS2*
cg10829227	11254	5.1	3.40×10-07	+++++-+++++	0.5651	19	47200595	*PRKD2*
cg07175797	11254	5.073	3.92×10-07	+++-+++-+++	0.389	16	50317656	
cg07012687	11251	5.048	4.46×10-07	+++-++++++-	0.2654	17	80195180	*SLC16A3*
cg12728588	11256	5.011	5.43×10-07	++++++++++	0.5798	1	36025489	*NCDN*
cg22069247	11253	4.93	8.23×10-07	+++-+-+++++	0.7434	2	232393256	*NMUR1*
cg12526091	11256	4.912	9.00×10-07	+++-+-+++++	0.8524	12	58245042	
cg14012686	11253	4.89	1.01×10-06	++-++++++-	0.6241	2	74785750	*C2orf65*
cg00153395	11255	4.888	1.02×10-06	+++-++++++	0.8356	1	65327523	*JAK1*
cg04583842	10491	4.836	1.33×10-06	+++-+++?+++	0.3046	16	88103117	*BANP*
cg03429645	11255	4.82	1.43×10-06	+++++++-+++	0.666	3	100053188	*NIT2*
cg04286697	11256	4.78	1.75×10-06	+++-++++++	0.2851	2	232259623	*B3GNT7*
cg08796240	11256	4.775	1.80×10-06	++++++++++-	0.2479	16	70733832	*VAC14*
cg19769147	11255	4.744	2.09×10-06	+++++-++++-	0.1715	14	105860954	*PACS2*
cg17822325	11253	4.722	2.33×10-06	+++++-+++++	0.9251	1	31896462	*SERINC2*
cg06096336	11254	4.707	2.51×10-06	++-++-+++++	0.4324	2	231989800	*PSMD1;HTR2B*
cg10401362	11254	4.691	2.71×10-06	-++-+-+++++	0.001998	7	157185402	*DNAJB6*
cg07467649	11256	4.689	2.75×10-06	+++-+-+++++	0.3612	7	72856462	*BAZ1B*
cg13747876	11256	4.676	2.93×10-06	+++++-++++-	0.4857	17	80195402	*SLC16A3*
cg25392060	11256	4.664	3.10×10-06	+++-++++++	0.5181	8	142297121	
cg22365240	11250	4.659	3.17×10-06	-++++++++++	0.983	2	105374995	
cg08631783	11253	4.62	3.83×10-06	-++-+-+-+++	0.1325	15	89418456	*ACAN*
cg19269039	11255	4.62	3.84×10-06	+++++++++-++	0.5069	1	111743200	*DENND2D*
cg03720762	11256	4.572	4.82×10-06	+++++++-+++	0.08729	17	17604184	*RAI1*
cg21604136	11255	4.568	4.93×10-06	++-+++++++	0.1963	1	9910137	*CTNNBIP1*
cg11931558	11252	4.546	5.46×10-06	++++++++++	0.8237	7	11013742	*PHF14*
cg24550880	11255	4.531	5.88×10-06	+++-+-+++++	0.4632	17	79420279	*BAHCC1*
cg07372520	11255	4.526	6.01×10-06	+++-+-+++++	0.4239	1	180086434	
cg05827190	11255	4.524	6.08×10-06	+++++-++-++	0.2259	4	681440	*MFSD7*
cg19743103	11253	4.524	6.08×10-06	+++--++++++	0.1436	1	6304220	*HES3*
cg17237086	11256	4.523	6.09×10-06	+++++-++++-	0.3516	22	40814966	*MKL1*
cg16745930	11254	-4.518	6.26×10-06	----------	0.1094	10	100220809	*HPSE2*
cg27305772	11256	4.507	6.57×10-06	+++++-+++++	0.8274	11	65630355	*MUS81*
cg02341197	11256	4.5	6.80×10-06	+++++-+++++	0.8595	21	34185927	*C21orf62*
cg23606718	11255	4.456	8.37×10-06	+++-++++++	0.8073	2	131513927	*FAM123C*
cg23282441	11254	4.449	8.63×10-06	++--++++++-	0.4265	10	73533927	*C10orf54;CDH23*
cg17743381	11256	4.437	9.10×10-06	+++++-+++++	0.7012	1	39024825	
cg05251389	11254	4.424	9.67×10-06	+++-+++-+++	0.7075	22	43525330	*BIK*
cg26610247	11255	4.424	9.70E-06	++---++++++	0.1656	8	142297175	

$ Order of cohorts in the direction column: Discovery and replication samples meta-analysis: CHS, FS, GS, HBCS, KORA, LBC1921, LBC1936, RS_BIOS, RS_III, WHI-EMPC, and ARIC; + hyper-methylation and – hypo-methylation.

Supplementary Table 4. Results for pathway analysis. Total pathways tested 1 775. Bonferroni threshold for significance = 0.05/1775 = $2.82x10^{-05}$.

Reactome pathways	Homo sapiens - REFLIST (20972)	Observed	Expected	Fold Enrichment	*p* value
Potassium Channels (R-HSA-1296071)	99	75	44.9	1.67	4.26×10-02
Downstream signal transduction (R-HSA-186763)	327	218	148.3	1.47	6.99×10-05
Signaling by PDGF (R-HSA-186797)	**354**	**236**	**160.54**	**1.47**	**1.98×10-05**
Signaling by SCF-KIT (R-HSA-1433557)	311	206	141.04	1.46	2.52×10-04
Muscle contraction (R-HSA-397014)	202	133	91.61	1.45	4.74×10-02
DAP12 signaling (R-HSA-2424491)	330	217	149.66	1.45	2.00×10-04
Downstream signaling of activated FGFR4 (R-HSA-5654716)	315	207	142.86	1.45	4.03×10-04
Downstream signaling of activated FGFR3 (R-HSA-5654708)	315	207	142.86	1.45	4.03×10-04
Downstream signaling of activated FGFR2 (R-HSA-5654696)	315	207	142.86	1.45	4.03×10-04
NGF signaling via TRKA from the plasma membrane (R-HSA-187037)	361	237	163.72	1.45	6.09×10-05
Signaling by FGFR3 (R-HSA-5654741)	319	209	144.67	1.44	4.33×10-04
Downstream signaling of activated FGFR1 (R-HSA-5654687)	318	208	144.22	1.44	5.16×10-04
Signaling by FGFR4 (R-HSA-5654743)	318	208	144.22	1.44	5.16×10-04
Signaling by EGFR (R-HSA-177929)	336	219	152.38	1.44	3.27×10-04
Neuronal System (R-HSA-112316)	338	220	153.29	1.44	3.38×10-04
Signaling by FGFR1 (R-HSA-5654736)	323	210	146.48	1.43	6.81×10-04
Signaling by ERBB4 (R-HSA-1236394)	315	204	142.86	1.43	1.27×10-03
Signaling by NGF (R-HSA-166520)	**440**	**284**	**199.54**	**1.42**	**1.31×10-05**
Axon guidance (R-HSA-422475)	**546**	**350**	**247.62**	**1.41**	**5.20×10-07**
DAP12 interactions (R-HSA-2172127)	344	220	156.01	1.41	1.13×10-03
Signaling by FGFR2 (R-HSA-5654738)	347	221	157.37	1.4	1.41×10-03
Interleukin-3, 5 and GM-CSF signaling (R-HSA-512988)	250	159	113.38	1.4	4.81×10-02
VEGFA-VEGFR2 Pathway (R-HSA-4420097)	307	195	139.23	1.4	7.06×10-03
NCAM signaling for neurite out-growth (R-HSA-375165)	260	165	117.91	1.4	3.85×10-02
Signaling by FGFR (R-HSA-190236)	353	224	160.09	1.4	1.53×10-03
Signaling by VEGF (R-HSA-194138)	316	200	143.31	1.4	6.65×10-03
Extracellular matrix organization (R-HSA-1474244)	292	184	132.42	1.39	2.01×10-02
Gastrin-CREB signalling pathway via PKC and MAPK (R-HSA-881907)	421	258	190.93	1.35	3.17×10-03
Developmental Biology (R-HSA-1266738)	**806**	**481**	**365.53**	**1.32**	**3.95×10-06**
Hemostasis (R-HSA-109582)	590	346	267.57	1.29	3.12×10-03
Transmembrane transport of small molecules (R-HSA-382551)	656	374	297.5	1.26	1.39×10-02
Disease (R-HSA-1643685)	899	497	407.7	1.22	1.16×10-02
Immune System (R-HSA-168256)	1604	865	727.43	1.19	2.28×10-04
Metabolism (R-HSA-1430728)	1972	1034	894.32	1.16	1.50×10-03

Supplementary Table 5. Differentially expressed probes associated with methylation at cg04987734

probeID	Effect	SE	*p* value	*p* value FDR	Gene	Protein Product	Chr	Cytoband
ILMN_1731233	2198.656865	400.6655942	5.70x10-08	0.00256	GZMH	NP_219491.1	14	14q12a
ILMN_1678238	428.2186866	83.04433811	3.28x10-07	0.00368	ZNF683	NP_775845.2	1	1p36.11b
ILMN_1768482	1564.643903	302.4642458	3.01x10-07	0.00368	CD8A	NP_741969.1	2	2p11.2e
ILMN_2353732	1520.927767	294.7777577	3.23x10-07	0.00368	CD8A	NP_741969.1	2	2p11.2e
ILMN_1760374	226.9467354	47.69009953	2.37x10-06	0.02124	CD8A	NP_001759.3	2	2p11.2e
ILMN_1813338	131.6868524	28.62194534	4.99x10-06	0.03734	LAG3	NP_002277.3	12	12p13.31d

Supplementary Table 6. Differentially expressed probes associated with methylation at cg12325605

ProbeID	Effect	SD	*p* value	*p* value FDR	Gene	Protein_Product	Chr	Cytoband
ILMN_3307729	-8.8331	1.0783	1.22E-15	2.59E-11	CXXC5	NP_057547.5	5	5q31.3a
ILMN_1693826	-6.3099	0.8402	1.81E-13	1.92E-09	HAVCR2	NP_116171.3	5	5q33.3a
ILMN_1792389	-6.7942	0.9238	5.36E-13	3.80E-09	RNF165	NP_689683.2	18	18q21.1a
ILMN_1745256	-7.2891	1.0111	1.46E-12	7.77E-09	CXXC5	NP_057547.4	5	5q31.3a
ILMN_1756439	-5.0451	0.7212	6.19E-12	2.63E-08	SCRN1	NP_055581.2	7	7p15.1b
ILMN_1701237	-7.5309	1.1002	1.67E-11	5.91E-08	SH2D1B	NP_444512.2	1	1q23.3b
ILMN_1751572	-4.0370	0.6086	6.55E-11	1.99E-07	TLE1	NP_005068.2	9	9q21.31d-q21.32a
ILMN_1695025	5.4926	0.8334	8.59E-11	2.28E-07	CD2	NP_001758.1	1	1p13.1b
ILMN_1796423	-9.9822	1.5515	2.30E-10	5.43E-07	CLIC3	NP_004660.2	9	9q34.3e
ILMN_1779071	-5.8855	0.9446	8.01E-10	1.70E-06	FEZ1	NP_005094.1	11	11q24.2b
ILMN_1767934	-3.9321	0.6349	1.00E-09	1.93E-06	PCSK5	NP_006191.2	9	9q21.13c
ILMN_2329114	-4.3857	0.7155	1.47E-09	2.60E-06	COLQ	NP_536800.1		3p24.3e
ILMN_1778240	-6.4774	1.0674	2.11E-09	3.45E-06	GFOD1	NP_061861.1	6	6p24.1a-p23b
ILMN_1713124	-8.5259	1.4224	3.28E-09	4.83E-06	AKR1C3	NP_003730.4	10	10p15.1c
ILMN_2261416	6.9993	1.1691	3.41E-09	4.83E-06	CD3D	NP_000723.1	11	11q23.3d
ILMN_1662026	-3.7472	0.6389	6.93E-09	9.20E-06	BTK	NP_000052.1	X	Xq22.1c
ILMN_1717197	5.7097	0.9868	1.09E-08	1.36E-05	CD3G	NP_000064.1	11	11q23.3d
ILMN_1778321	-4.6121	0.8019	1.32E-08	1.56E-05	SLC2A6	NP_060055.1	9	9q34.2a
ILMN_1694268	-3.6794	0.6423	1.50E-08	1.68E-05	HES6	NP_061115.2	2	2q37.3c
ILMN_1666594	-5.0374	0.8876	2.03E-08	2.15E-05	IRF8	NP_002154.1	16	16q24.1b
ILMN_1725750	3.3173	0.5861	2.20E-08	2.16E-05	LOC644695	XP_937403.1	17	17q21.32c
ILMN_2116714	-4.3420	0.7675	2.24E-08	2.16E-05	SLC39A1	NP_055252.2	1	1q21.3d
ILMN_3238058	-4.1020	0.7301	2.78E-08	2.56E-05	LOC151162		2	
ILMN_1729987	-4.1893	0.7520	3.62E-08	3.20E-05	SRC	NP_938033.1	20	20q11.23b
ILMN_1814194	-5.2012	0.9416	4.68E-08	3.98E-05	TCF4	NP_003190.1	18	18q21.2d
ILMN_1789394	-2.9872	0.5424	5.12E-08	4.19E-05	CATSPER1	NP_444282.2	11	11q13.1d
ILMN_1686623	-6.3585	1.1654	6.76E-08	5.16E-05	CSF1R	NP_005202.2	5	5q33.1c
ILMN_2062468	-4.4479	0.8154	6.80E-08	5.16E-05	IGFBP7	NP_001544.1	4	4q12e
ILMN_1763000	-4.3631	0.8141	1.14E-07	8.10E-05	ADAP2	NP_060874.1	17	17q11.2c
ILMN_1778010	4.6008	0.8587	1.14E-07	8.10E-05	IL32	NP_001012654.1	16	16p13.3d
ILMN_2368530	6.6308	1.2405	1.22E-07	8.38E-05	IL32	NP_001012651.1	16	16p13.3d

R39 R38 R37 R36 R35 R34 R33 R32 R31 R30 R29 R28 R27 R26 R25 R24 R23 R22 R21 R20 R19 R18 R17 R16 R15 R14 R13 R12 R11 R10 R9 R8 R7 R6 R5 R4 R3 R2 R1

ProbeID	Effect	SD	*p* value	*p* value FDR	Gene	Protein_Product	Chr	Cytoband
ILMN_1742544	-4.7371	0.8875	1.28E-07	8.46E-05	MEF2C	NP_002388.2	5	5q14.3f
ILMN_1730917	-2.6764	0.5052	1.57E-07	0.000101092	KMO	NP_003670.1	1	1q43e
ILMN_1780861	-4.4612	0.8437	1.66E-07	0.000103557	LOC653506	XP_932862.1	17	
ILMN_2044453	-3.6438	0.6942	2.03E-07	0.000123126	LPAR5	NP_065133.1	12	12p13.31d
ILMN_2402817	-3.7349	0.7200	2.80E-07	0.000164997	ZBTB16	NP_001018011.1	11	11q23.2a
ILMN_1760374	7.2576	1.4066	3.23E-07	0.000185165	CD8A	NP_001759.3	2	2p11.2e
ILMN_1680772	4.2635	0.8368	4.49E-07	0.000244702	LOC642083	XP_947821.1		
ILMN_2366212	-7.1843	1.4094	4.44E-07	0.000244702	CD79B	NP_001035022.1	17	17q23.3b
ILMN_1684349	-7.2084	1.4212	5.05E-07	0.000255661	IL2RB	NP_000869.1	22	22q12.3d
ILMN_1726769	-4.1116	0.8120	5.26E-07	0.000255661	CNDP2	NP_060705.1	18	18q22.3d
ILMN_1785439	-6.0634	1.1977	5.30E-07	0.000255661	CD79B	NP_067613.1	17	17q23.3b
ILMN_1791912	-5.7486	1.1347	5.19E-07	0.000255661	SIDT2	NP_001035545.1	11	11q23.3b
ILMN_2073289	-4.8478	0.9564	5.13E-07	0.000255661	MTSS1	NP_055566.2	8	8q24.13d
ILMN_1726189	-3.6717	0.7313	6.54E-07	0.000304129	MS4A14	NP_115986.3	11	11q12.2a
ILMN_2103919	3.0130	0.6003	6.59E-07	0.000304129	LRFN3	NP_078785.1	19	19q13.12a
ILMN_1659075	-4.6181	0.9215	6.85E-07	0.000309544	HLA-DOA	NP_002110.1	6	6p21.32a
ILMN_2044471	-4.4943	0.8989	7.26E-07	0.000321102	NCR3	NP_667341.1	6	6p21.33a
ILMN_1688775	-3.7605	0.7549	7.96E-07	0.000344996	METRNL	XP_946559.1		17q25.3h
ILMN_1678238	11.2618	2.2705	8.85E-07	0.000361337	ZNF683	NP_775845.2	1	1p36.11b
ILMN_1687440	-5.0770	1.0229	8.71E-07	0.000361337	HIPK2	NP_073577.2	7	7q34b
ILMN_1783149	-3.5526	0.7156	8.66E-07	0.000361337	CDH23	NP_071407.3	10	10q22.1d-q22.1e
ILMN_1790549	-3.7799	0.7647	9.63E-07	0.00038595	TSPAN3	NP_005715.1	15	15q24.3a
ILMN_1706015	-3.4385	0.7002	1.13E-06	0.000429816	FAM43A	NP_710157.2	3	3q29d
ILMN_2111229	-3.1022	0.6318	1.13E-06	0.000429816	BZRAP1	NP_004749.1	17	17q22d
ILMN_2313730	-5.1127	1.0412	1.13E-06	0.000429816	RHOC	NP_786886.1	1	1p13.2c
ILMN_2352131	-3.1882	0.6501	1.17E-06	0.000435294	ERBB2	NP_004439.2	17	17q12c
ILMN_1761733	-4.3548	0.8891	1.20E-06	0.000440691	HLA-DMB	NP_002109.1	6	6p21.32a
ILMN_3306672	4.4223	0.9081	1.38E-06	0.000497553	PATL2		15	15q21.1a
ILMN_1664063	-3.8971	0.8062	1.64E-06	0.000581938	FAM129C	NP_775815.2	19	19p13.11d
ILMN_1676099	-3.7607	0.7789	1.69E-06	0.000585771	SPON2	NP_036577.1	4	4p16.3c
ILMN_1864900	5.7175	1.1847	1.71E-06	0.000585771	MIAT		22	22q12.1a
ILMN_1673305	-4.6841	0.9727	1.80E-06	0.000588247	RHOC	NP_001036143.1	1	1p13.2c

ProbeID	Effect	SD	*p* value	*p* value FDR	Gene	Protein_Product	Chr	Cytoband
ILMN_1873034	5.7478	1.1926	1.77E-06	0.000588247			14	
ILMN_2255133	-4.7376	0.9838	1.80E-06	0.000588247	BCL11A	NP_075044.2	2	2p16.1a
ILMN_1660462	5.3059	1.1043	1.90E-06	0.000610758	MCOLN2	NP_694991.2	1	1p22.3f
ILMN_2353732	9.5669	1.9943	1.97E-06	0.000623914	CD8A	NP_741969.1	2	2p11.2e
ILMN_1691539	3.1592	0.6628	2.28E-06	0.000711975	LAT	NP_001014987.1	16	16p11.2e
ILMN_1768482	9.1598	1.9445	2.98E-06	0.000916306	CD8A	NP_741969.1	2	2p11.2e
ILMN_1681415	-2.8167	0.6015	3.39E-06	0.001029329	SIGLEC7	NP_057627.2	19	19q13.33d
ILMN_1886655	4.9436	1.0573	3.51E-06	0.001051247			7	
ILMN_2399174	-3.5327	0.7590	3.89E-06	0.001147502	TRAK1	NP_001036111.1	3	3p22.1b-p22.1a
ILMN_1741143	-4.1482	0.8933	4.09E-06	0.001172732	TXK	NP_003319.2	4	4p12a
ILMN_3219806	-4.2541	0.9160	4.08E-06	0.001172732	LOC643384		4	4p16.2b
ILMN_1746565	4.3241	0.9324	4.20E-06	0.001190381	CD6	NP_006716.2	11	11q12.2a
ILMN_1718565	-5.1447	1.1177	4.95E-06	0.001382921	CDKN1C	NP_000067.1	11	11p15.4d
ILMN_1701906	-4.2508	0.9253	5.15E-06	0.001409677	CD300C	NP_006669.1	17	17q25.1b
ILMN_1871233	2.9576	0.6439	5.18E-06	0.001409677			14	
ILMN_1874689	-4.7707	1.0468	6.11E-06	0.001642627			9	
ILMN_1714820	3.8115	0.8434	7.29E-06	0.001935205	ITGB1	NP_391988.1	10	10p11.22b
ILMN_1800512	-3.8659	0.8562	7.41E-06	0.001943433	HMOX1	NP_002124.1	22	22q12.3c
ILMN_1754894	-3.7813	0.8447	8.85E-06	0.00229089	C1orf162	NP_777556.1	1	1p13.2d
ILMN_1663080	-3.1416	0.7058	9.94E-06	0.002513266	LFNG	NP_001035257.1	7	7p22.2c
ILMN_1672503	-3.8779	0.8712	9.93E-06	0.002513266	DPYSL2	NP_001377.1	8	8p21.2a
ILMN_1773963	-3.3688	0.7584	1.04E-05	0.002590003	GNA15	NP_002059.1	19	19p13.3f
ILMN_1813338	5.2365	1.1812	1.08E-05	0.002659763	LAG3	NP_002277.3	12	12p13.31d
ILMN_1706426	2.8529	0.6456	1.15E-05	0.002709352	DSTN	NP_001011546.1	20	20p12.1a
ILMN_1733421	2.5721	0.5819	1.14E-05	0.002709352	PRKCQ	NP_006248.1	10	10p15.1a
ILMN_2328666	-2.3492	0.5309	1.12E-05	0.002709352	CD83	NP_004224.1	6	6p23b
ILMN_2379644	-4.4255	1.0004	1.12E-05	0.002709352	CD74	NP_004346.1	5	5q33.1c
ILMN_1778668	-4.8829	1.1057	1.16E-05	0.002716532	TAGLN	NP_003177.2	11	11q23.3b
ILMN_1736311	-2.7525	0.6264	1.28E-05	0.002960015	POU2F2	NP_002689.1	19	19q13.2c
ILMN_1653026	-3.7163	0.8475	1.34E-05	0.00302643	PLAC8	NP_057703.1	4	4q21.22a
ILMN_1822307	3.9161	0.8931	1.34E-05	0.00302643			7	
ILMN_3233930	-3.0165	0.6883	1.35E-05	0.003027151	LOC390557	XP_001727025.1	15	15q13.2a

R39 R38 R37 R36 R35 R34 R33 R32 R31 R30 R29 R28 R27 R26 R25 R24 R23 R22 R21 R20 R19 R18 R17 R16 R15 R14 R13 R12 R11 R10 R9 R8 R7 R6 R5 R4 R3 R2 R1

ProbeID	Effect	SD	*p* value	*p* value FDR	Gene	Protein_Product	Chr	Cytoband
ILMN_1814726	-3.1457	0.7188	1.39E-05	0.003073461	SCARB2	NP_005497.1	4	4q21.1a
ILMN_1745994	-3.9861	0.9115	1.41E-05	0.003089278	GAS7	NP_958839.1	17	17p13.1b
ILMN_1792538	-4.4597	1.0244	1.54E-05	0.003336479	CD7	NP_006128.1	17	17q25.3g
ILMN_1780825	-3.5547	0.8177	1.58E-05	0.003394903	RRAS	NP_006261.1	19	19q13.33b
ILMN_1723004	-3.1812	0.7338	1.67E-05	0.003541635	CD72	NP_001773.1	9	9p13.3b-p13.3a
ILMN_2221564	4.0620	0.9376	1.69E-05	0.003555535	LYAR	NP_060286.1	4	4p16.2b
ILMN_1719998	-3.3767	0.7814	1.78E-05	0.003613389	C9orf45		9	9q33.2b
ILMN_1735155	-3.1336	0.7252	1.77E-05	0.003613389	GLB1	NP_000395.1	3	3p22.3c
ILMN_1772218	-4.5773	1.0596	1.79E-05	0.003613389	HLA-DPA1	NP_291032.2	6	6p21.32a
ILMN_2305407	-3.6706	0.8496	1.78E-05	0.003613389	ZBTB16	NP_005997.2	11	11q23.2a
ILMN_1810289	-3.7048	0.8582	1.81E-05	0.003629459	FER1L3	NP_579899.1	10	10q23.33a-q23.33b
ILMN_1668277	-4.9016	1.1372	1.86E-05	0.003679059	BLK	NP_001706.2	8	8p23.1b
ILMN_1690566	-2.8496	0.6613	1.87E-05	0.003679059	RASSF4	NP_114412.2	10	10q11.21c
ILMN_1684346	-3.0655	0.7120	1.91E-05	0.003712668	TNFAIP8L1	NP_689575.1	19	19p13.3d
ILMN_2059549	-3.6837	0.8571	1.97E-05	0.003804589	SYK	NP_003168.2	9	9q22.2b
ILMN_1732452	-3.0147	0.7097	2.45E-05	0.004690535	MAPKAPK3	NP_004626.1	3	3p21.31b
ILMN_1728478	-3.8697	0.9116	2.48E-05	0.004702724	CXCL16	NP_071342.1	17	17p13.2c
ILMN_1693014	-3.4645	0.8184	2.61E-05	0.004864539	CEBPB	NP_005185.2	20	20q13.13e
ILMN_1696038	1.9781	0.4673	2.61E-05	0.004864539	LOC150051	XP_097792.4	21	21q22.11a
ILMN_2404154	-4.2650	1.0090	2.68E-05	0.004954898	SERPINA1	NP_001002235.1	14	14q32.13a
ILMN_1803745	-2.5375	0.6016	2.79E-05	0.005109897	SUOX	NP_000447.2	12	12q13.2c
ILMN_2325168	-2.9110	0.6906	2.82E-05	0.00512525	ARRB1	NP_064647.1	11	11q13.4c
ILMN_1711702	3.3172	0.7874	2.85E-05	0.005130269	CLEC2D	NP_037401.1	12	12p13.31a
ILMN_1742001	-6.1886	1.4716	2.95E-05	0.005258147	CD160	NP_008984.1	1	1q21.1b
ILMN_1768551	3.4146	0.8133	3.03E-05	0.005369899	LOC197135	XP_945795.1		15q21.1a
ILMN_1714197	-2.9296	0.7000	3.22E-05	0.005643611	ACSS2	NP_061147.1	20	20q11.22b
ILMN_2397721	-2.7664	0.6619	3.29E-05	0.00572596	GLB1	NP_001073279.1	3	3p22.3c
ILMN_1731233	11.6270	2.7876	3.41E-05	0.005862693	GZMH	NP_219491.1	14	14q12a
ILMN_2336595	-2.9515	0.7077	3.42E-05	0.005862693	ACSS2	NP_001070020.1	20	20q11.22b
ILMN_1851610	2.0046	0.4823	3.63E-05	0.006168592			11	
ILMN_3307757	-2.3004	0.5546	3.77E-05	0.006346449	C17orf87	NP_996986.1	17	17p13.2b
ILMN_1734276	-1.9918	0.4804	3.80E-05	0.006349518	PMEPA1	NP_954638.1	20	20q13.31a
ILMN_1810275	-3.5811	0.8646	3.86E-05	0.006407996	SLC7A7	NP_003973.2	14	14q11.2e-q11.2f

ProbeID	Effect	SD	*p* value	*p* value FDR	Gene	Protein_Product	Chr	Cytoband
ILMN_1803652	-2.6409	0.6379	3.90E-05	0.00641685	C9orf91	NP_694590.2	9	9q32e
ILMN_1671791	-2.6715	0.6462	3.99E-05	0.006417735	PCK2	NP_004554.2	14	14q12a
ILMN_1676003	-2.3921	0.5782	3.94E-05	0.006417735	PNOC	NP_006219.1	8	8p21.1d
ILMN_1786601	-2.6975	0.6524	3.98E-05	0.006417735	PLAGL2	NP_002648.1	20	20q11.21b
ILMN_1707077	-2.7917	0.6769	4.17E-05	0.006614673	SORT1	NP_002950.3	1	1p13.3b
ILMN_2307032	-3.3344	0.8087	4.18E-05	0.006614673	OSBPL5	NP_663613.1	11	11p15.4d
ILMN_3202591	3.5730	0.8668	4.20E-05	0.006614673	LOC440311			15q26.2b
ILMN_2055477	-2.5074	0.6101	4.43E-05	0.006911223	EXOSC7	NP_055819.1	3	3p21.31k
ILMN_1752899	-3.8148	0.9299	4.57E-05	0.007074816	BCL11A	NP_075044.2	2	2p16.1a
ILMN_2316386	-4.4503	1.0852	4.60E-05	0.007074816	GPBAR1	NP_001070659.1	2	2q35e
ILMN_2130411	-2.4328	0.5942	4.73E-05	0.007225214	KDELR1	NP_006792.1	19	19q13.32c
ILMN_1661646	-2.8419	0.6954	4.87E-05	0.007394904	BANK1	NP_001077376.1	4	4q24a
ILMN_1752478	-2.8757	0.7051	5.05E-05	0.007552722	DHRS3	NP_004744.2	1	1p36.22a-p36.21d
ILMN_2331082	-5.3401	1.3090	5.03E-05	0.007552722	MS4A7	NP_996821.1	11	11q12.2a
ILMN_1773352	6.2509	1.5348	5.18E-05	0.007686423	CCL5	NP_002976.2	17	17q12b
ILMN_1670134	-2.2578	0.5546	5.22E-05	0.007700711	FADS1	NP_037534.2	11	11q12.2b
ILMN_1752046	-2.7375	0.6734	5.34E-05	0.007824039	SH2B3	NP_005466.1	12	12q24.12a
ILMN_1744517	-3.5680	0.8781	5.39E-05	0.007835072	GNS	NP_002067.1	12	12q14.2b-q14.3a
ILMN_1796063	-2.7253	0.6737	5.81E-05	0.008285644	TRIM44	NP_060053.2	11	11p13a
ILMN_1802151	-3.8387	0.9485	5.77E-05	0.008285644	OSBPL5	NP_065947.1	11	11p15.4d
ILMN_1807211	-2.4380	0.6025	5.78E-05	0.008285644	NICN1	NP_115692.1	3	3p21.31d
ILMN_1804396	-6.0394	1.4945	5.91E-05	0.008365317	C14orf4	NP_078772.1	14	14q24.3c
ILMN_2355953	-2.5015	0.6197	6.02E-05	0.008463462	LILRB4	NP_001074907.1	19	19q13.42a-q13.42b
ILMN_1714397	-2.1833	0.5412	6.07E-05	0.008483436	CRYL1	NP_057058.2	13	13q12.11b
ILMN_1743340	-1.8839	0.4682	6.36E-05	0.008830401	CHST14	NP_569735.1	15	15q15.1a
ILMN_2150196	-5.9283	1.4749	6.46E-05	0.008915172	LRRC25	NP_660299.2	19	19p13.11c
ILMN_1699646	2.0318	0.5057	6.51E-05	0.008918614	APBB1	NP_001155.1	11	11p15.4c
ILMN_1667893	-3.0653	0.7633	6.57E-05	0.008940021	TNS3	NP_073585.8	7	7p12.3c
ILMN_1778681	-2.5495	0.6359	6.74E-05	0.009064824	EBF1	NP_076870.1	5	5q33.3c
ILMN_2414762	-2.9243	0.7294	6.74E-05	0.009064824	TLR10	NP_112218.2	4	4p14c
ILMN_2338348	3.2303	0.8081	7.08E-05	0.009457053	UBASH3A	NP_001001895.1	21	21q22.3b
ILMN_1743104	-2.1850	0.5473	7.23E-05	0.009590422	RBM4B	NP_113680.1	11	11q13.1e

ProbeID	Effect	SD	*p* value	*p* value FDR	Gene	Protein_Product	Chr	Cytoband
ILMN_1771376	-2.8599	0.7170	7.34E-05	0.009679975	PEA15	NP_003759.1	1	1q23.2d
ILMN_1762972	-3.4043	0.8541	7.42E-05	0.009728704	CHD9	NP_079410.4	16	16q12.2a
ILMN_1659227	-4.8310	1.2151	7.74E-05	0.010022073	CD79A	NP_001774.1	19	19q13.2c
ILMN_1723467	4.2912	1.0791	7.71E-05	0.010022073	ITGB1	NP_002202.2	10	10p11.22b
ILMN_1664464	-7.0537	1.7803	8.19E-05	0.010419177	PTGDS	NP_000945.3	9	9q34.3e
ILMN_1702501	-2.0464	0.5162	8.12E-05	0.010419177	RPS6KA2	NP_001006933.1	6	6q27c
ILMN_1778136	-2.1120	0.5330	8.17E-05	0.010419177	ZMYND15	NP_115641.1	17	17p13.2c
ILMN_1727315	-2.1605	0.5460	8.37E-05	0.010585176	DENND1A	NP_079096.2	9	9q33.2b
ILMN_1683146	-3.0598	0.7749	8.65E-05	0.010876066	FTH1	NP_002023.2	11	11q12.3a
ILMN_1729281	-2.0979	0.5321	8.87E-05	0.011011292	SPHK2	NP_064511.2	19	19q13.33a
ILMN_1746618	-1.9972	0.5065	8.84E-05	0.011011292	PAQR7	NP_848509.1	1	1p36.11b
ILMN_1696419	3.7034	0.9405	9.05E-05	0.011121368	STOM	NP_004090.4	9	9q33.2a
ILMN_1789233	-2.8792	0.7312	9.06E-05	0.011121368	VPS37C	NP_060436.4	11	11q12.2b
ILMN_2381899	3.6122	0.9181	9.17E-05	0.011190716	OPTN	NP_001008214.1	10	10p13e
ILMN_1861609	1.7308	0.4412	9.61E-05	0.011665404				
ILMN_1898771	3.0018	0.7658	9.72E-05	0.011731236			14	
ILMN_3236036	-3.4601	0.8835	9.86E-05	0.01183069	LOC283663		15	15q21.3d
ILMN_1669497	-3.4179	0.8756	0.000103938	0.012401377	OSBPL10	NP_060254.2	3	3p23a
ILMN_1710207	-2.5059	0.6429	0.00010622	0.012602807	C10orf6	NP_060591.2	10	10q24.31a
ILMN_2342066	-2.1738	0.5579	0.000107021	0.012627307	METRNL	NP_001004431.1	17	17q25.3h
ILMN_1811335	2.3216	0.5961	0.000107678	0.012634639	LOC653472	XP_937252.1	15	
ILMN_1653292	-3.0326	0.7813	0.000113703	0.013058407	PFKFB4	NP_004558.1	3	3p21.31e
ILMN_1714602	-2.7953	0.7202	0.000113642	0.013058407	CD86	NP_008820.2	3	3q13.33c
ILMN_1803162	3.4771	0.8955	0.000112981	0.013058407	RTEL1	NP_057518.1	20	20q13.33e
ILMN_2380967	-2.4302	0.6261	0.000113749	0.013058407	DNASE1L1	NP_001009934.1	X	Xq28g
ILMN_3235853	3.2732	0.8438	0.000114609	0.0130864	S1PR1	NP_001391.2	1	1p21.2a
ILMN_1714364	-2.5938	0.6696	0.000117312	0.013323376	PTK2	NP_005598.3	8	8q24.3c
ILMN_1693242	-2.5692	0.6641	0.00011954	0.013362079	ZNF296	NP_660331.1	19	19q13.32a
ILMN_1771179	2.5324	0.6544	0.000119169	0.013362079	CYB561	NP_001906.3	17	17q23.3a
ILMN_2128428	-2.9386	0.7594	0.000119152	0.013362079	DAB2	NP_001334.1	5	5p13.1c
ILMN_2068122	1.9272	0.4999	0.000126217	0.014034507	TMEM65	NP_919267.1	8	8q24.13d
ILMN_1682699	-2.8447	0.7389	0.000129105	0.014273683	PBX2	NP_002577.2	6	6p21.32b
ILMN_1697440	-1.9299	0.5018	0.000131056	0.014273683	PRPF4	NP_004688.2	9	9q32c

R39 R38 R37 R36 R35 R34 R33 R32 R31 R30 R29 R28 R27 R26 R25 R24 R23 R22 R21 R20 R19 R18 R17 R16 R15 R14 R13 R12 R11 R10 R9 R8 R7 R6 R5 R4 R3 R2 R1

ProbeID	Effect	SD	*p* value	*p* value FDR	Gene	Protein_Product	Chr	Cytoband
ILMN_1802843	-1.9720	0.5126	0.000130303	0.014273683	PRCC	NP_005964.3	1	1q23.1a
ILMN_2337655	-5.0875	1.3227	0.0001308	0.014273683	WARS	NP_004175.2	14	14q32.2b
ILMN_1707051	-2.4364	0.6348	0.000135134	0.01464278	NFATC1	NP_765978.1	18	18q23d
ILMN_2052790	-2.1252	0.5564	0.000145492	0.015621819	NONO	NP_031389.3	X	Xq13.1d
ILMN_3284119	3.1666	0.8294	0.000146376	0.015621819	LOC399804			10q23.33d
ILMN_3302919	-2.8817	0.7547	0.000146347	0.015621819	MYOF	NP_038479.1	10	10q23.33a-q23.33b
ILMN_1714861	-3.6214	0.9497	0.000149244	0.015848188	CD68	NP_001242.1	17	17p13.1d
ILMN_1699027	1.8284	0.4798	0.000150863	0.015940443	LOC642194	XP_947871.1		
ILMN_1671509	-1.9629	0.5158	0.000153614	0.016123893	CCL3	NP_002974.1	17	17q12b
ILMN_1699521	2.2283	0.5856	0.000154118	0.016123893	KIAA1641	XP_949154.1		2q11.2b
ILMN_1762308	-2.5751	0.6775	0.000156645	0.016228448	LOC654191	XP_945735.1		
ILMN_2331087	-3.9013	1.0261	0.000155886	0.016228448	MS4A7	NP_996821.1	11	11q12.2a
ILMN_1807825	-3.5043	0.9227	0.000158572	0.016348356	LY86	NP_004262.1	6	6p25.1a
ILMN_1715583	-2.3036	0.6067	0.000159377	0.016351904	BOP1	NP_056016.1	8	8q24.3g
ILMN_1765633	2.1837	0.5763	0.000164151	0.01640608	FAM10A7		7	7q33a
ILMN_1780057	-2.7719	0.7311	0.000162776	0.01640608	RENBP	NP_002901.1	X	Xq28f
ILMN_1793729	-3.5827	0.9451	0.000163162	0.01640608	C15orf39	NP_056307.2	15	15q24.2a
ILMN_2157441	-3.6844	0.9710	0.000160868	0.01640608	HLA-DRA	NP_061984.2	6	6p21.32b
ILMN_2310968	-2.6744	0.7059	0.00016454	0.01640608	RUFY1	NP_001035541.1	5	5q35.3d
ILMN_2338323	3.0149	0.7956	0.000163824	0.01640608	CDC25B	NP_004349.1	20	20p13b
ILMN_1689655	-3.3820	0.8940	0.0001682	0.016692624	HLA-DRA	NP_061984.2	6	6p21.32b
ILMN_1766637	-2.7787	0.7351	0.000169967	0.016789586	GLA	NP_000160.1	X	Xq22.1c
ILMN_1666019	1.9315	0.5126	0.000178331	0.017403806	ADNP	NP_056154.1	20	20q13.13f
ILMN_1782704	-5.1677	1.3716	0.000178643	0.017403806	CD19	NP_001761.3	16	16p11.2e
ILMN_1811779	-2.1668	0.5749	0.000177511	0.017403806	MGC24103		9	9p22.3a
ILMN_1736441	-2.2281	0.5917	0.000180026	0.017419015	PDXP	NP_064711.1	22	22q13.1a
ILMN_2363058	-2.2201	0.5897	0.00018044	0.017419015	PAOX	NP_997011.1	10	10q26.3f
ILMN_2059535	-3.8580	1.0251	0.000181607	0.01745231	PPM1F	NP_055449.1	22	22q11.22a
ILMN_1666269	-5.0382	1.3413	0.000186763	0.017867015	CTSZ	NP_001327.2	20	20q13.32b
ILMN_1672878	-2.6640	0.7100	0.000189698	0.01798866	ABR	NP_001083.2	17	17p13.3f-p13.3e
ILMN_2210129	1.8249	0.4863	0.000189729	0.01798866	PRIM1	NP_000937.1	12	12q13.3a
ILMN_1699644	-1.8522	0.4943	0.000193517	0.018266272	3-Mar	XP_001127871.1		5q23.2e-q23.2f

R39 R38 R37 R36 R35 R34 R33 R32 R31 R30 R29 R28 R27 R26 R25 R24 R23 R22 R21 R20 R19 R18 R17 R16 R15 R14 R13 R12 R11 R10 R9 R8 R7 R6 R5 R4 R3 R2 R1

ProbeID	Effect	SD	*p* value	*p* value FDR	Gene	Protein_Product	Chr	Cytoband
ILMN_1689324	1.7792	0.4750	0.000194531	0.01828073	LOC650751	XP_944926.2		
ILMN_2371700	-2.0453	0.5464	0.000196445	0.018379269	UCHL5IP	NP_059988.3	X	Xq28f
ILMN_1698367	2.6198	0.7004	0.00019853	0.018492882	CD84	NP_003865.1	1	1q23.2d-q23.3a
ILMN_1797522	-2.9317	0.7844	0.000201121	0.018599572	DUSP3	NP_004081.1	17	17q21.31b
ILMN_1826637	1.7086	0.4572	0.000201427	0.018599572			16	
ILMN_1873677	-1.7454	0.4674	0.000203276	0.018689064			17	
ILMN_1729287	-1.9091	0.5119	0.000207166	0.018964624	NMUR1	NP_006047.2	2	2q37.1b
ILMN_1738883	-1.6905	0.4537	0.00021031	0.01916984	RNF135	NP_115698.3	17	17q11.2c
ILMN_1741200	-2.2264	0.5979	0.000212285	0.019267132	RFX5	NP_001020774.1	1	1q21.2d
ILMN_2129505	-2.8649	0.7704	0.000216191	0.019538174	CYBASC3	NP_705839.2	11	11q12.2b
ILMN_1663866	-4.1908	1.1289	0.000221797	0.019959821	TGFBI	NP_000349.1	5	5q31.1f-q31.2a
ILMN_1740466	-3.1524	0.8499	0.0002246	0.02012677	FAM46A	NP_060103.2	6	6q14.1e
ILMN_2405297	-2.3486	0.6338	0.000227658	0.020315144	NOTCH2	NP_077719.2	1	1p12a
ILMN_1748601	3.0315	0.8185	0.0002292	0.020367186	CD8B	NP_004922.1	2	2p11.2e
ILMN_1773620	-2.7858	0.7533	0.000234163	0.020721453	SMARCC2	NP_620706.1	12	12q13.2c
ILMN_1662451	-4.8943	1.3239	0.000235385	0.02074316	FCER2	NP_001993.2	19	19p13.2e
ILMN_1769013	-3.7462	1.0148	0.00024007	0.020952656	ASGR1	NP_001662.1	17	17p13.1d
ILMN_1803810	-1.9871	0.5382	0.000239901	0.020952656	RRBP1	NP_001036041.1	20	20p12.1a
ILMN_1815658	1.7212	0.4663	0.000240722	0.020952656	SNHG8			3p21.1d
ILMN_1664016	-2.6543	0.7208	0.000248602	0.021181046	ARHGEF18	NP_056133.2	19	19p13.2e
ILMN_1724897	-2.0145	0.5474	0.000251324	0.021181046	C14orf93	NP_068763.1	14	14q11.2f
ILMN_1750167	-2.0969	0.5697	0.000250505	0.021181046	PRR3	NP_079539.2	6	6p21.33b
ILMN_1763447	-3.3364	0.9054	0.000246224	0.021181046	PLXNB2	NP_036533.2	22	22q13.33b
ILMN_1775235	-3.0146	0.8187	0.000248981	0.021181046	AFF3	NP_001020279.1	2	2q11.2c-q11.2d
ILMN_1785177	-1.9701	0.5346	0.000246342	0.021181046	DNAJC14	NP_115740.5	12	12q13.2c
ILMN_2242937	-2.3649	0.6416	0.000245225	0.021181046	ARSB	NP_000037.2	5	5q14.1c
ILMN_2400935	-1.6604	0.4511	0.000250374	0.021181046	TAGLN	NP_003177.2	11	11q23.3b
ILMN_1703955	2.1921	0.5959	0.000252344	0.021182931	FBXO32	NP_680482.1	8	8q24.13c
ILMN_1664698	-2.5694	0.6998	0.00025941	0.021526815	UNC119	NP_473376.1	17	17q11.2a
ILMN_1680738	-2.4993	0.6807	0.000259481	0.021526815	C5orf13	NP_004763.1	5	5q22.1b
ILMN_1713732	-2.1876	0.5957	0.000258805	0.021526815	ABL1	NP_005148.2	9	9q34.12a
ILMN_1802888	-3.9015	1.0640	0.000264194	0.02183249	ZNF185	NP_009081.2	X	Xq28e
ILMN_1724504	-2.1592	0.5892	0.000266367	0.021926777	SETD3	NP_115609.2	14	14q32.2b

ProbeID	Effect	SD	*p* value	*p* value FDR	Gene	Protein_Product	Chr	Cytoband
ILMN_1692163	-1.8034	0.4927	0.000270531	0.02218353	NSDHL	NP_057006.1	X	Xq28e
ILMN_1813846	-2.2827	0.6238	0.000271949	0.022214087	P2RX4	NP_002551.2	12	12q24.31b
ILMN_2193325	-3.1021	0.8499	0.000281677	0.022920541	MMP23B	NP_008914.1	1	1p36.33a
ILMN_1739794	3.2032	0.8779	0.000283288	0.022963603	CD3E	NP_000724.1	11	11q23.3d
ILMN_1723944	5.6822	1.5587	0.000286723	0.023084253	TARP	NP_001003806.1	7	7p14.1e
ILMN_1890134	1.9752	0.5419	0.00028695	0.023084253			14	
ILMN_1743441	1.7850	0.4904	0.000293097	0.023489798	GSTTP2		22	22q11.23b
ILMN_1808587	-2.1571	0.5937	0.000300591	0.023999833	ZFHX3	NP_008816.3	16	16q22.3b
ILMN_1690085	-1.9427	0.5363	0.000312725	0.024782304	STK11IP	NP_443134.2	2	2q35f
ILMN_1756595	-2.8012	0.7731	0.000311743	0.024782304	SH3TC1	NP_061859.3	4	4p16.1d
ILMN_2098325	-1.9698	0.5453	0.000324861	0.025648324	C8orf33	NP_075568.1	8	8q24.3h
ILMN_1652650	2.5699	0.7128	0.000333903	0.026071411	SH3KBP1	NP_001019837.1	X	Xp22.12b
ILMN_1709112	1.5875	0.4403	0.00033315	0.026071411	CHM	NP_000381.1	X	Xq21.2a
ILMN_1716704	2.4965	0.6922	0.000332098	0.026071411	NLRC5	NP_115582.3	16	16q13b-q13c
ILMN_1655623	1.6398	0.4556	0.000341056	0.02653237	LOC731884			
ILMN_1808590	-2.5268	0.7032	0.000349407	0.027082849	GUCY1A3	NP_000847.2	4	4q32.1b
ILMN_1794927	-3.6138	1.0065	0.000353076	0.027267773	LOC90925	NP_787066.1		
ILMN_1800739	-2.4926	0.6949	0.000357818	0.027533872	SPINT2	NP_066925.1	19	19q13.2a
ILMN_1707475	-2.1294	0.5942	0.000361993	0.027754569	UBE2E2	NP_689866.1	3	3p24.3a
ILMN_3192316	2.8962	0.8084	0.000363407	0.027762704	LOC100129237			8q12.1b
ILMN_1700762	1.9176	0.5355	0.000366132	0.02777109	PBX4	NP_079521.1	19	19p13.11a
ILMN_1781155	-2.8383	0.7926	0.000365705	0.02777109	LYN	NP_002341.1	8	8q12.1a
ILMN_1661954	1.6391	0.4581	0.000370275	0.027985398	C21orf24	NP_001001789.1	21	21q22.2a
ILMN_1764709	-3.7027	1.0360	0.000375596	0.02808774	MAFB	NP_005452.2	20	20q12b
ILMN_1789879	1.6086	0.4499	0.000373212	0.02808774	WDR35	NP_001006658.1	2	2p24.1d
ILMN_2370738	-2.0114	0.5628	0.000375466	0.02808774	SLC24A4	NP_705933.1	14	14q32.12b
ILMN_1653404	-2.0023	0.5604	0.00037705	0.028097471	NKIRAS2	NP_060065.2	17	17q21.2b
ILMN_2377496	-2.2771	0.6378	0.000380967	0.028290141	ERCC1	NP_001974.1	19	19q13.32a
ILMN_3237448	-2.0131	0.5644	0.000385896	0.028556276	BEND5	NP_078879.2	1	1p33b
ILMN_2209115	-3.6586	1.0260	0.000387392	0.028567469	MAK	NP_005897.1	6	6p24.2a
ILMN_2058468	-2.7860	0.7821	0.000392587	0.02885039	BACH2	NP_068585.1	6	6q15d
ILMN_2339863	3.0338	0.8525	0.000397951	0.029143744	VPS28	NP_057292.1	8	8q24.3h

ProbeID	Effect	SD	*p* value	*p* value FDR	Gene	Protein_Product	Chr	Cytoband
ILMN_1657862	-2.2411	0.6306	0.000405371	0.029222009	AHCY	NP_000678.1	20	20q11.22a
ILMN_1687707	1.7282	0.4863	0.000405472	0.029222009	LOC646795	XP_939035.2	22	22q13.1a
ILMN_1739749	-1.9693	0.5537	0.000401052	0.029222009	B3GALT6	NP_542172.2	1	1p36.33a
ILMN_1749109	-3.2796	0.9230	0.000405899	0.029222009	PSAP	NP_001035930.1	10	10q22.1e
ILMN_1805403	1.6266	0.4575	0.000402261	0.029222009	LOC390547	XP_372553.3	15	
ILMN_2158003	-2.4236	0.6836	0.000417838	0.029979843	KIAA1683	NP_079525.1	19	19p13.11c
ILMN_1769779	-2.9529	0.8331	0.000419312	0.029984309	PTP4A3	NP_009010.2	8	8q24.3d
ILMN_3182069	-1.8670	0.5269	0.000421295	0.030025063	LOC100129697	XP_001732874.1		16q24.3a
ILMN_1658885	-2.4807	0.7012	0.000429613	0.030309712	DAGLB	NP_631918.1	7	7p22.1b
ILMN_1686862	2.4224	0.6848	0.000430998	0.030309712	HLX	NP_068777.1	1	1q41d
ILMN_1710017	-4.0021	1.1307	0.000427272	0.030309712	CD79B	NP_000617.1	17	17q23.3b
ILMN_1718610	-2.5276	0.7143	0.000429181	0.030309712	ARHGAP17	NP_060524.4	16	16p12.1b
ILMN_1695311	-3.5963	1.0172	0.000434028	0.030422046	HLA-DMA	NP_006111.2	6	6p21.32a
ILMN_1727281	-2.0448	0.5788	0.00043828	0.030518626	TAF6L	NP_006464.1	11	11q12.3b
ILMN_1740430	1.5486	0.4383	0.000437745	0.030518626	SLC2A4RG	NP_064446.2	20	20q13.33e
ILMN_1694731	-2.9616	0.8387	0.000440517	0.03057415	CLCN7	NP_001278.1	16	16p13.3e
ILMN_1667932	2.7858	0.7899	0.00044797	0.030789632	LOC652726	XP_947444.1		
ILMN_1740385	-1.8347	0.5201	0.000446961	0.030789632	CEP164	NP_055771.4	11	11q23.3b-q23.3c
ILMN_2398847	-1.8239	0.5169	0.000445073	0.030789632	ARHGAP17	NP_060524.4	16	16p12.1b
ILMN_1669015	-2.0219	0.5737	0.000451756	0.030949666	XPNPEP1	NP_065116.2	10	10q25.1e
ILMN_1662306	1.7054	0.4843	0.000456804	0.031111871	RABL3	NP_776186.2	3	3q13.33b
ILMN_1769299	-3.0771	0.8738	0.000457054	0.031111871	MTMR11	NP_870988.2	1	1q21.2a
ILMN_1664434	-2.3418	0.6658	0.000463679	0.03121346	TCF3	NP_003191.1	19	19p13.3h
ILMN_1679169	1.8988	0.5397	0.000462501	0.03121346	LOC642223	XP_941468.1		
ILMN_1727045	-2.6579	0.7559	0.000465894	0.03121346	RASGRP3	NP_733772.1	2	2p22.3d
ILMN_1753426	-1.9519	0.5550	0.000464677	0.03121346	KIAA0556	NP_056017.1	16	16p12.1a-p11.2e
ILMN_1857017	-2.0588	0.5852	0.000462332	0.03121346			7	
ILMN_1678579	-1.9951	0.5677	0.000469296	0.031274186	CPT2	NP_000089.1	1	1p32.3c
ILMN_1906721	1.6297	0.4638	0.000469746	0.031274186			6	
ILMN_2093343	-3.1428	0.8959	0.000480029	0.031858948	PLAC8	NP_057703.1	4	4q21.22a
ILMN_1755710	-1.7670	0.5039	0.00048225	0.031906651	EFNA4	NP_872631.1	1	1q22a
ILMN_1776088	-2.2219	0.6340	0.000486978	0.03211936	NAT9	NP_056469.2	17	17q25.1b
ILMN_1742200	1.5686	0.4480	0.000491894	0.032343189	LOC440225		15	15q11.2a-q11.2b

ProbeID	Effect	SD	*p* value	*p* value FDR	Gene	Protein_Product	Chr	Cytoband
ILMN_1653504	2.8076	0.8036	0.000505818	0.032851844	EDG1	NP_001391.2	1	1p21.2a
ILMN_1655740	1.4846	0.4246	0.00050126	0.032851844	SNAI2	NP_003059.1	8	8q11.21c
ILMN_1774604	-2.8151	0.8056	0.000505406	0.032851844	PNKD	NP_072094.1	2	2q35e
ILMN_2085862	-3.2599	0.9327	0.000503804	0.032851844	SLC15A3	NP_057666.1	11	11q12.2a
ILMN_1819608	-2.9732	0.8513	0.000508631	0.032933854			11	
ILMN_1655469	-2.1610	0.6196	0.000517882	0.033113318	TSPAN3	NP_005715.1	15	15q24.3a
ILMN_1685521	-4.4527	1.2760	0.000514095	0.033113318	KLRF1	NP_057607.1	12	12p13.31a
ILMN_1785356	-2.4155	0.6924	0.000516015	0.033113318	DENND5A	NP_056028.2	11	11p15.4a
ILMN_2230902	-2.9999	0.8603	0.000519198	0.033113318	CTNNA1	NP_001894.2	5	5q31.2d
ILMN_2394381	-1.9881	0.5701	0.000518051	0.033113318	CLN3	NP_000077.1	16	16p11.2e
ILMN_1691431	1.7279	0.4963	0.000529077	0.03361348	FAM18B	XP_944589.1		17p11.2f
ILMN_1755862	-2.0555	0.5905	0.000530206	0.03361348	PFAS	NP_036525.1	17	17p13.1c
ILMN_1810910	1.5700	0.4514	0.000536436	0.033907211	CFH	NP_001014975.1	1	1q31.3c
ILMN_1660582	-1.5746	0.4530	0.000540236	0.034046108	LIG3	NP_039269.2	17	17q12a
ILMN_1753608	-2.1571	0.6207	0.000541905	0.034050254	TMEM131	NP_056163.1	2	2q11.2b
ILMN_1728107	-2.8232	0.8133	0.000550129	0.034363645	GNG7	NP_443079.1	19	19p13.3g
ILMN_1797975	3.4245	0.9865	0.000549307	0.034363645	CXCR3	NP_001495.1	X	Xq13.1d
ILMN_2329773	2.5004	0.7208	0.000555061	0.034570018	RAB27A	NP_004571.2	15	15q21.3b
ILMN_1769451	-2.2663	0.6539	0.000560834	0.034786512	ILVBL	NP_006835.2	19	19p13.12b
ILMN_1795835	-2.6060	0.7520	0.000561812	0.034786512	LOC338758	XP_936452.1	12	12q21.33b
ILMN_1846776	1.8616	0.5374	0.000564247	0.034835695			21	
ILMN_1754121	-2.6263	0.7584	0.000567115	0.034911298	CSK	NP_004374.1	15	15q24.1b
ILMN_1718542	1.6527	0.4778	0.000574619	0.035196872	LOC643430	XP_933208.1	15	
ILMN_1791771	-2.5461	0.7361	0.000575069	0.035196872	HCK	NP_002101.2	20	20q11.21b
ILMN_1654313	2.2066	0.6385	0.000581221	0.035268475	LOC730995	XP_001130291.1		
ILMN_1805225	-2.2361	0.6470	0.000581093	0.035268475	LPCAT3	NP_005759.4	12	12p13.31d
ILMN_1815734	-2.4922	0.7209	0.000579238	0.035268475	FCHSD2	NP_055639.1	11	11q13.4b
ILMN_1716973	-2.2967	0.6656	0.000593402	0.035905024	POLM	NP_037416.1	7	7p13d
ILMN_2296950	-2.3616	0.6849	0.000597766	0.036066346	APOBEC3F	NP_001006667.1	22	22q13.1c
ILMN_1807152	-1.6186	0.4697	0.000602922	0.036274398	GOLGA1	NP_002068.1	9	9q33.3a
ILMN_1668514	-2.2452	0.6519	0.000606842	0.036304511	PIP5K1C	NP_036530.1	19	19p13.3e
ILMN_1684289	-2.0509	0.5954	0.000606393	0.036304511	PNPO	NP_060599.1	17	17q21.32b

R39 R38 R37 R36 R35 R34 R33 R32 R31 R30 R29 R28 R27 R26 R25 R24 R23 R22 R21 R20 R19 R18 R17 R16 R15 R14 R13 R12 R11 R10 R9 R8 R7 R6 R5 R4 R3 R2 R1

ProbeID	Effect	SD	*p* value	*p* value FDR	Gene	Protein_Product	Chr	Cytoband
ILMN_1712026	-2.6221	0.7616	0.000609716	0.036374028	NLRP3	NP_001073289.1	1	1q44e
ILMN_2404625	2.4967	0.7255	0.000612788	0.036454854	LAT	NP_001014987.1	16	16p11.2e
ILMN_1715804	-2.0853	0.6061	0.000614792	0.036471909	PITPNA	NP_006215.1	17	17p13.3e
ILMN_1788605	1.7500	0.5089	0.000619625	0.036656251	LOC644623		13	13q14.3d
ILMN_1667232	-7.6810	2.2395	0.000639269	0.037292583	KIR2DL3	NP_055326.3	19	19q13.42b
ILMN_1691717	-2.9327	0.8552	0.000640917	0.037292583	RHBDF2	NP_001005498.2	17	17q25.1d
ILMN_1717781	-1.9505	0.5684	0.000634519	0.037292583	VPS11	NP_068375.3	11	11q23.3e
ILMN_1901009	1.8175	0.5298	0.000637337	0.037292583			7	
ILMN_2065606	-1.7011	0.4960	0.000640424	0.037292583	TOMM40L	NP_115550.2	1	1q23.3a
ILMN_2382500	-1.7504	0.5104	0.000640779	0.037292583	MS4A14	NP_001073160.1	11	11q12.2a
ILMN_1677827	-2.6315	0.7679	0.00064569	0.037467686	TLR7	NP_057646.1	X	Xp22.2
ILMN_1774062	-1.8475	0.5395	0.000652308	0.037732698	SLC25A5	NP_001143.1	X	Xq24c
ILMN_1858762	1.5793	0.4613	0.000653811	0.037732698			2	
ILMN_1777118	-1.9585	0.5728	0.000664826	0.038264448	INTS9	NP_060720.1	8	8p21.1c-p21.1b
ILMN_1784535	1.7582	0.5149	0.000675438	0.038770119	LOC389137	XP_371655.3	3	3q11.2b
ILMN_1661484	-1.8744	0.5492	0.000679811	0.038915952	ZBTB45	NP_116181.1	19	19q13.43c
ILMN_1687430	-1.7113	0.5018	0.000686466	0.039191287	EIF2B4	NP_056451.2	2	2p23.3a
ILMN_3240740	-2.0011	0.5871	0.000690741	0.039329666	EIF3L	NP_057175.1	22	22q13.1a-q13.1b
ILMN_2367113	2.0380	0.5986	0.000700205	0.039761887	CASP6	NP_116787.1	4	4q25c
ILMN_1701857	1.6540	0.4860	0.000703813	0.039860206	LOC644373	XP_937247.1	1	
ILMN_1691341	4.2165	1.2397	0.000708835	0.040037866	IL7R	XP_942460.1		5p13.2c
ILMN_1663132	-2.2508	0.6621	0.000713225	0.040089369	ADCK2	NP_443085.2	7	7q34c
ILMN_1700507	-2.2024	0.6479	0.000713522	0.040089369	CLASP1	NP_056097.1	2	2q14.2e-q14.3a
ILMN_1653856	-2.3331	0.6883	0.000738767	0.040638982	STS-1	NP_116262.2	11	11q24.1b
ILMN_1663113	-2.2314	0.6583	0.000739281	0.040638982	TTLL12	NP_055955.1	22	22q13.2c
ILMN_1678729	-1.9951	0.5881	0.000731974	0.040638982	SIL1	NP_001032722.1	5	5q31.2d
ILMN_1712305	-3.2429	0.9568	0.000739472	0.040638982	CYBRD1	NP_079119.2	2	2q31.1c
ILMN_1714952	1.5985	0.4718	0.000743793	0.040638982	ZNF703	XP_001129663.1		8p12a
ILMN_1742187	1.9893	0.5865	0.000732909	0.040638982	MAN1A1	NP_005898.2	6	6q22.31a
ILMN_1746276	-2.3012	0.6780	0.000726998	0.040638982	EPC1	NP_079485.1	10	10p11.22b
ILMN_1777740	-2.1105	0.6221	0.000731812	0.040638982	C8orf55	NP_057731.1	8	8q24.3e
ILMN_1821270	2.1457	0.6334	0.000743745	0.040638982			14	
ILMN_1906423	1.5127	0.4465	0.000743316	0.040638982			2	

ProbeID	Effect	SD	*p* value	*p* value FDR	Gene	Protein_Product	Chr	Cytoband
ILMN_3247533	2.6712	0.7885	0.000744353	0.040638982	LOC100133840	XP_001714436.1		
ILMN_1779257	-1.6974	0.5013	0.000749137	0.040795325	CD40	NP_001241.1	20	20q13.12b
ILMN_1701173	-1.9660	0.5809	0.000752336	0.040864759	KCNK6	NP_004814.1	19	19q13.2a
ILMN_1742507	1.6788	0.4962	0.000755292	0.040920615	LRRC45	NP_659436.1	17	17q25.3g
ILMN_2302118	-2.4412	0.7217	0.000758304	0.040979274	CCDC50	NP_777568.1	3	3q28d
ILMN_1757347	-1.8729	0.5542	0.000766582	0.041012679	C22orf9	NP_056079.1	22	22q13.31b
ILMN_1797298	-1.8256	0.5402	0.000766646	0.041012679	ARMC7	NP_078861.1	17	17q25.1c
ILMN_1798288	-2.0622	0.6099	0.000762438	0.041012679	MOBKL2C	NP_660322.2	1	1p33d
ILMN_1811049	-3.2655	0.9663	0.000766528	0.041012679	POU2AF1	NP_006226.1	11	11q23.1b
ILMN_1698578	1.5425	0.4568	0.000773679	0.041012985	LOC440978	XP_496659.2	3	
ILMN_1721383	1.5114	0.4475	0.000771268	0.041012985	TWIST2	XP_950696.1		2q37.3c
ILMN_1722738	-2.3288	0.6897	0.000774376	0.041012985	ROGDI	NP_078865.1	16	16p13.3b
ILMN_2243553	-2.3587	0.6984	0.000772182	0.041012985	ZNF275	NP_001073954.1	X	Xq28f
ILMN_1795063	-2.1297	0.6312	0.000781231	0.041273112	ZADH2	NP_787103.1	18	18q22.3d
ILMN_1813840	-2.0396	0.6047	0.000784811	0.041359319	AOF2	NP_055828.2	1	1p36.12a
ILMN_1709334	-2.0267	0.6013	0.000792168	0.041643734	TM9SF1	NP_006396.2	14	14q12a
ILMN_2284591	-1.9793	0.5875	0.000795015	0.041690195	OPA3	NP_079412.1	19	19q13.32a
ILMN_1661264	-2.0959	0.6224	0.000800545	0.041701655	SHMT2	NP_005403.2	12	12q13.3b
ILMN_2058795	-2.4777	0.7357	0.000798799	0.041701655	PGCP	NP_057218.1	8	8q22.1d-q22.1e
ILMN_3228688	-3.9243	1.1655	0.000801124	0.041701655	LOC730415	XP_001720887.1		
ILMN_1696933	-1.4398	0.4281	0.000811439	0.042135298	NLRP3	NP_004886.3	1	1q44e
ILMN_1697220	-1.9550	0.5815	0.000816028	0.042270239	NT5E	NP_002517.1	6	6q14.3c
ILMN_1668312	-1.5537	0.4628	0.000830762	0.042690407	SLC2A9	NP_001001290.1	4	4p16.1b
ILMN_1767470	-2.5217	0.7512	0.000831038	0.042690407	SCPEP1	NP_067639.1	17	17q22c
ILMN_2310896	-2.5805	0.7688	0.00083218	0.042690407	NLRP3	NP_004886.3	1	1q44e
ILMN_2311826	-1.5692	0.4674	0.000828842	0.042690407	USP6NL	NP_001073960.1	10	10p14a
ILMN_1682761	-2.4123	0.7189	0.000834997	0.042714448	C17orf87	NP_996986.1	17	17p13.2b
ILMN_1756928	-3.2682	0.9741	0.000836671	0.042714448	RTN1	NP_066959.1	14	14q23.1b-q23.1c
ILMN_1892608	1.5845	0.4731	0.000853954	0.043492267			X	
ILMN_1697268	-3.0111	0.9001	0.000866088	0.044004739	EMILIN2	NP_114437.2	18	18p11.32a-p11.31e
ILMN_1663379	-2.0289	0.6080	0.000892429	0.044983286	FBXL15	NP_077302.2	10	10q24.32b
ILMN_1684040	3.2169	0.9641	0.000893122	0.044983286	C6orf190	NP_001010923.1	6	6q22.33b

R39 R38 R37 R36 R35 R34 R33 R32 R31 R30 R29 R28 R27 R26 R25 R24 R23 R22 R21 R20 R19 R18 R17 R16 R15 R14 R13 R12 R11 R10 R9 R8 R7 R6 R5 R4 R3 R2 R1

ProbeID	Effect	SD	*p* value	*p* value FDR	Gene	Protein_Product	Chr	Cytoband
ILMN_1691071	-5.2848	1.5840	0.00089382	0.044983286	FCRLA	NP_116127.3	1	1q23.3b
ILMN_2197247	-1.6956	0.5079	0.000888023	0.044983286	POLR3A	NP_008986.2	10	10q22.3c
ILMN_1894166	1.6314	0.4893	0.000900238	0.045199208				
ILMN_1656082	-2.3334	0.7002	0.000906162	0.045389318	H3F3A	NP_002098.1	1	1q42.12c
ILMN_2354191	2.4821	0.7454	0.000914481	0.045698241	CD8B	NP_004922.1	2	2p11.2e
ILMN_1674160	2.6102	0.7841	0.000917573	0.045745086	BIN1	NP_647596.1	2	2q14.3d
ILMN_1768598	-1.9102	0.5740	0.00092109	0.045763922	LAIR1	NP_068352.1	19	19q13.42a
ILMN_1811264	-1.9224	0.5777	0.00092226	0.045763922	C15orf57	NP_443081.1	15	15q15.1b
ILMN_1697218	-2.0392	0.6132	0.000928826	0.045953462	MED22	NP_852468.1	9	9q34.2a
ILMN_1742705	-1.9389	0.5831	0.000930407	0.045953462	SLC39A11	NP_631916.1	17	17q24.3c-q25.1a
ILMN_1707943	-1.9806	0.5959	0.000934289	0.045969338	C19orf39	NP_787067.1	19	19p13.2b
ILMN_3201937	-1.5381	0.4628	0.000935058	0.045969338	LOC645381		X	Xq11.1c
ILMN_1690454	-1.9485	0.5873	0.000953363	0.046637325	C3orf54	NP_976248.1	3	3p21.31c
ILMN_1692100	-1.4446	0.4353	0.000951278	0.046637325	ZNF35	NP_003411.3	3	3p21.32a
ILMN_1778709	1.8975	0.5720	0.000955233	0.046637325	PICALM	NP_009097.2	11	11q14.2a
ILMN_1762312	-2.4816	0.7488	0.000966202	0.047064673	FOXRED1	NP_060017.1	11	11q24.2c
ILMN_2086890	-1.6410	0.4953	0.000968904	0.047088274	ANGPT1	NP_001137.2	8	8q23.1b-q23.1c
ILMN_1717029	1.9539	0.5901	0.000975966	0.047122974	FLJ33590	NP_776182.1	2	2q37.3g
ILMN_1746686	-1.7395	0.5253	0.000976616	0.047122974	POLR1C	NP_976035.1	6	6p21.1c
ILMN_2053490	-1.6797	0.5074	0.000978493	0.047122974	FAM53B	NP_055476.3	10	10q26.13d
ILMN_2388142	-1.8669	0.5639	0.000977833	0.047122974	CD99L2	NP_604394.1	X	Xq28c
ILMN_1781752	-1.8092	0.5481	0.001013832	0.048714387	CLEC16A	NP_056041.1	16	16p13.13c
ILMN_2328224	-1.8853	0.5717	0.001024176	0.049100348	MADD	NP_569829.1	11	11p11.2b
ILMN_1727709	-3.2402	0.9831	0.001030643	0.049260545	GPBAR1	NP_733800.1	2	2q35e
ILMN_1784651	-1.7552	0.5326	0.001032157	0.049260545	NAGA	NP_000253.1	22	22q13.2b
ILMN_1662707	1.6659	0.5056	0.001035399	0.049304483	SEC24B	NP_006314.2	4	4q25c
ILMN_1670539	-2.4662	0.7489	0.001041001	0.049460371	LOC92017	NP_001009607.1	16	16p13.13a-p13.12b
ILMN_1687921	-2.1841	0.6635	0.001044983	0.049538725	JMJD8	NP_001005920.2	16	16p13.3f
ILMN_1724863	-2.2403	0.6812	0.001056305	0.049963924	TICAM1	NP_891549.1	19	19p13.3d

Supplementary Table 7. Differentially expressed probes associated with methylation at cg14023999

probeID	Effect	SD	*p* value	*p* value FDR	Gene	Protein_Product	Chr	Cytoband
ILMN_1743032	-3.1308	0.5973	2.11E-07	0.00448515	CTSS	NP_004070.3	1	1q21.2c
ILMN_1785095	3.2162	0.6340	5.03E-07	0.005341343	ATP6V0E2	NP_660265.2	7	7q36.1b

Supplementary Table 8. Results from Pathway analysis of 422 differential expressed genes associated with differential methylation at cg12325605, Significant pathways are highlighted in bold. Significance is based on a Bonferroni count of 1774 tested pathways

Reactome pathways	Homo sapiens REFLIST (21002)	Observed	Expected	over/under	Fold Enrichment	*P* value
Phosphorylation of CD3 and TCR zeta chains (R-HSA-202427)	26	6	0.45	+	13.46	1.27E-02
PD-1 signaling (R-HSA-389948)	27	6	0.46	+	12.96	1.57E-02
Antigen activates B Cell Receptor (BCR) leading to generation of second messengers (R-HSA-983695)	62	8	1.06	+	7.53	2.64E-02
Immunoregulatory interactions between a Lymphoid and a non-Lymphoid cell (R-HSA-198933)	**156**	**17**	**2.67**	**+**	**6.36**	**5.61E-06**
Adaptive Immune System (R-HSA-1280218)	**819**	**42**	**14.04**	**+**	**2.99**	**7.37E-07**
Immune System (R-HSA-168256)	**1594**	**70**	**27.32**	**+**	**2.56**	**6.93E-10**
Innate Immune System (R-HSA-168249)	784	34	13.44	+	2.53	1.77E-03
Developmental Biology (R-HSA-1266738)	798	31	13.68	+	2.27	4.56E-02
Unclassified (UNCLASSIFIED)	11733	148	201.12	-	0.74	0.00E+00

Supplementary Table 9. Results from Pathway analysis of all (428) differential expressed genes, Significance is based on a Bonferroni count of 1774 tested pathways. Significant pathways are highlighted in bold

Reactome pathways	Homo sapiens REFLIST (21002)	Genes observed	Genes expected	over/ under	Fold Enrichment	*p* value
Phosphorylation of CD3 and TCR zeta chains (R-HSA-202427)	26	6	0.45	+	13.39	1.31×10-02
PD-1 signaling (R-HSA-389948)	27	6	0.47	+	12.89	1.62×10-02
Antigen activates B Cell Receptor (BCR) leading to generation of second messengers (R-HSA-983695)	62	8	1.07	+	7.49	2.75×10-02
Immunoregulatory interactions between a Lymphoid and a non-Lymphoid cell (R-HSA-198933)	**156**	**17**	**2.69**	**+**	**6.32**	**6.09×10-06**
Adaptive Immune System (R-HSA-1280218)	**819**	**43**	**14.12**	**+**	**3.05**	**2.61×10-07**
Immune System (R-HSA-168256)	**1594**	**72**	**27.47**	**+**	**2.62**	**1.01×10-10**
Innate Immune System (R-HSA-168249)	784	35	13.51	+	2.59	7.18×10-04
Unclassified (UNCLASSIFIED)	11733	148	202.24	-	0.73	0.00×10+00

CHAPTER 5

Physical health factors for depression

CHAPTER 5.1

Multiple episodes of late-life depression and cognitive decline

Jovanova Olivera, Wolters Frank, Ikram Arfan, Tiemeier Henning, Schmitz Norbert.

Under review Jama Neurology

ABSTRACT

Importance: Population-based studies consistently reported a cross-sectional but no prospective association between depression and cognitive decline. Clinical studies suggest that the number of depressive episodes a patient experienced, may explain the association between depression and cognitive decline. To date, no population-based study has investigated the prospective relation between single-episode and multiple-episode depression and subsequent cognitive decline.

Objectives: To determine the association between depressive episodes and cognitive change over a 12-year period.

Design, Settings and Participants: This study was embedded in the Rotterdam Study, a prospective population-based cohort of persons aged ≥ 55 years. 6 221 non-demented participants, with a mean (*SD*) age of 68.6 (8.1) years were enrolled between 1997 and 2001 and were followed for cognitive functioning until 2012.

Main outcomes and Measures: Depressive symptoms questionnaire and general practitioner records were used to detect 751 (12.1%) persons with single-episode and 350 (5.6%) persons with multiple-episode among 6 221 individuals (depression follow-up 1993 - 2004). Cognitive decline was determined during three consecutive examination rounds with a test-battery including a letter-digit-substitution test, Stroop test, verbal fluency test, and mini-mental state examination (cognition follow-up 1997 - 2011). The association between depressive episodes and cognitive decline was modelled with a latent growth curve adjusted for sociodemographic and lifestyle factors, and comorbidities.

Results: Severity of depressive symptoms at baseline was not related to cognitive decline during on average 12 years follow-up. Likewise, participants who experienced a single-episode depression had no excess decrease in cognitive functioning during follow-up (beta, -0.088; *S.E.*, 0.12; p = 0.45). In contrast, participants who experienced multiple-episodes of depression had a stronger decrease in the cognitive functioning than persons without depression (beta, -0.656; *S.E.*, 0.16; $p < 0.001$).

Conclusions and relevance: In older adults, multiple-episode but not single-episode depression is associated with accelerated cognitive decline. This finding is in line with clinical studies showing that chronic psychiatric disease may have impact on brain resilience. Elderly persons suffering recurrent depression could be considered a target group for prevention and management of cognitive impairment.

INTRODUCTION

Depression and cognitive decline are common chronic and disabling conditions affecting older people.[1,2] Non-demented elderly diagnosed with depression often report cognitive complaints such as concentration loss, memory impairment, word-finding difficulties, or limitations in daily activities.[2] Depression and cognitive impairment often co-occur[1,3-9] and cognitive complaints frequently persist even after the remission of the depressive episode.[2,7,10,11] Yet, whether depression-associated pathology contributes to the subsequent cognitive decline is uncertain.[2]

While it is often speculated that depression can negatively affect neurocognitive functioning, it is unclear whether depression truly contributes to subsequent cognitive decline.[12] A substantial number of clinical and population-based studies addressed the prospective association between depression and cognitive complaints, but findings have been inconsistent.[8,9,12-25] Clinical studies repeatedly reported greater cognitive decline following depression.[12,19,20,22] In contrast, population-based studies generally found no relation between depressive symptoms assessed at baseline and subsequent change in cognitive function.[8,15-18,21] This discrepancy between epidemiological and clinical findings may relate to the variation in severity of depression between settings.

In order to reconcile these discrepant findings researchers have more recently studied persistent depressive symptoms in population-based studies.[9,22,26] Persistent depression was defined as symptom ratings repeatedly assessed above a certain threshold. These studies showed that the persistence of symptoms across a 2 - 5 year follow-up period was paralleled by a decline in cognitive functioning.[9,22,26] This essentially cross-sectional design of two repeatedly assessed measures suggests that multiple episodes[23] of depressive symptoms rather than a single episode may be associated with cognitive decline. Against this background, a recent review argued that the relationship between severity of depression and cognitive disability is poorly understood and understudied.[12]

Longitudinal population-based studies that assess depressive episodes over a long period prior to cognitive decline are lacking.[12] We aimed to investigate in an ongoing population-based study whether the frequency of exposure to depressive episodes in non-demented elderly is associated with cognitive decline if followed for many years. We hypothesized that persons with multiple episodes of depression have a faster cognitive decline over a 12 years follow-up period than individuals that experienced none or a single episode of depression.

METHODS

Study population

This study was embedded in the Rotterdam Study, a prospective population-based cohort that started in 1990 among inhabitants aged ≥ 55 years living in the district Ommoord, Rotterdam, the Netherlands.[27] The medical ethics committee approved the study according to the Population Study Act of the Dutch Ministry of Health (Wet Bevolkingsonderzoek). Written informed consent was obtained from all participants.

The first cohort of the Rotterdam Study included 7 983 participants and was expanded with an additional 3 011 participants, who had reached age of 55 years or moved into the study area, in the year 2000. The details of the objectives and the study design of the Rotterdam Study have been previously described elsewhere.[27] In brief, all participants underwent home interview and an extensive set of examinations in a research centre during follow-up examination rounds that were conducted on average every 3 - 4 years.[27] The current study included participants assessed for depressive episodes and cognitive functioning, during the 1997 - 1999 examination of the first cohort, and the 2000 - 2001 examination of the second cohort; this defines the baseline of the present study (Supplement 1 and Figure 1).

7 808 persons participated in this baseline examination of this study (Supplement 1). Of these, 7 533 participants were screened for depressive symptoms (96.5% response rate). We excluded 1 269 participants, who did not have full cognitive testing (at least two completed cognitive tests were required). The subjects with no cognitive assessment were on average older (74.9 versus 69.0 years), more likely to be female (66.9% versus 56.7%) and had experienced more depression (20% versus 18%) than those participants with cognitive assessment. Additionally, we excluded 92 persons with prevalent dementia and 8 persons with bipolar disorder. This resulted in a cohort of 6 221 persons eligible for the study.

Table 1. Study population by depressive episodes status (*N* = 6 221)

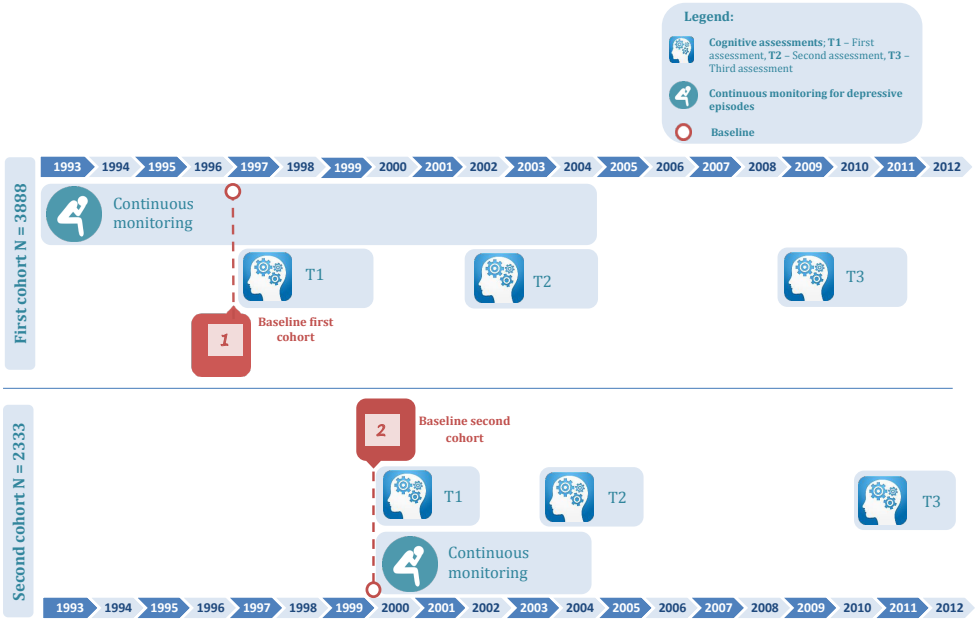
	Depressive episodes		
	No depression ***N* = 5120**	***Single episode*** ***N* = 751**	***Multiple episodes*** ***N* = 350**
Age, years	68.4 (8)	70.1 (8.5)	69.6 (7.9)
Sex, female	2757 (53.8)	495 (65.9)	270 (77.1)
Highest education attained			
Primary	614 (12.0)	122 (16.2)	57 (16.3)
Lower/intermediate	2203 (43.0)	340 (45.3)	159 (45.4)
Vocational	1566 (30.6)	196 (26.1)	111 (31.7)
Higher	710 (13.9)	86 (11.5)	20 (5.7)
Smoking habits			
Never smoker	1642 (32.1)	270 (36)	119 (34)
Current smoker	850 (16.6)	157 (20.9)	82 (23.4)
Past smoker	2628 (51.3)	324 (43.1)	149 (42.6)
Alcohol consumption			
No consumption	484 (9.5)	88 (11.7)	51 (14.6)
Current consumers	4345 (84.9)	589 (78.54)	257 (73.4)
Past consumers	291 (5.7)	73 (9.7)	42 (12)
Body mass index, kg/m²	26.9 (3.9)	27.1 (4)	27.5 (4.5)
IADL*	15.6 (7.3)	15.3 (7.5)	15.3 (7.5)
Cardiovascular medication	2088 (40.8)	362 (48.2)	169 (48.3)
Psychotropic medication	631 (12.3)	217 (28.9)	156 (44.6)
Cancer	386 (7.5)	54 (7.2)	38 (10.9)
Diabetes mellitus	252 (4.9)	54 (7.2)	20 (5.2)
Myocardial infarctions	212 (4.1)	34 (4.5)	15 (4.3)
Stroke	152 (3.0)	38 (5.1)	20 (5.7)

Descriptive characteristics of study population were assessed at baseline and present the non-imputed data. The values represent counts (percentage) or means (standard deviations).***IADL** – instrumental activity of daily living.

Assessment of depressive episodes

Participants were monitored for the occurrence of depressive episodes (including: depressive symptoms, depressive syndromes, dysthymia and major depressive disorders) during an average of 8 years. The time-lines for assessment of depressive episodes are illustrated in detail in Figure 1. The occurrence of depressive episodes was assessed continuously four years prior to baseline and before the first re-examination of the first cohort (first cohort 1993 - 2004). Likewise, depressive episodes were assessed prior to the first re-examination of the second cohort (second cohort 1999 - 2004).

Figure 1. Diagram of the timeline



In order to enhance case-findings of depressive episodes during the interval between examinations and at examination rounds the information was collected with different assessment methods.[28] First, we used information from the screening for depressive symptoms, performed at each examination. All participants were screened for depression with the validated Dutch versions of the Centre for Epidemiologic Studies Depression (CES-D) scale.[29] The 20-item CES-D measure self-reported frequency of experienced depressive symptoms. Participants could score from 0 up to 60. A cut-off score of ≥ 16 defined participants with clinically relevant depressive symptoms, i.e., a depressive episode.[29]

To identify the depressive episodes that occurred and may have remitted in the intervals between follow-up examination rounds, we continuously monitored data on the occurrence of depressive episodes from the general practitioners (GP) records. The Netherlands have a primary health care system, thus all specialist must report to the GP. The GP records are elaborate and include diagnoses obtained from mental-health professionals substantiated by DSM-classification.[28] The occurrences inferred from the different data sources were consistent.[28]

With this information, participants were either classified as persons with depressive episodes (single-episode (*N* = 751) and multiple-episode (*N* = 350) depression) and persons with no episodes of depression (*N* = 5 120). Only depressive episodes occurring before the first re-assessment of cognitive functioning were included to minimize the overlap between depressive episode assessment and cognitive decline (Figure 1). Depression often has a chronic course and

only 60% of the persons suffering recover in the first year, 70 - 80% recover during the second year but remission of depression is not always well documented.[28] Therefore, to distinguish between episodes we assumed that one depressive episode may last for a maximum of two years.

Cognitive functioning

Cognitive function was assessed during three consecutive examination rounds (Figure 1) with a battery of cognitive tests that included the Mini Mental State Examination (MMSE), Stroop test, a letter-digit substitution task (LDST), and a verbal fluency test (VF). On the MMSE, participants could score from 0 to 30 points.[30] The 40-items Stroop test[31] scored time needed to name the ink colour of incongruously printed name of colours. The LDST[32] assessed the total number of correct digits assigned to letters according to a code key during 60 seconds. VF was assessed with the animal naming test.[33] The total number of named animals within one minute were scored. Lower scores indicated worse performance on all tests, except for the Stroop test.

We used principal component analyses of the four cognitive test scores to derive a more robust summary test score (G-factor). The G-factor was calculated at each wave using a multiple indicator measurement model. The model was constrained by requiring corresponding loadings across waves to be equal in the assumption of measurement invariance (Supplement 3). For each participant, z-scores were calculated for the G-factor and each test separately by dividing the difference between individual test score and mean test score by the standard deviation.

Table 2. Descriptive of cognitive performance tests measured over time for depression

Depression status	Cognitive test					
	T1* (1997-1999)		T2* (2002-2004)		T3* (2009-2011)	
	N	Mean (SD)	*N*	Mean (SD)	*N*	Mean (SD)
	Letter-digit substitution task					
No depression	5036	28 (7.2)	3797	27.4 (7.5)	2480	26.5 (8.2)
Single episode	739	26.2 (7.5)	485	26.1 (7.5)	262	26.5 (9.4)
Multiple episodes	341	26.1 (7)	249	24.4 (7.4)	129	24.4 (9.3)
	Verbal fluency test					
No depression	5094	21.4 (5.7)	3888	20.7 (5.5)	2560	20.8 (6.3)
Single episode	747	20.8 (6.1)	502	20.5 (5.5)	274	20.9 (7.5)
Multiple episodes	347	20.6 (5.8)	261	19.8 (5.4)	138	19 (5.2)
	Stroop test					
No depression	4970	58.0 (27.3)	3 523	60.1 (29.9)	2466	61.8 (31.4)
Single episode	730	64.8 (34.9)	455	65.2 (37.8)	265	61.4 (29.4)
Multiple episodes	340	62.9 (29.7)	231	70.6 (38.5)	136	67.9 (40.4)
	Mini–mental state examination					
No depression	5119	27.8 (1.8)	3998	27.6 (2.2)	2718	27.4 (3)
Single episode	751	27.5 (2.1)	546	27.2 (2.7)	294	27.1 (3.1)
Multiple episodes	350	27.5 (2)	287	26.9 (3.1)	159	26.8 (2.9)

T1: time point one; ***T2:*** time point two; ***T3:*** time point three.

Covariates

Data on socio-demographic background and life-style factors were assessed at baseline using standardized interview questions. Highest level of education was classified in four categories: low, intermediate-low, intermediate-high, and high education. Smoking and alcohol consumption habits were classifed as current, past and non-consumers. Presence of common chronic diseases (stroke, diabetes mellitus, myocardial infarction and cancer) was based on self-report, examination, medical record information, and drug utilization. Psychotropic and cardiovascular medications were self-reported during interview, and verifed by medication prescriptions. Height and weight were measured and body mass index (BMI) was calculated as weight divided by height squared (kg/m2). To assess physical self-maintenance, we used the instrumental-Activities of daily living scale (IADL).

DATA ANALYSIS

A descriptive analysis provided an overview of the study sample at baseline. Also, we assessed crude mean differences in cognitive test scores between persons with no depression, single-episode, and multiple-episode depression at each of the three examination rounds, by using repeated measures analysis of variance.

In order to analyse the temporal relationship between depression and cognitive change we used a latent growth curve model (LGCM) approach. This approach models repeated measures into latent variables with a random intercept and random slope, allowing individual cases to have individual trajectories of change over time. This methodology permitted us to effectively study the within-person (intra-individual) and the between-person (inter-individual) change in cognitive functioning over time.

In the present study, a trajectory of cognition was modelled from the G-factor of cognitive functioning (constructed as described above). The trajectory was estimated by using an intercept (the average level of baseline scores) and a slope (linear rate of change over time). The time points for the slope growth factor were fixed at 0, 1, and 2 to define a linear growth model and the factor loadings were held constant across time to create an LGM process with latent slope and intercept.
This study examined the association of depression with the change rates (slope) of the G-factor of cognition measured at 3 time points (Figure 1). First, we evaluated the course of cognition over time while controlling for sex, age, education and cohort (Supplement 3). Second, we conditioned cognitive performance on baseline depressive symptoms and depressive episodes

status while controlling for sex, age, education and cohort. Third, we constructed a fully adjusted model were we additionally controlled the conditional models for other potential confounders; including various comorbidities, life-style, and demographic factors (Table 3).
Parameters were estimated using a maximum likelihood estimation, including estimating a random effect and estimator with robust standard errors. The general goodness of fit of each model was evaluated using the Comparative Fit Index (CFI) and the Root Mean Square Error of Approximation (RMSEA). Conventional rules suggest that a CFI ≥ 0.90 indicates adequate fit and CFI ≥ 0.95 indicates excellent fit; a RMSEA ≤ 0.08 indicates adequate fit (≤ 0.05 indicates excellent fit).[34,35]

Multiple imputations were used in to account for missing data on potential confounding variables (missing values: alcohol consumption 0.01%, BMI 0.5%, IADL 8.1%, cardiovascular and psycotropic medication 4.1%, cancer 0.01%, diabetes mellitus 0.2%, myocardial infarction 8%, education 0.6%,). All analyses were run in the complete case and ten imputed data sets. In this manuscript we present results from the imputed data. Analyses were conducted in M-Plus (Version 7.3).

In a sensitivity analyses we conducted the analyses for the initial cohort and the extension cohort separately, to account for the slight discrepancy in the assessment of past depression (Supplement 4).

RESULTS

Table 1 provides the demographics and clinical characteristics of the participants by depression status at baseline. Participants with single-episode (70.1 ± 8.5 years) and with multiple-episode (69.6 ± 7.9 years) were on average older than participants without depression (68.4 ± 8 years). Table 2 presents the mean cognitive status of the three depression groups at each assessment wave. The overall decline in all cognitive tests across follow-up in all three groups is shown. For example, persons with multiple episodes of depression had a lower MMSE score at all time points (27.5 at baseline and 26.8 at the end of follow-up) than those without depression (27.8 at baseline and 27.4 at the end of follow-up).

We first tested the association of depressive symptoms at baseline studied continuously with decline of cognitive functioning across follow-up (Supplement 2). Although persons with more depressive symptoms had worse cognitive function at baseline (see intercept, Supplement 2), we found no prospective association between depressive symptoms and cognitive decline (beta [*B*], -0.006; standard error [*S.E.*], 0.01; $p = 0.28$; see slope).

Second, we investigated the association between frequency of depressive episodes and cognitive functioning (Table 3). All latent growth models had an excellent fit, with no CFI lower than 0.06 and no RMCEA higher than 0.04. In line with the analysis of continuous depressive symptoms, we observed that a single or multiple episodes of depression were related to poor cognitive functioning at baseline (Table 3, see intercept). A single episode of depression was not associated with subsequent cognitive decline (*B*, -0.09; *S.E.*, 0.12; $p = 0.45$; see slope), but persons with multiple-episode depression had faster cognitive decline than those without depression (*B*, -0.66; *S.E.*, 0.2; $p < 0.001$). This finding was not significantly changed after adjustment for various potential confounders (Table 3).

Table 3. Initial level and rate of change of cognitive functioning and single or multiple-episode depression using the Latent growth curve modelling (*N* = 6 221)

Conditional model	Intercept			Slope		
	Estimate	**S.E.**	***p* value**	**Estimate**	**S.E.**	***p* value**
Depressive episodes						
No depression		reference			reference	
Single episode	-0.911	0.20	<0.001	-0.088	0.12	0.45
Multiple episodes	-1.088	0.28	<0.001	-0.645	0.16	<0.001
Sex (female)	0.680	0.13	<0.001	0.198	0.07	0.007
Age	-2.945	0.08	<0.001	-1.045	0.05	<0.001
Cohort	-1.002	0.15	<0.001	0.075	0.08	0.36
Education	2.043	0.08	<0.001	0.091	0.04	0.034

Model fit indices: RMSEA* (95% CI) = 0.038 (0.036; 0.040); CFI** = 0.967

Fully adjusted model	Intercept			Slope		
	Estimate	**S.E.**	***p* value**	**Estimate**	**S.E.**	***p* value**
Depressive episodes						
No depression		reference			reference	
Single episode	-0.581	0.20	0.003	-0.088	0.12	0.45
Multiple episodes	-0.545	0.28	0.053	-0.656	0.16	<0.001
Sex (female)	1.034	0.14	<0.001	0.153	0.08	0.051
Age	-2.678	0.08	<0.001	-1.023	0.05	<0.001
Cohort	-0.914	0.15	<0.001	0.094	0.08	0.25
Education	1.958	0.08	<0.001	0.090	0.04	0.036
Body mass index	-0.022	0.02	0.18	-0.012	0.01	0.20
Cancer	0.506	0.24	0.033	-0.121	0.14	0.38
Diabetes mellitus	-0.855	0.29	0.003	-0.541	0.17	0.002
Stroke	-1.611	0.36	<0.001	-0.428	0.26	0.09
Myocardial infarction	0.258	0.33	0.47	-0.003	0.19	0.99
Psychotropic therapy	-0.661	0.18	<0.001	0.061	0.11	0.57
Cardiovascular therapy	-0.152	0.15	0.30	-0.173	0.08	0.030
IADL***	1.318	0.12	<0.001	-0.063	0.08	0.44
Smoking habits	0.327	0.08	<0.001	-0.033	0.04	0.42
Alcohol consumption	0.050	0.16	0.75	-0.047	0.10	0.62

Model fit indices: RMSEA* (95% CI) = 0.028 (0.027; 0.030); CFI** = 0.965

* RMSEA: Root Mean Square Error of Approximation **CFI: Comparative Fit Index;***IADL instrumental activity of daily living.

DISCUSSION

In this population-based cohort of older adults we investigated the impact of late-life depression on the change in cognitive functioning. We observed that increased baseline depressive symptoms (continuously assessed) did not predict cognitive decline over a 12 year follow-up. However, persons who had experienced multiple-episodes of depression were prone to faster cognitive decline during follow-up compared to those who experienced only one or no depressive episode. These results emphasise that the persistence or recurrence rather than the severity of depression at one time point relates to cognitive decline.

Depression and cognitive decline are two conditions consistently co-occurring more often than expected. Impaired cognition is an important symptom of depression and it is evident that clinical and community-dwelling individuals with depression have impaired cognitive functioning compared to non-depressed individuals.[13,36,37]

In contrast to the cross-sectional associations, the prospective association of depression and cognitive decline has been a challenge for both clinical and population based studies. Clinical studies mostly report prospective association in patients with depression[7,12], while population-based studies often did not find a prospective association.[8,15,17,19] Consistent with previously published population-based studies, we also observed that depressive symptoms at baseline did not predict cognitive decline.[8,14-18,20] The discordance in findings between clinical and population-based studies is probably not a power problem, since these cohorts were sufficiently large and include many individuals with depressive symptoms. Rather, patients with more severe symptoms are typically selected in clinical studies due to referral patterns. Thus, in the general population studies persons screened positively for depressive symptoms may not have severe or chronic enough depressive problems to impact the cognitive functioning.

We continuously monitored persons over longer periods of time for the occurrence of depressive episodes and observed that not a single episode of depression but recurrent depression was associated with a stronger decline of cognitive functioning over the longer term. A few previously published studies support our results and suggest that persistent symptoms of depression are related to concurrent cognitive decline.[9,21,26] For example, in the EVA study, the authors observed that more depressive symptoms at baseline were associated with lower MMSE scores only if the depressive symptoms persisted during the entire follow-up of 4 years.[26] Also Kohler et al, reported a similar association between the co-occurring persistent depressive symptoms and the decline in cognition during 6 years follow-up.[21] However, in these studies persistent depressive symptoms were defined during the same period as the cognitive functioning. In the present study, there is only a short overlap in the assessments periods of repeated depression and cognitive functioning

assessments. Hence, the longitudinal association reflects the decline of the cognitive function after the assessment of depression.

How does multiple-episode depression affect cognitive functioning?

The neuropsychological consequences of late-life depression are unclear. Various potential mechanisms may explain why multiple-episode but not single-episode depression is a risk factor for long-term cognitive decline. First, elderly people experience chronic stress due to age-related life-style, environmental, and health changes. Exposure to such a chronic stress increases the likelihood of developing mental disorders with a more chronic/recurrent course. Chronic stress can activate the hypothalamus-pituitary–adrenal (HPA) axis, which responds with cortisol release. This depression or stress-related prolonged hypercortisolemia is known to have adverse neurotoxic effect such as hippocampal neuronal degeneration and loss.[38] Indeed, hyper-activation of the HPA axis with consequent hippocampal and hypothalamic atrophy were associated with both recurrent depression and cognitive decline.[9,22,38-40] Importantly, other biological pathways also link depression and cognitive decline such as metabolic changes or inflammation.[41,42] Studies have shown that severe or repeated depression is related to subsequent increase of interleukin-6 and other inflammatory markers.[43]

Second, depression and cognitive decline might reflect one shared underlying pathophysiological mechanism. For example, cerebrovascular disease causes disruption of prefrontal systems that can give rise to both mood and executive functions and thereby explain the relationship between depression and cognitive decline.[44] This could be due to microvascular dysfunction which underlies amyloid-β formation, and white-matter changes.[23] Indeed, a vascular hypothesis is supported by neuroimaging data showing that cerebrovascular lesions were related to both persistent depression and increased risk of dementia.[45] Importantly, vascular small vessel disease can be both a cause and a consequence of depression, thus the two mechanisms above need not be exclusive. In this study we carefully adjusted for vascular morbidity at baseline including; stroke, diabetes mellitus, and myocardial infarction. Thus confounding by pre-existing vascular diseases is not very likely but this does not control for vascular change as a consequence of depression.

Limitations

We continuously monitored participants, both men and women, for depressive episodes for almost 12 years. This large population study combines a comprehensive follow-up of depression and detailed cognitive functioning assessments. Yet, several limitations need to be addressed. First, different aspects of severity of depression such as age of first onset, number and duration of the depressive episodes, duration of the illness, and treatment resistance might be more

meaningful for the subsequent cognitive functioning but are not easily measured in population-based settings.[12] Second, the cognitive battery assessment did not include all cognitive domains of interest. However, we constructed G-factor from four cognitive functioning measures which yields a good but not ideal proxy of the nominal five items G-factor.[46] Third, persons with depression may have cognitive impairment prior to the occurrence of depression and a perception of cognitive decline can be precipitate depression.[3,16] However, we excluded participants with prevalent dementia and controlled for baseline cognition function, thus reverse causality cannot easily explain our results. Fourth, the first cohort but not the second was assessed for history of depression four years prior to baseline (Figure). However, the sensitivity analysis stratified by cohort did not change the results (Supplement 4).

In conclusion, our study demonstrates that individuals who experienced multiple episodes of depression had accelerated decline in cognitive functioning over an average follow-up of 12 years compared to individuals without depression. In contrast, those who experienced a single episode of depression followed a similar trajectory of cognitive decline as those without depression. This result can help to reconcile the contrasting findings from clinical and population-based studies, and has implications for the prognosis and management of persons with chronic depression.

REFERENCES

1. Pantzar A, Laukka EJ, Atti AR, Fastbom J, Fratiglioni L, Backman L. Cognitive deficits in unipolar old-age depression: a population-based study. *Psychol Med.* 2014;44(5):937-947.
2. Alexopoulos GS. Depression in the elderly. *Lancet.* 2005;365(9475):1961-1970.
3. Perrino T, Mason CA, Brown SC, Spokane A, Szapocznik J. Longitudinal relationships between cognitive functioning and depressive symptoms among Hispanic older adults. *J Gerontol B Psychol Sci Soc Sci.* 2008;63(5):P309-317.
4. Air T, Weightman MJ, Baune BT. Symptom severity of depressive symptoms impacts on social cognition performance in current but not remitted major depressive disorder. *Front Psychol.* 2015;6:1118.
5. Shimada H, Park H, Makizako H, Doi T, Lee S, Suzuki T. Depressive symptoms and cognitive performance in older adults. *J Psychiatr Res.* 2014;57:149-156.
6. Hasselbalch BJ, Knorr U, Hasselbalch SG, Gade A, Kessing LV. The cumulative load of depressive illness is associated with cognitive function in the remitted state of unipolar depressive disorder. *Eur Psychiatry.* 2013;28(6):349-355.
7. Hasselbalch BJ, Knorr U, Hasselbalch SG, Gade A, Kessing LV. Cognitive deficits in the remitted state of unipolar depressive disorder. *Neuropsychology.* 2012;26(5):642-651.
8. Gale CR, Allerhand M, Deary IJ, Team HAS. Is there a bidirectional relationship between depressive symptoms and cognitive ability in older people? A prospective study using the English Longitudinal Study of Ageing. *Psychol Med.* 2012;42(10):2057-2069.
9. Goveas JS, Espeland MA, Woods NF, Wassertheil-Smoller S, Kotchen JM. Depressive symptoms and incidence of mild cognitive impairment and probable dementia in elderly women: the Women's Health Initiative Memory Study. *J Am Geriatr Soc.* 2011;59(1):57-66.
10. Austin MP, Mitchell P, Goodwin GM. Cognitive deficits in depression: possible implications for functional neuropathology. *Br J Psychiatry.* 2001;178:200-206.
11. Uher R, Payne JL, Pavlova B, Perlis RH. Major depressive disorder in DSM-5: implications for clinical practice and research of changes from DSM-IV. *Depress Anxiety.* 2014;31(6):459-471.
12. McClintock SM, Husain MM, Greer TL, Cullum CM. Association between depression severity and neurocognitive function in major depressive disorder: a review and synthesis. *Neuropsychology.* 2010;24(1):9-34.
13. Ganguli M, Du Y, Dodge HH, Ratcliff GG, Chang CC. Depressive symptoms and cognitive decline in late life: a prospective epidemiological study. *Arch Gen Psychiatry.* 2006;63(2):153-160.
14. Wilson RS, Mendes De Leon CF, Bennett DA, Bienias JL, Evans DA. Depressive symptoms and cognitive decline in a community population of older persons. *J Neurol Neurosurg Psychiatry.* 2004;75(1):126-129.
15. Bassuk SS, Berkman LF, Wypij D. Depressive symptomatology and incident cognitive decline in an elderly community sample. *Arch Gen Psychiatry.* 1998;55(12):1073-1081.
16. Vinkers DJ, Gussekloo J, Stek ML, Westendorp RG, van der Mast RC. Temporal relation between depression and cognitive impairment in old age: prospective population based study. *BMJ.* 2004;329(7471):881.
17. Geerlings MI, Schoevers RA, Beekman AT, et al. Depression and risk of cognitive decline and Alzheimer's disease. Results of two prospective community-based studies in The Netherlands. *Br J Psychiatry.* 2000;176:568-575.
18. Dufouil C, Fuhrer R, Dartigues JF, Alperovitch A. Longitudinal analysis of the association between depressive symptomatology and cognitive deterioration. *Am J Epidemiol.* 1996;144(7):634-641.
19. Royall DR, Palmer R, Chiodo LK, Polk MJ. Depressive symptoms predict longitudinal change in executive control but not memory. *Int J Geriatr Psychiatry.* 2012;27(1):89-96.

20. Yaffe K, Blackwell T, Gore R, Sands L, Reus V, Browner WS. Depressive symptoms and cognitive decline in nondemented elderly women: a prospective study. *Arch Gen Psychiatry.* 1999;56(5):425-430.
21. Kohler S, van Boxtel MP, van Os J, et al. Depressive symptoms and cognitive decline in community-dwelling older adults. *J Am Geriatr Soc.* 2010;58(5):873-879.
22. Freiheit EA, Hogan DB, Eliasziw M, et al. A dynamic view of depressive symptoms and neurocognitive change among patients with coronary artery disease. *Arch Gen Psychiatry.* 2012;69(3):244-255.
23. Zeki Al Hazzouri A, Vittinghoff E, Byers A, et al. Long-term cumulative depressive symptom burden and risk of cognitive decline and dementia among very old women. *J Gerontol A Biol Sci Med Sci.* 2014;69(5):595-601.
24. Dotson VM, Resnick SM, Zonderman AB. Differential Association of Concurrent, Baseline, and Average Depressive Symptoms with Cognitive Decline in Older Adults. *Am J Geriatr Psychiatry.* 2008;16(4):318-330.
25. Graziane JA, Beer JC, Snitz BE, Chang CC, Ganguli M. Dual Trajectories of Depression and Cognition: A Longitudinal Population-Based Study. *Am J Geriatr Psychiatry.* 2016;24(5):364-373.
26. Paterniti S, Verdier-Taillefer MH, Dufouil C, Alperovitch A. Depressive symptoms and cognitive decline in elderly people. Longitudinal study. *Br J Psychiatry.* 2002;181:406-410.
27. Hofman A, Brusselle GG, Darwish Murad S, et al. The Rotterdam Study: 2016 objectives and design update. *Eur J Epidemiol.* 2015;30(8):661-708.
28. Luijendijk HJ, van den Berg JF, Dekker MJ, et al. Incidence and recurrence of late-life depression. *Arch Gen Psychiatry.* 2008;65(12):1394-1401.
29. Beekman AT, Deeg DJ, Van Limbeek J, Braam AW, De Vries MZ, Van Tilburg W. Criterion validity of the Center for Epidemiologic Studies Depression scale (CES-D): results from a community-based sample of older subjects in The Netherlands. *Psychol Med.* 1997;27(1):231-235.
30. Ott A, Breteler MM, van Harskamp F, Stijnen T, Hofman A. Incidence and risk of dementia. The Rotterdam Study. *Am J Epidemiol.* 1998;147(6):574-580.
31. Houx PJ, Jolles J, Vreeling FW. Stroop interference: aging effects assessed with the Stroop Color-Word Test. *Exp Aging Res.* 1993;19(3):209-224.
32. Welsh KA, Butters N, Mohs RC, et al. The Consortium to Establish a Registry for Alzheimer's Disease (CERAD). Part V. A normative study of the neuropsychological battery. *Neurology.* 1994;44(4):609-614.
33. Bleecker ML, Bolla-Wilson K, Agnew J, Meyers DA. Age-related sex differences in verbal memory. *J Clin Psychol.* 1988;44(3):403-411.
34. Marsh HW, Hau K-T, Wen Z. In Search of Golden Rules: Comment on Hypothesis-Testing Approaches to Setting Cutoff Values for Fit Indexes and Dangers in Overgeneralizing Hu and Bentler's (1999) Findings. *Structural Equation Modeling: A Multidisciplinary Journal.* 2004;11(3):320-341.
35. Marsh HW, Wen Z, Hau KT. Structural equation models of latent interactions: evaluation of alternative estimation strategies and indicator construction. *Psychol Methods.* 2004;9(3):275-300.
36. La Rue A, Swan GE, Carmelli D. Cognition and depression in a cohort of aging men: results from the Western Collaborative Group Study. *Psychol Aging.* 1995;10(1):30-33.
37. Lichtenberg PA, Ross T, Millis SR, Manning CA. The relationship between depression and cognition in older adults: a cross-validation study. *J Gerontol B Psychol Sci Soc Sci.* 1995;50(1):P25-P32.
38. Sapolsky RM. The possibility of neurotoxicity in the hippocampus in major depression: a primer on neuron death. *Biol Psychiatry.* 2000;48(8):755-765.
39. Lupien SJ, McEwen BS, Gunnar MR, Heim C. Effects of stress throughout the lifespan on the brain, behaviour and cognition. *Nat Rev Neurosci.* 2009;10(6):434-445.
40. Videbech P, Ravnkilde B. Hippocampal volume and depression: a meta-analysis of MRI studies. *Am J Psychiatry.* 2004;161(11):1957-1966.
41. McIntyre RS, Cha DS, Soczynska JK, et al. Cognitive deficits and functional outcomes in major depressive disorder: determinants, substrates, and treatment interventions. *Depress Anxiety.* 2013;30(6):515-527.

42. Yaffe K, Kanaya A, Lindquist K, et al. The metabolic syndrome, inflammation, and risk of cognitive decline. *JAMA.* 2004;292(18):2237-2242.

43. Gimeno D, Kivimaki M, Brunner EJ, et al. Associations of C-reactive protein and interleukin-6 with cognitive symptoms of depression: 12-year follow-up of the Whitehall II study. *Psychol Med.* 2009;39(3):413-423.

44. Butters MA, Young JB, Lopez O, et al. Pathways linking late-life depression to persistent cognitive impairment and dementia. *Dialogues Clin Neurosci.* 2008;10(3):345-357.

45. Alexopoulos GS. The vascular depression hypothesis: 10 years later. *Biol Psychiatry.* 2006;60(12):1304-1305.

46. Jensen AR. *THE g FACTOR The Science of Mental Ability.* Westport, Connecticut, United States of America: Praeger Publishers, 88 Post Road West, Westport, CT 06881 An imprint of Greenwood Publishing Group, Inc.; 1998.

CHAPTER 5.2

Myocardial infarction and the long-term risk for depression

Jovanova Olivera, Luik Annemarie, Leening Marten, Noordam Raymond, Aarts Nikkie, Hofman Albert, Franco Oscar, Dehghan Abbas, Tiemeier Henning

Psychological Medicine. 2016 Mar; 46(9):1951-60.

ABSTRACT

Background: The association between myocardial infarction (MI) and depression is well described. Yet, the underlying mechanisms are unclear and the contribution of psychological factors is uncertain. We aimed to determine the risk of recognized (RMI) and unrecognized myocardial infarction (UMI) on depression, as both have a similar impact on cardiovascular health but differ in psychological epiphenomena.

Methods and results: Participants of the Rotterdam Study, 1 823 men aged ≥ 55, were followed for the occurrence of depression. RMI and UMI were ascertained using electrocardiography and medical history at baseline. We determined the strength of the association of RMI and UMI with mortality, and we studied the relation of RMI and UMI with depressive symptoms and the occurrence of major depression.

Results: The risk of mortality was similar in men with RMI (adjusted hazard ratio (*HR*): 1.71, 95%CI: 1.45-2.03) and UMI (*HR*: 1.58, 95%CI: 1.27-1.97). Men with RMI had on average (unstandardized regression coefficient (*B*): 1.14, 95%CI: 0.07-2.21) higher scores for depressive symptoms. By contrast, we found no clear association between UMI and depressive symptoms (*B*: 0.55, 95%CI: -0.51-1.62) in men. Analysis including occurrence of major depression as the outcome were consistent with the pattern of association.

Conclusion: The discrepant association of RMI and UMI with mortality compared to depression suggests that the psychological burden of having experienced an MI contributes to the long-term risk of depression.

INTRODUCTION

The increased occurrence of depression after myocardial infarction (MI) is well-recognized in clinical practice.[1-3] In those that experienced MI, the prevalence of depression exceeds 30%.[2,4] Several biological and pathophysiological pathways including inflammation and endothelial dysfunction might link MI and depression[5,6], but post-MI depression could also occur as a result of illness perception. Alterations in mood can be part of the coping process after surviving a cardiac event, treatment and recovery.[4]

MI is traditionally described as a clinical syndrome characterized with precordial discomfort and pain. Nevertheless, a substantial proportion of MIs are not clinically recognized by patients or treating physicians.[7-9] The available epidemiological, clinical and basic science evidence provides support for a similar ischemic pathology underlying both recognized (RMI) and unrecognized (UMI) myocardial infarctions.[10-14] Many studies demonstrated that UMI has a similar long-term impact on physical health as RMI. In men, and to a lesser extent in women, both RMI and UMI are associated with increased risk of all-cause mortality, heart failure and sudden death.[14,15] In the Rotterdam Study, men with UMI were also at increased risk for dementia, stroke and cerebral small vessel diseases.[16-19]

Whether there is a common mechanism that engenders RMI and clinically UMI remains controversial.[12] One theory suggests that the cardiac-pain mechanism is the main reason for the difference in the perception of the discomfort.[20] A second theory suggests that various humoral and neurological factors together determine whether the symptoms arise from myocardial ischemia.[21] Third, despite similar pathophysiology the size, location or duration of the myocardial ischemia could explain the absence of pain in UMI.[10,14,21] There is a lack of empirical evidence to supporting any of the above theories and the reported long-term impacts of UMI on physical health suggests that differences in size and duration of ischemia are unlikely explanations. Finally, psychological factors regarding the recognition and willingness of the patient to seek medical attention have been discussed.[12]

To test the impact of RMI and UMI in the present study population, we first compared the association of RMI and UMI with mortality in men and women (note, mortality has been studied in a slightly different sample of the Rotterdam Study previously[15]). By definition, the awareness of having survived a life-threatening health event is absent in persons with UMI. This characteristic of UMI provides us with the opportunity to explore the potential role of psychological mechanisms underlying the relation between MI and depression, but only if the long-term impact can be established and if depression can be assessed independently of help seeking behavior. In this case our approach provides us with the unique opportunity to disentangle the pathophysiological from the psychological mechanisms in the association between MI and depression.

The main aim of our study was to determine the relation of RMI and UMI with depression in older persons. If MI would cause depression mainly due to psychological mechanisms, we would expect that individuals, who are unaware of having had an MI, are not at risk of developing depression. Indeed we hypothesized that RMI but not UMI is associated with an increased risk of depression.

METHODS

The study was part of the Rotterdam Study, a population-based cohort designed to study the occurrence and determinants of diseases.[22] In 1990, all residents aged ≥ 55 of Ommoord, a district in Rotterdam, the Netherlands, were invited to participate. Informed consent to retrieve information from treating physicians was obtained from all participants. The medical ethics committee approved the study according to the Wet Bevolkingsonderzoek-Population Study Act executed by the Ministry of Health, the Netherlands.

In the Rotterdam Study, four examination rounds have taken place. In each round, participants underwent home interviews and subsequent physical examinations at the research center. During the second examination (1993 - 1995), 4 940 persons were interviewed for depressive symptoms. This round constituted the baseline for the current study.[23] We excluded 461 persons without electrocardiography, due to technical problems, or lack of personnel to operate the electrocardiography device. Another 23 persons were excluded because they withdrew the informed consent. We excluded persons with a Mini Mental State Exam (MMSE) score of ≤ 26 at baseline (*N* = 287)[24], no valid information on the MMSE (*N* = 129), and bipolar disorder (*N* = 3). In total, 4 037 participants were eligible with a valid electrocardiography and evaluation of depressive symptoms. These participants were followed from baseline to occurrence of major depression (MDD), death, loss to follow-up, or the end of the study (January 1st, 2010).

History of myocardial infarction

MI status was assessed as described in detail previously.[9] In short, at the research center visit, a 10-second 12-lead electrocardiogram was recorded using an ACTA Gnosis IV-electrocardiography recorder (Esaote Biomedica, Florence, Italy) at a sampling frequency of 500 Hz and sored digitally. Electrocardiograms were processed by the Modular electrocardiogram Analysis System (MEANS) to obtain electrocardiogram measurement and interpretation.[25] Two research doctors blinded for other clinical information validated the electrocardiograms selected by MEANS. The diagnosis of MI using MEANS is mainly driven by pathological Q-waves and auxiliary criteria, such as QR ratio and R-wave progression. ST-T changes were not considered as criteria for MI by MEANS, but were taken into account by the clinicians validating and ascertaining the diagnosis of MI.

A senior cardiologist, ascertained the final diagnosis of MI. Assessment of clinically RMI was done as previously reported.[9] A history of RMI was defined by self-reported MI confirmed by clinical data (GP records, and Nationwide Medical Registry), and confirmed or not confirmed with electrocardiogram characteristics matching MI.[9] A history of UMI included all participants without documentation or self-reported MI, but with electrocardiogram characteristics matching MI.[9] All UMIs were therefore Q-wave MIs. Persons without indication of MI on electrocardiography and no medical documentation of an earlier MI constituted the reference group.[9]

Vital status

The information on vital status is collected continuously, on a weekly basis from the Central Register of Population of Rotterdam. The follow-up for mortality started from the assessment of depression until loss to follow-up or September 5th, 2013 (mean follow-up 13.9 years).[9,15]

Depression

At baseline, all participants were screened for depression with the validated Dutch versions of the Centre for Epidemiologic Studies Depression (CES-D) scale[26], or the Hospital Anxiety and Depression Scale (HADS-D).[27] The reliability of these questionnaires was tested formally: the Cronbach alphas were 0.82 and 0.83, respectively. The CES-D consists of 20 items related to assess depressive symptoms. All symptoms were self-rated on a 0-3 scale, dependent upon the frequency of the experienced symptom.[28] Participants could score from 0 up to 60. The HADS-D consists of 7 items measuring depressive symptoms. Again, participants rated the symptoms on a 0 - 3 scale. Standardized Z-score of the CES-D and HADS-D scores at baseline was used to control for the severity of depressive symptoms and adjust for pre-existing depressive symptoms.
During first and second follow-up assessments (at the examination rounds in 1997 - 1999 and 2002-2004) the CES-D scale was used to determine depressive symptoms.

From 1993 till 2010, participants were followed for the occurrence of MDD, which was determined from two sources of information as described previously.[23] At follow-up examinations all participants who scored 16 or above on the CES-D scale, were invited to the research center for a semi-structured clinical interview to diagnose depressive disorders. A trained clinician diagnosed depression according to the Diagnostic and Statistical Manual of Mental Disorders, 4th revised edition (DSM-IV) by using the Dutch version of the Schedules for Clinical Assessment in Neuropsychiatry (SCAN).[29] Second, medical records of general practitioners were continuously monitored for occurrence of an episode of MDD from baseline onward (mean follow-up 12.5 years). Considering the information about possible MDD provided by the psychiatric interviews and medical records, we defined the occurrence of MDD episodes, as the first event that

chronologically occurred in one of the two data sources (*note*, all record retrievals were blinded for prior episodes). The occurrences inferred from the different data sources were consistent.[23] However, one assessment (SCAN) was repeatedly cross-sectional and the other assessment reflects continuous monitoring, thus such comparison must be interpreted carefully to account for the different assessment methods.

Covariates

Potential confounders were selected on the basis of prior knowledge.[1,2] Marital status, education, smoking, and alcohol use were assessed at baseline during the home interview. Education was classified as low, intermediate, or high. Participants were classified as never married or divorced, married, and widowed. Smoking habits were classified as never, past, or current smoker and, similarly, alcohol consumption was classified as never, past, or current consumer. The MMSE was used to measure cognitive performance.[24]

Blood pressure was measured twice in the sitting position at the right upper arm with a random-zero sphygmomanometer. The average of these consecutive measurements was calculated. Diabetes was considered present if fasting plasma glucose was 7.0 mmol/l or higher, non-fasting glucose or an oral glucose tolerance test result of 11.1 mmol/l or higher, or if the participant used blood glucose-lowering medication.[30] A history of stroke was defined as a self-reported stroke verified by medical records.[31] Height and weight were measured without heavy clothing to calculate body mass index (BMI).

DATA ANALYSIS

First we determined the relative mortality risk of participants with RMI and UMI using Cox proportional hazard regression analyses. These analyses were adjusted for age, level of education, smoking status, and alcohol consumption, history of stroke, diabetes, systolic blood pressure, and baseline depressive symptoms. Results of previous studies suggested sex difference in the prognosis of RMI and UMI.[15] Also, the interaction term of history of MI with sex and mortality was statistically significant ($p < 0.05$). Therefore we stratified all analyses by sex.

Second, we studied whether RMI and UMI are associated with depressive symptoms with linear regression. The models were adjusted for baseline depressive symptoms in order to correct for pre-existing depression. Generalized Estimating Equation (GEE) approach was used to estimate the pooled effect of RMI and UMI on repeatedly assessed depressive symptoms. The interaction term of history of MI with sex and depressive symptoms was statistically significant ($p < 0.05$).

All analyses related to the association between RMI/UMI and depressive symptoms were tested using two models. The first model was adjusted for age and baseline depressive symptoms. The second model was additionally adjusted for level of education, marital status, smoking status, alcohol consumption, BMI, history of stroke, diabetes mellitus, and systolic blood pressure.
Third, we assessed the association between RMI/UMI and MDD using Cox proportional hazard regression analyses. These analyses included two models with the same covariates in those testing depressive symptoms described above. The Breslow-Generalized Wilcoxon test was used to compare the hazard curves of RMI and UMI.

Multiple imputations were used in order to account for missing data on potential confounding variables. Smoking status was missing in 2.9%, marital status in 16.9%, BMI in 3.5%, alcohol consumption in 1.4%, systolic blood pressure in 1.5%, and education in 0.3% of the participants. All analyses were rerun in the complete case and five imputed data sets. In this paper we present results from the imputed data. All analyses were performed using IBM SPSS Statistics, version 21 (IBM Corp., Somers, NY USA).

RESULTS

Baseline characteristics of 4 037 participants are presented in Table 1. The mean age of men was 68.1 years (6.9 *SD*) and women were on average 68.5 years (7.7 *SD*). Of the men, 12.3% (*N* = 220) had a RMI, 6.3% (*N* = 115) an UMI and 81.4% (*N* = 1 488) had no evidence of a MI. Of the women, 2.4% (*N* = 53) had had a RMI, 4.7% (*N* = 104) had had an UMI and 92.9% (*N* = 2 057) were free of MI. In total, 273 persons had a RMI, of these 138 had electrocardiography evidence of MI. The remaining 135 RMI cases were identified by clinical records and therefore represent non-Q-wave MIs.

First, we assessed the mortality risk of RMI and UMI (Table 2). In men, the average follow-up was 13.1 years (23 846 person-years) during which 943 men died. RMI and UMI were associated with an increased risk of mortality: recognized MI (*HR*: 1.71, 95%CI: 1.45-2.03); UMI (*HR*: 1.58, 95%CI: 1.27-1.97). In women, the average follow-up was 13.9 years (32 438 person-years) during which 825 women died. Women with RMI (*N* = 53) had a higher mortality risk (*HR*: 1.81, 95%CI: 1.33-2.46). However, there was no association between UMI (*N* = 104) and mortality risk in women (*HR*: 0.99, 95%CI: 0.77-1.30).

Next, we studied the association between RMI and UMI and repeated assessments of depressive symptoms after a mean of 3.8 and 8.5 years The mean CES-D score at the first and second follow-up assessment was 4.2 (6.7 *SD*) and 6.7 (7.7 *SD*), respectively.

Table 1. Baseline characteristics of the study population (*N* = 4 037)

	History of myocardial infarction					
	Men (*N*=1 823)			Women (*N*=2214)		
	No MI (*N*=1488)	Recognized MI (*N*=220)	Unrecognized MI (*N*=115)	No MI (*N*=2057)	Recognized MI (*N*=53)	Unrecognized MI (*N*=104)
Age, years, mean (SD)	67.6 (6.9)	69.8 (6.5)	70.1 (7.2)	68.2 (7.6)	74.3 (8.7)	71.8(7.9)
Depressive symptoms at baseline, score, mean (SD)*	4.3 (4.7)	4.9 (4.9)	4.3 (5.6)	5.4 (5.9)	7.2 (5.7)	6.1 (5.9)
Smoking status,						
Never smoker, *N* %	132 (8.9)	11 (5.1)	7 (6.4)	944 (48)	23 (46)	58 (56.9)
Past smoker, *N* %	1041 (70.6)	167 (77)	78 (71.6)	658 (33.5)	21 (42)	28 (27.5)
Current smoker, *N* %	302 (20.5)	39 (18)	24 (22)	363 (18.5)	6 (12)	16 (15.7)
Alcohol consumption,						
Never consumer, *N* %	82 (5.5)	26 (11.9)	3 (2.6)	293 (14.5)	14 (27.5)	19 (19)
Past consumer, *N* %	100 (6.8)	23 (10.6)	11 (9.6)	345 (17.1)	10 (19.6)	24 (24)
Current consumer, *N* %	1299 (87.7)	169 (77.5)	101 (87.8)	1378 (68.4)	27 (52.9)	57 (57)
MMSE, score, mean (SD)	28.2 (1.2)	28.2 (1.2)	28.1 (1.4)	28.1 (1.3)	27.9 (1.1)	28.2 (1.3)
BMI, kg/m2, mean (SD)	25.8 (2.9)	26.2 (3.1)	26.3 (3.2)	26.6 (3.9)	27.4 (4.5)	27.6 (4.5)
Systolic blood pressure, mmHg, mean (SD)	139.8 (21.6)	138.4 (20.4)	146.3 (21)	140 (21.9)	134.1 (19.2)	146.3 (23.4)
Blood pressure lowering medication, *N* %	274 (19.4)	81 (39.3)	22 (20.2)	521 (26.4)	18 (37.5)	30 (30.3)
History of cardiovascular diseases, *N* %	80 (5.4)	205 (93.2)	99 (86.1)	61 (3)	50 (94.3)	86 (82.7)
History of Stroke, *N* %	39 (2.6)	13 (5.9)	7 (6.1)	32 (1.6)	4 (7.5)	3 (2.9)
History of CABG, *N* %	36 (2.5)	59 (27.6))	0 (0)	9 (0.4))	8 (15.4))	0 (0)
History of PCI, *N* %	10 (0.7)	12 (5.6)	0 (0)	6 (0.3)	4 (7.7)	1 (1)
Diabetes mellitus, *N* %	121 (8.1)	40 (18.2)	17 (16.5)	162 (7.9)	7 (13.2)	11 (10.6)
Highest education attained,						
Low education, *N* %	282 (19)	59 (26.78)	26 (23.2)	729 (35.5)	27 (50.9)	35 (34)
Intermediate education, *N* %	935 (63)	131 (59.5)	67 (59.8)	1209 (58.9)	25 (47.2)	61 (59.2)
High education, *N* %	268 (18)	30 (13.6)	19 (17)	113 (5.5)	1 (1.9)	7 (6.8)
Marital status,						
Never married or divorced, *N* %	57 (4.6)	9 (5.6)	6 (7)	249 (14.3)	4 (11.1)	7 (8.1)
Married or living together, *N* %	1055 (84.7)	140 (86.4)	66 (76.7)	970 (55.7)	14 (38.9)	41 (47.7)
Widowed, *N* %	133 (10.7)	13 (8)	14 (16.3)	522 (30)	18 (50)	38 (44.2)

Table presents complete data, missing values are not imputed here.*CES-D/HADS-D z-score.

Table 2. Recognized and unrecognized MI and all-cause mortality risk with Cox regression (*N* = 4 037)

	All-cause mortality	
	HR (95% CI)	***p* value**
Men (*N* = 1 823)		
Recognized MI (*N* = 220)		
Multivariate adjusted*	1.71 (1.45;2.03)	<0.001
Unrecognized MI (*N* = 115)		
Multivariate adjusted*	1.58 (1.27;1.97)	<0.001
Women (*N* = 2 214)		
Recognized MI (*N* = 53)		
Multivariate adjusted*	1.90 (1.39;2.61)	<0.001
Unrecognized MI (*N* = 104)		
Multivariate adjusted*	0.99 (0.77;1.30)	0.99

1 768 (943 men) of 4 037 person died during 56 284 person-years of follow-up.
* Multivariate adjusted: additionally adjusted for age, level of education, smoking status, Alcohol consumption, history of stroke, diabetes mellitus and systolic blood pressure.

In men (Table 3), RMI predicted higher scores of depressive symptoms at the first (*B*: 1.05, 95%CI: 0.16-1.95) and second assessment (*B*: 1.29, 95%CI: 0.07-2.52). UMI was not associated with depressive symptoms at the first (*B*: 0.79, 95%CI: –0.34-1.97), or second assessment (*B*: 0.54, 95%CI: –1.14-2.21). In the combined analysis of the two assessments, we observed that RMI (*B*: 1.14, 95%CI: 0.07-2.21) but not UMI (*B*: 0.55, 95%CI: –0.51-1.62) was associated with more depressive symptoms. In women (Table 4), RMI was not significantly associated with depressive symptoms in the first (*B*: -1.37, 95%CI: –3.64-0.90) or in the second assessment (*B*: 1.48, 95%CI: –1.94-4.89). Moreover, there were no significant associations in the combined analysis of the two depressive symptoms assessments (*B*: -0.86, 95%CI: –3.13-1.42). UMI was not associated with depressive symptoms in any analysis. For this reason we restricted the additional analysis on occurrence of MDD in men only.

In an additional sensitivity analysis we explored the relation of RMI and UMI with the occurrence of MDD. Men were followed up for 11.8 years ± 4.71 *SD* (21 664 person years). In total, 52 men (10 of those with RMI but only 3 of those with UMI) developed MDD during follow-up. We calculated occurrence of MDD during follow-up and cannot present point prevalence data. Men with a history of RMI were at increased risk to develop MDD (*HR*: 2.18, 95%CI: 1.04-4.57), but not those with an UMI (*HR*: 0.99, 95%CI: 0.77-1.30). The difference between the risk of RMI and UMI was statistically significant ($p = 0.024$).

We repeated the analyses (see Supplementary material 1) excluding RMI cases that did not meet electrocardiography criteria for prior Q-wave MI, to reduce possible misclassification and bias by person (over) reporting symptoms. In these analyses, the effect estimates were very similar. However, due to loss of statistical power the effect was no longer statistically significant.

Table 3. The longitudinal association of recognized or unrecognized MI with depressive symptoms in men (*N* = 1 823)

	Depressive symptoms							
		First assessment[§] (*N* = 1471)			Second assessment[§] (*N* = 1113)		Combined analysis (*N* = x, with 2584 observations)	
Exposure	***N***	**B (95% CI)**	***p* value**	***N***	**B (95% CI)**	***p* value**	**B (95% CI)**	***p* value**
Recognized MI								
Adjusted for age and baseline depressive symptoms*	159	0.92 (0.29;1.81)	0.043	102	1.26 (0.04;2.48)	0.043	1.06 (-0.09;2.13)	0.052
Multivariate adjusted**	159	1.05 (0.16;1.95)	0.021	102	1.29 (0.07;2.52)	0.039	1.14 (0.07;2.21)	0.037
Unrecognized MI								
Adjusted for age and baseline depressive symptoms*	84	0.85 (-0.33;2.04)	0.16	51	0.50 (-1.18;2.18)	0.56	0.62 (-0.44;1.67)	0.25
Multivariate adjusted**	84	0.79 (-0.40;1.97)	0.19	51	0.54 (-1.14;2.21)	0.53	0.55 (-0.51;1.62)	0.31

Analysis were performed with linear regression and generalized estimating equations.

[§] CES-D mean score (SD); **First assessment**: Recognized MI 4.2 (7.9), Unrecognized MI 3.8 (6.5) and No MI 2.9 (5.3); **Second assessment:** Recognized MI 6.7 (8.1), Unrecognized MI 5.4 (5.9) and No MI 5.1 (6.3)

*Model adjusted for age and baseline depression (corrected for baseline depressive symptoms);

**Multivariate adjusted: additionally adjusted for level of education, marital status, body mass index, smoking status, alcohol consumption, and history of stroke, diabetes mellitus and systolic blood pressure

Table 4. The longitudinal association of recognized or unrecognized MI with depressive symptoms in women (*N* = 2 214)

		Depressive symptoms						
		First assessment[§] (*N* = 1831)			Second assessment[§] (*N* = 1506)		Combined analysis (*N* = x, with 3337 observations)	
Exposure	***N***	**B (95% CI)**	***p* value**	***N***	**B (95% CI)**	***p* value**	**B (95% CI)**	***p* value**
Recognized MI								
Adjusted for age and baseline depressive symptoms*	36	-1.11 (-3.37;1.16)	0.34	19	1.24 (-2.20;4.68)	0.48	-0.73 (-3.0;1.54)	0.53
Multivariate adjusted**	36	-1.37 (-3.64;0.90)	0.24	19	1.48 (-1.94;4.89)	0.40	-0.86 (-3.13;1.42)	0.46
Unrecognized MI								
Adjusted for age and baseline depressive symptoms*	84	0.12 (-1.383;1.62)	0.87	62	0.57 (-1.37;2.50)	0.57	0.23 (-1.18;1.64)	0.75
Multivariate adjusted**	84	-0.04 (-1.54;1.46)	0.96	62	0.39 (-1.54;2.33)	0.69	0.12 (-1.30;1.54)	0.87

Analysis were performed with linear regression and generalized estimating equations.

[§] CES-D mean score (SD); **First assessment**: Recognized MI 5.2 (8.2), Unrecognized MI 5.9 (8.2) and No MI 4.9 (7.2); **Second assessment:** Recognized MI 10.2 (9.2), unrecognized MI 9.1 (8.4) and No MI 7.6 (8.3)

*Model adjusted for age and baseline depression (corrected for baseline depressive symptoms);

**Multivariate adjusted: additionally adjusted for level of education, marital status, body mass index, smoking status, alcohol consumption, and history of stroke, diabetes mellitus and systolic blood pressure.

DISCUSSION

Men with RMI or UMI had an increased mortality risk in this population-based study, as demonstrated previously.[15] However, the risks of RMI and UMI for depression did not correspond to the mortality risks. Whereas men with RMI had an increased risk to develop depression, men with UMI were at similar risk to develop depression compared to persons without MI. In line with other studies, we confirmed a much weaker association between unrecognized MI and survival in women.[8] This has been attributed to the increased likelihood of lead-placement artefacts due to breast tissue.[32] Therefore, the results of the association between unrecognized MI and depression in women are hard to interpret.

A history of UMI conveyed a long-term increased risk of both depressive symptoms and MDD in men. Many studies of cardiovascular disease and depression focused on the prognostic value of short-term incidence of depression in patients shortly after the diagnosis of MI.[2,33,34] One such study reported that patients with heart disease have 45% more depressive symptoms compared to those without heart disease.[35] The long-term risk of psychological distress in post-MI patients as well as depression in patients with cardiovascular disease have previously been studied for follow-up periods of 5-8 years.[35,36] The present study addresses a period of 11 years after experiencing MI. In men, RMI was associated with an increased risk of MDD that remained elevated even over a long period. However, in men UMI was not associated with depressive symptoms or occurrence of MDD, although it was a risk factor for mortality. This discrepancy in the risks of UMI for mortality and depression, when compared to RMI, can be explained in multiple ways.

First, previous studies have suggested that patients with UMI may have less severe symptoms and mostly do not consult physicians.[12,37] Others have argued that, the size of myocardial tissue damage underlying UMI may also be less severe.[38] Yet, our study showed that in men, UMI was almost as predictive of mortality as RMI. This suggests that the underlying pathology is of equal severity.

Second, it is conceivable that the mortality risk of RMI and UMI are similar only due to the medical interventions and behavioral changes occurring after RMI. Successful treatment of RMI certainly decreases mortality risk.[39] However, it is questionable whether treating MI successfully has a strong impact on the risk of depression.[40]

Third, depression may also be an adverse effect of the treatment of RMI. Patients with RMI are more often treated with interventions such as percutaneous coronary intervention to restore the coronary blood flow and the tissue perfusion.[39] These procedures have been associated with increased risk of depressive symptoms because of the intensity of the treatment procedure.[3]

Also, drugs used for blood pressure and heart rate control, such as beta-blockers, may worsen depressive symptoms.[39] However, we accounted for antihypertensive use and previous cardiac interventions at baseline to control for this effect.

Fourth, our results suggest that in men psychological factors contribute to the development of depression after MI. Although it remains difficult to disentangle psychological mechanisms from the pathophysiological mechanism underlying both MI and depression, we argue that psychological mechanisms are key to the different risk patterns of RMI and UMI in men. Due to improved clinical management of clinical MI, the survival has improved substantially, yet patients with MI often experience the event as life threatening.[41] Further, frightening events impact not only the survivors health, but also theirs families and working life.[42] These resulting lifestyle changes contribute to the psychological reaction following the MI experience and often resulting in depression.[43] Clinical studies of MI patients demonstrated that the psychological distress related to health problems makes men particularly vulnerable to depression.[36]

We did not observe any relation between RMI or UMI with depressive symptoms in women. Only 53 women but 220 men had RMI at baseline, thus our analysis were underpowered in women. Moreover, we found no relation between UMI and mortality in women, suggesting that this event does not signal pathophysiological change or that there is possible misclassification of UMI in women. Some studies reported a low accuracy of electrocardiography diagnostics in older women.[32] In this light the association of UMI with both mortality and depression was consistent.

Depression is an established risk factor for coronary heart disease[44] and may thus contribute to the occurrence of MI. Although we correct for baseline depressive symptoms, reverse causality could in theory influence the observed association between RMI and depression.[45] The Rotterdam Study started with screening for depression during 1993, this assessment constitutes the baseline of our study and no prior assessments were conducted.

A limitation of our study is the possible misclassification of UMI. The definition of UMI heavily relies on the presence of pathological Q-waves, therefore most non-Q-wave MIs remained undetected. Second, electrocardiogram evidence of a previous MI can disappear over time.[20] This again could lead to misclassification of UMI and thereby lower the effect estimates. Third, the known history of RMI might introduce assessment bias towards MDD diagnosis. In contrast, individuals with UMI have less specific symptomatology and mostly seek no medical care thus may avoid MDD medical diagnosis. However, the systematic approaches we used to ascertain events of MDD minimize the possibility of these biases.[23] Forth, due to the observational nature of the study design presence of residual confounding cannot be ruled out.

Our study has several strengths. It is based on a large well-described population-based cohort with a long follow-up. Data were gathered prospectively and without prior knowledge of MI status. Participants were monitored continuously during follow-up for the occurrence of depression and diagnoses were made by clinical interviews according to the DSM criteria. We used different data sources and approaches to collect and validate MDD.[23]

Furthermore, detailed information on RMI was collected and validated from the medical records of the participants.[9] Importantly, when we restricted the cases with RMI to those with electrocardiography-verified changes (excluding those with clinical symptoms only) results did not change. This suggest that possible misclassification cannot explain our findings. Possibly, persons (over)reporting more symptoms (without pathology) do not account for the observed association and did not bias the effect estimates (see (Macleod *et al.*, 2002) for example).[46]

In conclusion, RMI and UMI are both associated with a higher mortality risk in men. However, only man with RMI had a higher likelihood of depression during follow-up. Although pathophysiological differences between UMI and RMI may also explain this discrepancy, our results are consistent with the hypothesis that the psychological burden of experiencing MI contributes to the long-term risk of depression after MI. We can only carefully speculate about clinical implication as future studies must confirm this results and the underlying mechanisms but our findings suggest that effective long-term prevention of post-MI depression should address the psychological impact of experiencing an MI.

R1
R2
R3
R4
R5
R6
R7
R8
R9
R10
R11
R12
R13
R14
R15
R16
R17
R18
R19
R20
R21
R22
R23
R24
R25
R26
R27
R28
R29
R30
R31
R32
R33
R34
R35
R36
R37
R38
R39

REFERENCES

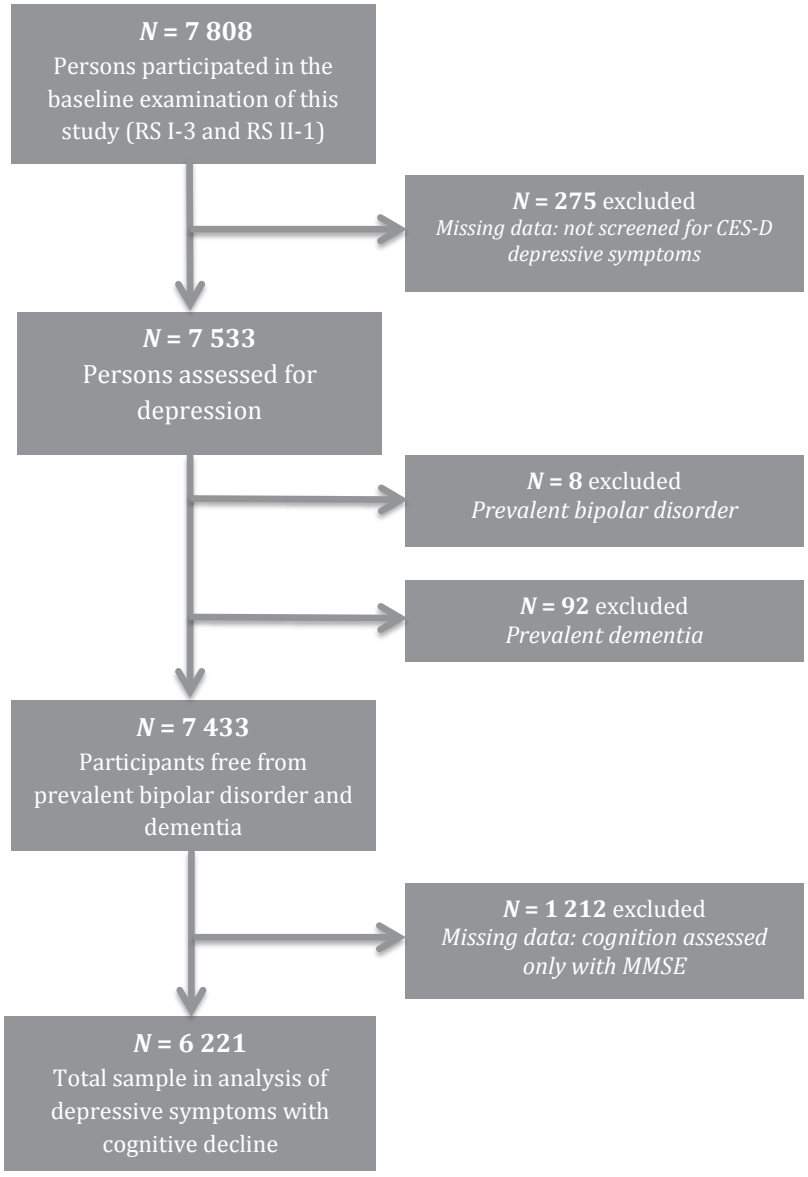
1. Denollet J, Strik JJ, Lousberg R, Honig A. Recognizing increased risk of depressive comorbidity after myocardial infarction: looking for 4 symptoms of anxiety-depression. *Psychotherapy and psychosomatics.* 2006;75(6):346-352.
2. Frasure-Smith N, Lesperance F, Talajic M. Depression following myocardial infarction. Impact on 6-month survival. *JAMA : the journal of the American Medical Association.* 1993;270(15):1819-1825.
3. Park MW, Park SH, Kwak DJ, Sung YM. Effects of substrate heating on the photovoltaic characteristics of dye-sensitized solar cells during two-step Ti film deposition by RF magnetron sputtering. *Journal of nanoscience and nanotechnology.* 2012;12(4):3309-3312.
4. Lesperance F, Frasure-Smith N. Depression in patients with cardiac disease: a practical review. *Journal of psychosomatic research.* 2000;48(4-5):379-391.
5. Tiemeier H, Hofman A, van Tuijl HR, Kiliaan AJ, Meijer J, Breteler MM. Inflammatory proteins and depression in the elderly. *Epidemiology.* 2003;14(1):103-107.
6. Harrison NA, Cooper E, Voon V, Miles K, Critchley HD. Central autonomic network mediates cardiovascular responses to acute inflammation: relevance to increased cardiovascular risk in depression? *Brain, behavior, and immunity.* 2013;31:189-196.
7. Stokes J, 3rd, Dawber TR. The silent coronary: the frequency and clinical characteristics of unrecognized myocardial infarction in the Framingham study. *Ann Intern Med.* 1959;50(6):1359-1369.
8. de Torbal A, Boersma E, Kors JA, et al. Incidence of recognized and unrecognized myocardial infarction in men and women aged 55 and older: the Rotterdam Study. *European heart journal.* 2006;27(6):729-736.
9. Leening MJ, Kavousi M, Heeringa J, et al. Methods of data collection and definitions of cardiac outcomes in the Rotterdam Study. *European journal of epidemiology.* 2012;27(3):173-185.
10. Cabin HS, Roberts WC. Quantitative comparison of extent of coronary narrowing and size of healed myocardial infarct in 33 necropsy patients with clinically recognized and in 28 with clinically unrecognized ("silent") previous acute myocardial infarction. *Am J Cardiol.* 1982;50(4):677-681.
11. Kannel WB, Abbott RD. Incidence and prognosis of unrecognized myocardial infarction. An update on the Framingham study. *N Engl J Med.* 1984;311(18):1144-1147.
12. Barsky AJ, Hochstrasser B, Coles NA, Zisfein J, O'Donnell C, Eagle KA. Silent myocardial ischemia. Is the person or the event silent? *JAMA.* 1990;264(9):1132-1135.
13. Kumar V, Abbas AK, Aster JC. *Robbins Basic Pathology.* Elsevier Health Sciences; 2012.
14. Schelbert EB, Cao JJ, Sigurdsson S, et al. Prevalence and prognosis of unrecognized myocardial infarction determined by cardiac magnetic resonance in older adults. *JAMA : the journal of the American Medical Association.* 2012;308(9):890-896.
15. Dehghan A, Leening MJ, Solouki AM, et al. Comparison of prognosis in unrecognized versus recognized myocardial infarction in men versus women >55 years of age (from the Rotterdam Study). *The American journal of cardiology.* 2014;113(1):1-6.
16. Ikram MA, Hollander M, Bos MJ, et al. Unrecognized myocardial infarction and the risk of stroke: the Rotterdam Study. *Neurology.* 2006;67(9):1635-1639.
17. Ikram MA, van Oijen M, de Jong FJ, et al. Unrecognized myocardial infarction in relation to risk of dementia and cerebral small vessel disease. *Stroke.* 2008;39(5):1421-1426.
18. Leening MJ, Elias-Smale SE, Felix JF, et al. Unrecognised myocardial infarction and long-term risk of heart failure in the elderly: the Rotterdam Study. *Heart (British Cardiac Society).* 2010;96(18):1458-1462.
19. Krijthe BP, Leening MJ, Heeringa J, et al. Unrecognized myocardial infarction and risk of atrial fibrillation: the Rotterdam Study. *Int J Cardiol.* 2013;168(2):1453-1457.

20. Sheifer SE, Manolio TA, Gersh BJ. Unrecognized myocardial infarction. *Annals of internal medicine.* 2001;135(9):801-811.
21. Cohn PF, Fox KM, Daly C. Silent myocardial ischemia. *Circulation.* 2003;108(10):1263-1277.
22. Hofman A, Darwish Murad S, van Duijn CM, et al. The Rotterdam Study: 2014 objectives and design update. *European journal of epidemiology.* 2013;28(11):889-926.
23. Luijendijk HJ, van den Berg JF, Dekker MJ, et al. Incidence and recurrence of late-life depression. *Archives of general psychiatry.* 2008;65(12):1394-1401.
24. Folstein MF, Folstein SE, McHugh PR. "Mini-mental state". A practical method for grading the cognitive state of patients for the clinician. *Journal of psychiatric research.* 1975;12(3):189-198.
25. van Bemmel JH, Kors JA, van Herpen G. Methodology of the modular ECG analysis system MEANS. *Methods of information in medicine.* 1990;29(4):346-353.
26. Beekman AT, Deeg DJ, Van Limbeek J, Braam AW, De Vries MZ, Van Tilburg W. Criterion validity of the Center for Epidemiologic Studies Depression scale (CES-D): results from a community-based sample of older subjects in The Netherlands. *Psychological medicine.* 1997;27(1):231-235.
27. Bjelland I, Dahl AA, Haug TT, Neckelmann D. The validity of the Hospital Anxiety and Depression Scale. An updated literature review. *Journal of psychosomatic research.* 2002;52(2):69-77.
28. Breslau N. Depressive symptoms, major depression, and generalized anxiety: a comparison of self-reports on CES-D and results from diagnostic interviews. *Psychiatry research.* 1985;15(3):219-229.
29. Piyavhatkul N, Krisanaprakornkit T, Paholpak S, Khiewyou J. Validity and reliability of WHO Schedules for Clinical Assessment in Neuropsychiatry (SCAN)-Thai version: Cognitive Impairment or Decline Section. *J Med Assoc Thai.* 2008;91(7):1129-1136.
30. Dehghan A, Kardys I, de Maat MP, et al. Genetic variation, C-reactive protein levels, and incidence of diabetes. *Diabetes.* 2007;56(3):872-878.
31. Wieberdink RG, Ikram MA, Hofman A, Koudstaal PJ, Breteler MM. Trends in stroke incidence rates and stroke risk factors in Rotterdam, the Netherlands from 1990 to 2008. *European journal of epidemiology.* 2012;27(4):287-295.
32. Murabito JM, Evans JC, Larson MG, Levy D. Prognosis after the onset of coronary heart disease. An investigation of differences in outcome between the sexes according to initial coronary disease presentation. *Circulation.* 1993;88(6):2548-2555.
33. Thombs BD, Bass EB, Ford DE, et al. Prevalence of depression in survivors of acute myocardial infarction. *J Gen Intern Med.* 2006;21(1):30-38.
34. Larsen KK. Depression following myocardial infarction--an overseen complication with prognostic importance. *Dan Med J.* 2013;60(8):B4689.
35. Polsky D, Doshi JA, Marcus S, et al. Long-term risk for depressive symptoms after a medical diagnosis. *Archives of internal medicine.* 2005;165(11):1260-1266.
36. Drory Y, Kravetz S, Hirschberger G, Israeli Study Group on First Acute Myocardial I. Long-term mental health of men after a first acute myocardial infarction. *Archives of physical medicine and rehabilitation.* 2002;83(3):352-359.
37. Stern S. Symptoms other than chest pain may be important in the diagnosis of "silent ischemia," or "the sounds of silence". *Circulation.* 2005;111(24):e435-437.
38. Lumley MA, Torosian T, Rowland LL, Ketterer MW, Pickard SD. Correlates of unrecognized acute myocardial infarction detected via perfusion imaging. *The American journal of cardiology.* 1997;79(9):1170-1173.
39. Steg PG, James SK, Atar D, et al. ESC Guidelines for the management of acute myocardial infarction in patients presenting with ST-segment elevation. *European heart journal.* 2012;33(20):2569-2619.
40. Ravven S, Bader C, Azar A, Rudolph JL. Depressive symptoms after CABG surgery: a meta-analysis. *Harvard review of psychiatry.* 2013;21(2):59-69.

41. Ladwig KH, Roll G, Breithardt G, Budde T, Borggrefe M. Post-infarction depression and incomplete recovery 6 months after acute myocardial infarction. *Lancet.* 1994;343(8888):20-23.
42. Wiles R. Patients' perceptions of their heart attack and recovery: the influence of epidemiological "evidence" and personal experience. *Social Science & Medicine.* 1998;46(11):1477-1486.
43. Lloyd GG, Cawley RH. Psychiatric morbidity after myocardial infarction. *Q J Med.* 1982;51(201):33-42.
44. Lesperance F, FrasureSmith N, Talajic M. Major depression before and after myocardial infarction: Its nature and consequences. *Psychosomatic Medicine.* 1996;58(2):99-110.
45. Trivedi MH. The link between depression and physical symptoms. *Prim Care Companion J Clin Psychiatry.* 2004;6(Suppl 1):12-16.
46. Macleod J, Davey Smith G, Heslop P, Metcalfe C, Carroll D, Hart C. Psychological stress and cardiovascular disease: empirical demonstration of bias in a prospective observational study of Scottish men. *BMJ.* 2002;324(7348):1247-1251.

APPENDIX of CHAPTER 5.1

Figure 1. Flow chart of study population



RS=Rotterdam Study, Roman numbers denote cohorts and arabic numbers examination waves, e.g. RS I-3 = third examination round of the first cohortMMSE= Mini Mental State Examination

Baseline depressive symptoms analysis – Supplement 1.

Initial level and rate of change of cognitive functioning and single or multiple-episode depression using the Latent growth curve modelling (*N* = 6 221)

Conditional model	Intercept			Slope		
	Estimate	S.E.	*p* value	Estimate	S.E.	*p* value
Depressive symptoms	-0.080	0.01	<0.001	-0.007	0.01	0.18
Sex (female)	0.758	0.14	<0.001	0.180	0.07	0.015
Age	-2.909	0.08	<0.001	-1.047	0.05	<0.001
Cohort	-0.843	0.15	<0.001	0.098	0.08	0.28
Education	2.031	0.08	<0.001	2.043	0.08	0.028

Model fit indices: RMSEA* (95% CI) = 0.039 (0.037; 0.041); CFI** = 0.967

Fully adjusted model	Intercept			Slope		
	Estimate	S.E.	*p* value	Estimate	S.E.	*p* value
Depressive symptoms	-0.050	0.01	<0.001	-0.006	0.01	0.28
Sex (female)	1.081	0.14	<0.001	0.139	0.08	0.078
Age	-2.665	0.08	<0.001	-1.022	0.05	<0.001
Cohort	-0.827	0.15	<0.001	0.120	0.08	0.14
Education	1.954	0.08	<0.001	0.090	0.04	0.037
Body mass index	-0.022	0.02	0.17	-0.013	0.01	0.15
Cancer	0.548	0.24	0.021	-0.138	0.14	0.31
Diabetes mellitus	-0.858	0.29	0.003	-0.529	0.17	0.002
Stroke	-1.523	0.36	<0.001	-0.433	0.26	0.09
Myocardial infarction	0.272	0.33	0.40	-0.04	0.19	0.83
Psychotropic therapy	-0.555	0.18	0.002	-0.004	0.11	0.96
Cardiovascular therapy	-0.156	0.14	0.27	-0.159	0.08	0.044
IADL	1.289	0.11	<0.001	-0.046	0.08	0.58
Smoking habits	0.327	0.08	<0.001	-0.032	0.04	0.43
Alcohol consumption	0.064	0.16	0.68	-0.050	0.10	0.60

Model fit indices: RMSEA* (95% CI) = 0.029 (0.027; 0.030); CFI** = 0.965
* RMSEA: Root Mean Square Error of Approximation **CFI: Comparative Fit Index;***IADL instrumental activity of daily living.

Unconditional Model			
	Estimate	S.E.	*p* value
Factor loadings of latent variable cognition at T1 (baseline)			
LDST	1.000		
VF	0.571	0.010	<0.001
STROOP	-3.629	0.062	<0.001
MMSE	0.215	0.004	<0.001
Intercept and slope of latent variable cognition			
Intercept	0.000	0.000	/
Slope	-2.812	0.143	0.001

Model fit indices: RMSEA* (95% CI) = 0.041 (0.039; 0.043); CFI** = 0.967
* RMSEA: Root Mean Square Error of Approximation **CFI: Comparative Fit Index;

Stratified analysis in to the respect of cohort

Cohort 1 analysis
Flow chart of study population.

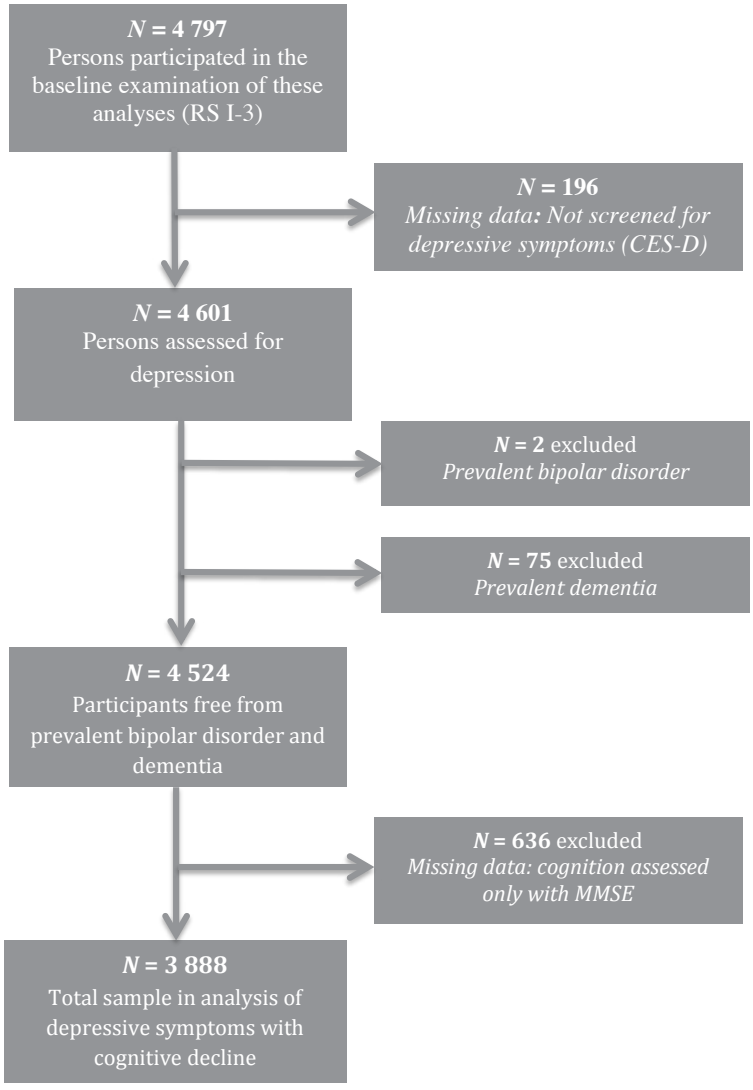


Table 1. Study population by depressive episodes status (*N* = 3 888)

	Depressive episodes		
	No depression *N* = 3143	**Single episode *N* = 499**	**Multiple episodes *N* = 246**
Age, years, mean (SD)	71.3 (6.8)	72.9 (7)	71.9 (6.4)
Sex*, female, *N* (%)	1726 (54.9)	326 (65.3)	192 (78)
Highest education attained,			
Primary, *N* (%)	474 (15.2)	90 (18.3)	49 (20.2)
Lower/intermediate, *N* (%)	1324 (42.5)	219 (44.5)	107 (44)
Vocational, N (%)	970 (31.1)	142 (28.9)	76 (31.1)
Higher, *N* (%)	348 (11.2)	41 (8.3)	11 (4.5)
IADL,** score, mean (SD)	17.7 (0.6)	17.6 (0.9)	17.5 (0.8)
Smoking habits,			
Never smoker, *N* (%)	1043 (33.2)	186 (37.3)	91 (37)
Current smoker, *N* (%)	470 (15.2)	98 (19.6)	51 (20.7)
Past smoker, *N* (%)	1622 (51.6)	215 (43.1)	104 (42.3)
Alcohol consumption,			
No consumption, *N* (%)	316 (10.1)	57 (11.4)	40 (16.3)
Current consumers, *N* (%)	2642 (84.1)	391 (78.4)	175 (71.1)
Past consumers, *N* (%)	185 (5.9)	51 (10.2)	31 (12.6)
Body mass index, mean (SD)	26.8 (3.9)	27.1 (4)	27.2 (4.1)
Cardio vascular medication, *N* (%)	1413 (45)	261 (52.3)	130 (52.8)
Central nervous system medication, *N* (%)	426 (13.6)	141 (28.3)	107 (43.5)
Cancer, *N* (%)	167 (5.3)	30 (6)	22 (8.9)
Diabetes mellitus, *N* (%)	136 (4.3)	24 (4.8)	12 (4.9)
Myocardial infarction, *N* (%)	106 (3.4)	24 (4.8)	5 (2)
Stroke, *N* (%)	109 (3.4)	22 (4.4)	14 (5.7)

Descriptive characteristics of study population were assessed at baseline. Values represent percentage (%), or mean (SD). *Female; **IADL: Instrumental activity of daily leaving score.

Table 2. Descriptive of cognitive performance tests measured over time for depression

Depressive episodes			Cognitive test			
Time point	**T1***		**T2***		**T3***	
	N	**Mean (SD)**	***N***	**Mean (SD)**	***N***	**Mean (SD)**
	Letter-digit substitution task					
No depression	3088	27.9 (7)	2196	25.6 (6.8)	1127	25.1 (8.3)
Single episode	492	25.1 (7.4)	289	24.5 (7.3)	132	25.1 (11.4)
Multiple episodes	239	25.4 (6.5)	174	23.2 (7.1)	83	23.4 (10.2)
	Verbal fluency test					
No depression	3125	21 (5.4)	2264	20.4 (5.4)	1323	19.9 (6.4)
Single episode	497	20.1 (6.1)	302	19.8 (5.1)	140	19.3 (7.5)
Multiple episodes	245	20.4 (5.9)	184	19.4 (5.5)	91	18.5 (5.3)
	Stroop test					
No depression	3048	61.5 (29.1)	2069	64.1 (31.6)	1260	66.9 (35.1)
Single episode	481	68.8 (37.5)	273	69.7 (42.2)	133	69.5 (36.6)
Multiple episodes	239	65.7 (32.8)	165	73.8 (39.2)	88	74 (40.4)
	The mini–mental state examination					
No depression	3142	27.8 (1.8)	2329	27.5 (2.3)	1434	27 (3.4)
Single episode	499	27.4 (2.2)	336	26.9 (2.8)	154	26.6 (3.6)
Multiple episodes	246	27.5 (2)	200	27 (2.6)	108	26.6 (3)

T1: time point one (1997-1999); ***T2:*** time point two (2002-2004); ***T3:*** time point three (2009-2011).

Table 3. Initial level (intercept) and rate of change (slope) of cognitive performance and late-life depression using the Latent growth curve modelling (*N* = 3 888)

Unconditional Model			
Estimate		**S.E.**	***p* value**
Factor loadings of latent variable cognition			
LDST T1	1.000		
VF T1	0.592	0.013	<0.001
STROOP T1	-3.947	0.088	<0.001
MMSE T1	0.235	0.006	<0.001
Intercept and slope of latent variable cognition			
Intercept	0.000	0.000	
Slope	-3.020	0.132	<0.001

Model fit indices: RMSEA* (95% CI) = 0.043 (0.040; 0.046); CFI** = 0.965

Conditional model	**Intercept**			**Slope**		
	Estimate	**S.E.**	***p* value**	**Estimate**	**S.E.**	***p* value**
Depressive episodes						
No depression		reference			reference	
Single episode	-1.017	0.24	<0.001	-0.128	0.16	0.44
Multiple episodes	-0.800	0.33	0.016	-0.661	0.21	0.001
Sex (female)	0.296	0.17	0.080	0.170	0.11	0.107
Age	-2.608	0.08	<0.001	-0.850	0.06	<0.001
Education	1.988	0.10	<0.001	0.110	0.06	0.072

Model fit indices: RMSEA* (95% CI) = 0.038 (0.036; 0.040); CFI** = 0.965

Fully adjusted	**Intercept**			**Slope**		
	Estimate	**S.E.**	***p* value**	**Estimate**	**S.E.**	***p* value**
Depressive episodes						
No depression		reference			reference	
Single episode	-0.836	0.24	<0.001	-0.155	0.17	0.35
Multiple episodes	-0.364	0.33	0.276	-0.718	0.21	0.001
Sex (female)	0.725	0.18	<0.001	0.105	0.11	0.357
Age	-2.375	0.09	<0.001	-0.832	0.07	<0.001
Education	1.901	0.10	<0.001	0.116	0.06	0.058
Body mass index	-0.011	0.02	0.61	-0.007	0.01	0.62
Cancer	0.571	0.34	0.605	-0.291	0.24	0.21
Diabetes mellitus	-0.755	0.39	0.050	-0.533	0.27	0.049
Stroke	-1.718	0.43	<0.001	-0.058	0.35	0.87
Myocardial infarction	0.600	0.46	0.19	0.112	0.29	0.70
Psychotropic therapy	-0.613	0.22	0.006	0.283	0.15	0.065
Cardiovascular therapy	-0.030	0.17	0.861	-0.262	0.11	0.016
iADL	1.258	0.13	<0.001	-0.002	0.10	0.98
Smoking habits	0.291	0.10	0.003	-0.052	0.06	0.38
Alcohol consumption	0.153	0.19	0.43	-0.077	0.14	0.72

Model fit indices: RMSEA* (95% CI) = 0.027 (0.025;0.029); CFI** = 0.964
Model 1: Unconditional model; Model 2: Conditional model; Model 3: Fully adjusted model. * RMSEA: Root Mean Square Error of Approximation **CFI: Comparative Fit Index.

Cohort 2 analysis.

Flow chart of study population.

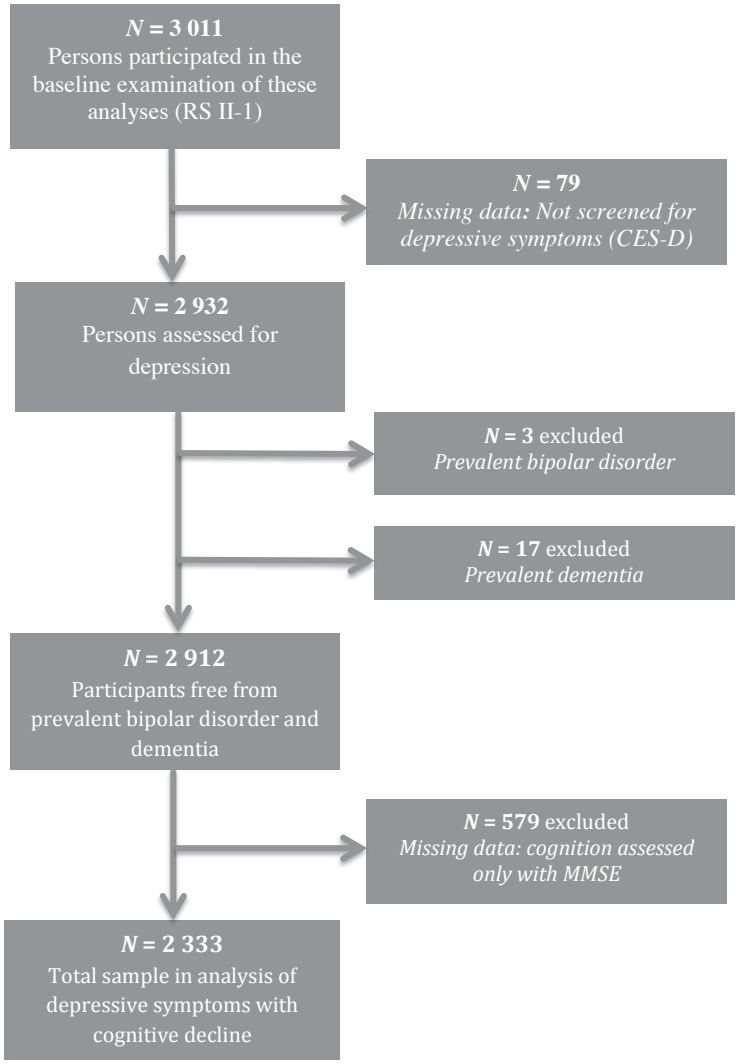


Table 1. Study population by depressive episodes status (*N* = 2 333)

	Depressive episodes		
	No depression *N* = 1977	**Single episode *N* = 252**	**Multiple episodes *N* = 104**
Age, years, mean (SD)	63.6 (7.6)	64.6 (8.6)	64.1 (8.5)
Sex*, female, *N* (%)	1031 (52.1)	169 (67.1)	78 (75)
Highest education attained,			
Primary, *N* (%)	140 (7.1)	32 (12.7)	8 (7.7)
Lower/intermediate, *N* (%)	879 (44.5)	121 (48)	52 (50)
Vocational, *N* (%)	596 (30.1)	54 (21.4)	35 (33.7)
Higher, *N* (%)	362 (18.3)	45 (17.9)	9 (8.7)
iADL,** score, mean (SD)	16.3 (6.3)	16.6 (5.3)	14.9(7.9)
Smoking habits,			
Never smoker, *N* (%)	599 (30.3)	84 (33.3)	28 (26.9)
Current smoker, *N* (%)	372 (18.8)	59 (23.4)	31 (29.8)
Past smoker, *N* (%)	1006 (50.9)	109 (43.3)	45 (43.3)
Alcohol consumption,			
No consumption, *N* (%)	168 (8.5)	31 (12.4)	11 (10.6)
Current consumers, *N* (%)	1703 (86.1)	198 (78.9)	82 (78.8)
Past consumers, *N* (%)	106 (5.4)	22 (8.8)	11 (10.6)
Body mass index, mean (SD)	27.1 (4)	27.1 (4.1)	28.2 (5.3)
Cardio vascular medication, *N* (%)	675 (34.1)	101 (40.1)	39 (37.5)
Central nervous system medication, *N* (%)	205 (10.4)	76 (30.2)	49 (47.1)
Cancer, *N* (%)	219 (11.1)	24 (9.5)	16 (15.4)
Diabetes mellitus, *N* (%)	116 (5.9)	30 (11.9)	8 (7.7)
Myocardial infarction, *N* (%)	106 (5.4)	10 (4)	10 (9.6)
Stroke, *N* (%)	44 (2.2)	16 (6.3)	6 (5.8)

Descriptive characteristics of study population were assessed at baseline. Values represent percentage (%), or mean (SD). *****Female; **iADL: Instrumental activity of daily leaving score.

Table 2. Descriptive of cognitive performance tests measured over time for depression

Depressive episodes			Cognitive test			
Time point	**T1***		**T2***		**T3***	
	N	**Mean (SD)**	***N***	**Mean (SD)**	***N***	**Mean (SD)**
	Letter-digit substitution task					
No depression	1948	29.9 (7.2)	1601	29.3 (7.9)	1209	28 (7.8)
Single episode	247	28.5 (7.2)	196	28.5 (7.2)	130	27.9 (6.5)
Multiple episodes	102	27.8 (7.8)	75	27.1 (7.2)	46	26.3 (7.3)
	Verbal fluency test					
No depression	1969	22.2 (5.4)	1624	21.2 (5.6)	1237	21.7 (6.1)
Single episode	250	22.3 (6.1)	200	21.6 (5.8)	134	22.7 (5.7)
Multiple episodes	102	21.3 (5.7)	77	20.7 (5.1)	47	20 (4.8)
	Stroop test					
No depression	1922	52.6 (23.3)	1454	54.5 (26.3)	1260	56.5 (26)
Single episode	249	57.1 (27.7)	182	58.5 (29.1)	132	53.3 (16.2)
Multiple episodes	101	56.1 (19.4)	66	62.5 (35.7)	48	56.8 (22.9)
	The mini–mental state examination					
No depression	1977	27.8 (1.8)	1669	27.7 (2.1)	1284	27.7 (2.4)
Single episode	252	27.7 (1.9)	210	26.5 (2.4)	140	27.6 (2.2)
Multiple episodes	104	27.4 (2.1)	87	26.5 (3.9)	51	27.3 (2.5)

T1: time point one (1997-1999); ***T2:*** time point two (2002-2004); ***T3:*** time point three (2009-2011).

Table 3. Initial level (intercept) and rate of change (slope) of cognitive performance and late-life depression using the Latent growth curve modelling (N = 2 333)

Unconditional Model			
Estimate		**S.E.**	***p* value**
Factor loadings of latent variable cognition			
LDST T1	1.000		
VF T1	0.565	0.018	<0.001
STROOP T1	-3.161	0.093	<0.001
MMSE T1	0.208	0.007	<0.001
Intercept and slope of latent variable cognition			
Intercept	0.000	0.000	
Slope	-2.075	0.131	<0.001

Model fit indices: RMSEA* (95% CI) = 0.040 (0.037; 0.045); CFI** = 0.974

Conditional model	**Intercept**			**Slope**		
	Estimate	**S.E.**	***p* value**	**Estimate**	**S.E.**	***p* value**
Depressive episodes						
No depression		reference			reference	
Single episode	-0.616	0.33	0.061	-0.026	0.15	0.87
Multiple episodes	-1.619	0.50	0.001	-0.577	0.24	0.018
Sex (female)	1.330	0.22	<0.001	0.213	0.10	0.029
Age	-2.526	0.11	<0.001	-1.057	0.06	<0.001
Education	2.039	0.13	<0.001	0.078	0.06	0.17

Model fit indices: RMSEA* (95% CI) = 0.037 (0.033; 0.041); CFI** = 0.973

Fully adjusted model	**Intercept**			**Slope**		
	Estimate	**S.E.**	***p* value**	**Estimate**	**S.E.**	***p* value**
Depressive episodes						
No depression		reference			reference	
Single episode	-0.106	0.33	0.747	0.039	0.16	0.80
Multiple episodes	-0.865	0.50	0.083	-0.451	0.50	0.072
Sex (female)	1.465	0.22	<0.001	0.184	0.10	0.071
Age	-2.280	0.11	<0.001	-1.034	0.06	<0.001
Education	1.951	0.13	<0.001	0.060	0.06	0.295
Body mass index	-0.031	0.03	0.22	-0.017	0.01	0.16
Cancer	0.368	0.32	0.25	-0.040	0.15	0.79
Diabetes mellitus	-1.092	0.41	0.008	-0.512	0.22	0.017
Stroke	-1.460	0.63	0.021	-1.052	0.36	0.003
Myocardial infarction	-0.080	0.46	0.86	-0.190	0.23	0.41
Psychotropic therapy	-0.631	0.30	0.036	-0.304	0.14	0.035
Cardiovascular therapy	-0.403	0.23	0.075	-0.045	0.11	0.674
iADL	1.339	0.21	<0.001	-0.112	0.12	0.36
Smoking habits	0.266	0.12	0.025	-0.003	0.05	0.95
Alcohol consumption	-0.071	0.26	0.79	-0.003	0.13	0.98

Model fit indices: RMSEA* (95% CI) = 0.026 (0.024;0.029); CFI** = 0.973
Model 1: Unconditional model; Model 2: Conditional model; Model 3: Fully adjusted model. * RMSEA: Root Mean Square Error of Approximation **CFI: Comparative Fit Index.

Appendix of Chapter 5.2

Supplementary material 1. All analyses reported in the manuscript were rerun excluding all participants with recognized MI that did not meet electrocardiography criteria for prior Q-wave MI.

Table 1. Baseline characteristics of the study population (*N* = 3 940)

	History of myocardial infarction					
	Men (*N* = 1 823)			Women (*N* = 2 214)		
	No MI (*N*=1488)	Recognized MI (*N*=144)	Unrecognized MI (*N*=115)	No MI (*N*=2057)	Recognized MI (*N*=32)	Unrecognized MI (*N*=104)
Age, years, mean (SD)	67.6 (6.9)	70.2 (6.7)	70.1 (7.2)	68.2 (7.6)	73.4 (8.9)	71.8(7.9)
Depressive symptoms at baseline, score, mean (SD)*	4.3 (4.7)	4.8 (5.4)	4.3 (5.6)	5.4 (5.9)	7.3 (5.2)	6.1 (5.9)
Smoking status,						
Never smoker, *N* %	132 (8.9)	9 (6.4)	7 (6.4)	944 (48)	16 (50)	58 (56.9)
Past smoker, *N* %	1041 (70.6)	106 (71.6)	78 (71.6)	658 (33.5)	13 (40.6)	28 (27.5)
Current smoker, *N* %	302 (20.5)	28 (22)	24 (22)	363 (18.5)	3 (9.4)	16 (15.7)
Alcohol consumption,						
Never consumer, *N* %	82 (5.5)	19 (13.2)	3 (2.6)	293 (14.5)	10 (32.3)	19 (19)
Past consumer, *N* %	100 (6.8)	16 (11.1)	11 (9.6)	345 (17.1)	5 (16.1)	24 (24)
Current consumer, *N* %	1299 (87.7)	109 (75.5)	101 (87.8)	1378 (68.4)	16 (51.6)	57 (57)
MMSE, score, mean (SD)	28.2 (1.2)	28.2 (1.2)	28.1 (1.4)	28.1 (1.3)	27.9 (1.1)	28.2 (1.3)
BMI, kg/m2, mean (SD)	25.8 (2.9)	26.2 (3.1)	26.3 (3.2)	26.6 (3.9)	27.8 (4.2)	27.6 (4.5)
Systolic blood pressure, mmHg, mean (SD)	139.8 (21.6)	137.6 (21.3)	146.3 (21)	140 (21.9)	132.4 (16.6)	146.3 (23.4)
Blood pressure lowering medication, *N* %	274 (19.4)	50 (37.9)	22 (20.2)	521 (26.4)	9 (33.3)	30 (30.3)
History of cardiovascular diseases, *N* %	80 (5.4)	133 (92.4)	99 (86.1)	61 (3)	31 (96.9)	86 (82.7)
History of Stroke, *N* %	39 (2.6)	7 (4.9)	7 (6.1)	32 (1.6)	2 (6.2)	3 (2.9)
History of CABG, *N* %	36 (2.5)	38 (27.5)	0 (0)	9 (0.4))	5 (15.6)	0 (0)
History of PCI, *N* %	10 (0.7)	7 (5)	0 (0)	6 (0.3)	2 (6.2)	1 (1)
Diabetes mellitus, *N* %	121 (8.1)	24 (16.7)	17 (16.5)	162 (7.9)	4 (12.5)	11 (10.6)
Highest education attained,						
Low education, *N* %	282 (19)	36 (25)	26 (23.2)	729 (35.5)	16 (50)	35 (34)
Intermediate education, *N* %	935 (63)	87 (60.4)	67 (59.8)	1209 (58.9)	18 (46.2)	61 (59.2)
High education, *N* %	268 (18)	21 (14.6)	19 (17)	113 (5.5)	0 (0)	7 (6.8)
Marital status,						
Never married or divorced, *N* %	57 (4.6)	6 (5.7)	6 (7)	249 (14.3)	3 (12.5)	7 (8.1)
Married or living together, *N* %	1055 (84.7)	89 (84.8)	66 (76.7)	970 (55.7)	7 (29.2)	41 (47.7)
Widowed, *N* %	133 (10.7)	10 (9.5)	14 (16.3)	522 (30)	14 (58.3)	38 (44.2)

Table presents complete data, missing values are not imputed here.
*CES-D/HADS-D z-score.

Table 2. Recognized and unrecognized MI and all-cause mortality risk with Cox regression (*N* = 3 940)

	All-cause mortality	
	HR (95% CI)	***p* value**
Men (*N* = 1 747)		
Recognized MI (*N* = 144)		
Multivariate adjusted*	1.99 (1.64;2.42)	<0.001
Unrecognized MI (*N* =115)		
Multivariate adjusted*	1.59 (1.27;1.98)	<0.001
Women (*N* = 2 193)		
Recognized MI (*N* = 32)		
Multivariate adjusted*	1.97 (1.32;2.94)	0.001
Unrecognized MI (*N* =104)		
Multivariate adjusted*	1.00 (0.77;1.30)	0.97

2 148 (1087 men) of 3 940 persons died during 55,170 person-years of follow-up.
*Multivariate adjusted: additionally adjusted for age, level of education, smoking status, alcohol consumption, history of stroke, diabetes mellitus and systolic blood pressure.

Table 3. The longitudinal association of recognized or unrecognized MI with depressive symptoms in men (*N* = 1 747)

		Depressive symptoms						
		First assessment§ (*N* = 1 228)			Second assessment§ (*N* = 1 073)		Combined analysis (*N* = x, with 2 486 observations)	
Exposure	***N***	**B (95% CI)**	***p* value**	***N***	**B (95% CI)**	***p* value**	**B (95% CI)**	***p* value**
Recognized MI								
Adjusted for age and baseline depressive symptoms*	102	1.25 (0.16;2.34)	0.024	61	1.18 (-0.35;2.7)	0.13	1.17 (-0.23;2.56)	0.10
Multivariate adjusted**	102	1.38 (0.30;2.47)	0.013	61	1.11 (-0.41;2.63)	0.15	1.23 (-0.16;2.61)	0.08
Unrecognized MI								
Adjusted for age and baseline depressive symptoms*	84	0.85 (-0.33;2.04)	0.16	51	0.50 (-1.18;2.18)	0.56	0.62 (-0.44;1.67)	0.25
Multivariate adjusted**	84	0.79 (-0.40;1.97)	0.19	51	0.54 (-1.14;2.21)	0.53	0.55 (-0.51;1.62)	0.31

Analysis were performed with linear regression and generalized estimating equations.
§ CES-D mean score (SD); **First assessment**: Recognized MI 4.6 (8.9), Unrecognized MI 3.8 (6.5) and No MI 2.9 (5.3); **Second assessment:** Recognized MI 6.2 (7.7), Unrecognized MI 5.4 (5.9) and No MI 5.1 (6.3)
*Model adjusted for age and baseline depression (corrected for baseline depressive symptoms);
**Multivariate adjusted: additionally adjusted for level of education, marital status, body mass index, smoking status, alcohol consumption, and history of stroke, diabetes mellitus and systolic blood pressure

Table 4. The longitudinal association of recognized or unrecognized MI with depressive symptoms in women (*N* = 2 193)

		Depressive symptoms						
		First assessment[§] (*N* = 1 819)			Second assessment[§] (*N* = 1 498)		Combined analysis (*N* = x, with 3 317 observations)	
Exposure	*N*	B (95% CI)	*p* value	*N*	B (95% CI)	*p* value	B (95% CI)	*p* value
Recognized MI								
Adjusted for age and baseline depressive symptoms*	24	0.15 (-2.61;2.91)	0.92	11	3.28 (-1.21;7.78)	0.15	0.45 (-2.55;3.44)	0.77
Multivariate adjusted**	24	-0.18 (-2.95;2.58)	0.90	11	3.80 (-0.78;8.38)	0.10	0.37 (-2.60;3.34)	0.81
Unrecognized MI								
Adjusted for age and baseline depressive symptoms*	84	0.12 (-1.38;1.62)	0.88	62	0.57 (-1.37;2.50)	0.57	0.23 (-1.18;1.64)	0.75
Multivariate adjusted**	84	-0.02 (-1.52;1.48)	0.98	62	0.44 (-1.48;2.63)	0.65	0.14 (-1.28;1.55)	0.85

Analysis were performed with linear regression and generalized estimating equations.

[§] CES-D mean score (SD); **First assessment**: Recognized MI 6.8 (9.6), Unrecognized MI 5.9 (8.2) and No MI 4.9 (7.2); **Second assessment:** Recognized MI 11.2 (8.7), Unrecognized MI 9.1 (8.4) and No MI 7.6 (8.3)

*Model adjusted for age and baseline depression (corrected for baseline depressive symptoms);

**Multivariate adjusted: additionally adjusted for level of education, marital status, body mass index, smoking status, alcohol consumption, and history of stroke, diabetes mellitus and systolic blood pressure.

CHAPTER 6

General Discussion

R1
R2
R3
R4
R5
R6
R7
R8
R9
R10
R11
R12
R13
R14
R15
R16
R17
R18
R19
R20
R21
R22
R23
R24
R25
R26
R27
R28
R29
R30
R31
R32
R33
R34
R35
R36
R37
R38
R39

In this thesis I present evidence for the possible role of several biological factors and physical health problems in the aetiology of depression. The studies described in the previous chapters were performed mostly within the frame of the Rotterdam Study (RS), a prospective population-based data-collection project. One of the studies in this thesis is a consortium based meta-analysis comprising eleven cohort studies from the Cohorts for Heart and Aging Research in Genomic Epidemiology (CHARGE) consortium and another study comprises data from the Erasmus Rucphen Family Study. The details of the conducted research, methods, analysis, and results have already been described in detail. In this chapter, I would like to give an overview of the key findings and emphasize their possible clinical implications. Studying mental health diseases such as depression is a challenge and faces a series of concerns that were already discussed in some detail. In this chapter I will also elaborate on the most important methodological issues that I encountered conducting this work. Finally, I will conclude and describe some perspectives and strategies for future research.

Main findings

Potential biomarkers for depression

Clinically translatable biomarkers in psychiatry are crucial for efficient diagnosis, therapy, and prognosis of a psychiatric disease.[1] Detection of biomarkers specific for depression is necessary for the psychiatric practice. Yet, validated clinically relevant biomarkers specific for depression have been very hard to identify.[2] In other medical disciplines, such as endocrinology, conditions like diabetes mellitus[3] are clinically diagnosed, predicted and the severity of the disease is defined by a well-established set of biomarkers. Unlike endocrinology, in psychiatry we lack such laboratory tools and non-invasive blood tests to diagnose and predict depression. All biological parameters that can be repeatedly measured with precision and are related to the occurrence of depression are considered potential biomarkers for depression.[2] Yet, biomarkers could be factors that are causally related to depression or are good predictors of depression (**Figure 1**). They may also be biological markers that just allow a better discrimination between depressed and non-depressed patients, and thus constitute diagnostic biomarkers (**Figure 1**). Therefore, the full definition for biomarkers (biological markers) postulated by the World Health Organization includes any substances, structures, or processes that can be objectively measured in the body or its products and influence or predict the incidence of outcome or disease.[4]

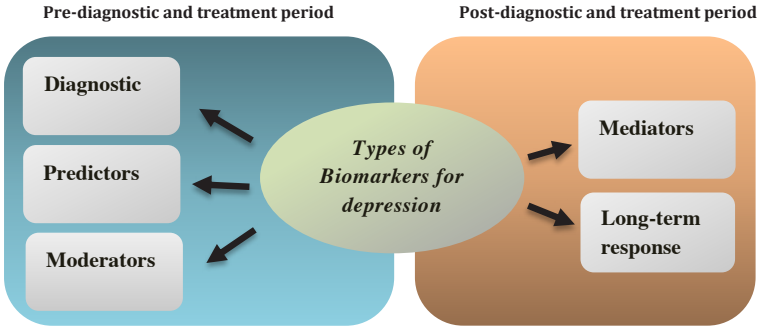
Epidemiological longitudinal designs have a unique role in refining or rejecting causal hypotheses[5] on the relation of certain biological risk factors with depression. Longitudinal associations may thus imply or dismiss possible biomarkers for depression. In **Chapter 2** we study whether

aetiological candidate markers, including blood extracted vitamin D serum levels (neuroendocrine factor), or three neuro-inflammatory markers (interleukin (IL)-6, alpha-1-antichymotrypsin (ACT), and C reactive protein (CRP)) are risk factors for depression.

In order to re-examine the extensively studied possible causal relation between vitamin D serum levels and depression, in **Chapter 2.1** we addressed the association between these two factors. Both the cross-sectional and the prospective association between vitamin D serum levels and subsequent change of depressive symptoms or incident major depressive disorder (MDD) were explored. Few studies have tested these associations. Prior longitudinal analyses produced inconsistent results that were mostly explained by small sample size, incomplete definition of depression, confounding, and short length of follow-up.[6-8] We observed no longitudinal association between vitamin D and depression after some of the limitations of previous studies, such as poor confounder control, were addressed. All authors should be careful when commenting on the presence or absences of causality, but our careful analysis sheds doubt on the causal relationship between vitamin D and depression. Yet, we observed a negative cross-sectional association between vitamin D and depression. Thus, we did speculate about a possible role of vitamin D serum levels as a biomarker of depression.

Biological factors need not be causally related to depression in order to be a biomarker for depression.[2] A marker that differs between those with depression and those without depression need not to be involved in the causal pathway of depression to be a diagnostic biomarker (see **Figure 1**). To optimally establish a temporal relationship, the longitudinal design should include repeated assessments of the biomarker of interest over a significant period of time.[9] Indeed, longitudinal designs such as the one we use, with a single measure of the biomarker at baseline are not ideal designs and are vulnerable to random measurement error.[10] Moreover, the long follow-up (almost 12 years follow-up time in the current study) may additionally bias the results by introducing regression dilution bias.[11] The longitudinal design we used is limited to determine the directionality of the association. Yet, I argue that the power of the cross-sectional analysis should also not be underestimated to capture two correlates. Vitamin D deficiency and depression are unquestionably correlated but vitamin D does not have the role of a long-term risk factor for depression.

Figure 1 ©taken and adapted from Gadad at al.[2] **Types of Biomarkers for depression**.



Biomarkers can be defined as biomarkers used before diagnostics and treatment initiation, and are classified as diagnostic, predictive, or moderators. Diagnostic biomarkers classify an patient with depressive disorder, predictive biomarkers determine overall likelihood of response/remission, and moderators determine likelihood of response/remission with a particular treatment. Mediators are biomarkers collected soon after treatment initiation and help predict overall likelihood of response/remission. Long-term treatment response biomarkers are factors that may also be indicative of ultimate outcome.

I doubt that vitamin D can be classified as a biomarker with clinical relevance for depression. Nowadays, depression is being diagnosed by self-reported depressive symptoms diagnosed by clinician's criteria such as the DSM criterion.[12] This nosologic method of depression diagnosis is an obstacle in further clinical validation of vitamin D as a predicting biomarker. Venkatasubramanian G. commented on biological factors with discriminating power but low or no predicting power: "If in heart research we look for characteristics that would help classify chest pain patients into those with and without coughing, we will get nowhere".[1] "On the other hand, testing specific enzymes elevations in persons with specific ECG changes compared to persons with no ECG changes is more informative and might bring more meaningful clinical discrimination".[1] In this notation, I rather see vitamin D deficiency as a biomarker of conditions such as skeletomuscular diseases, osteoporosis, increased risk of falls with subsequent fractures in patients with chronic psychiatric disease such as depression, than as a biomarker specific for depression.[13]

Biomarker research has focused and worked on identifying potential peripheral proteins of interest such as inflammatory markers.[14] Tumor necrosis factor alpha (TNF-alfa), interleukin-6 (IL-6), interleukin-1Beta (IL-1Beta), and C-reactive protein (CRP) are just some of the proteomic markers that have been associated with depression.[14] In **Chapter 2.2** we investigated both the cross-sectional and the longitudinal association between IL-6, alpha-1-antichymotrypsin (ACT), and CRP and depressive symptoms. As previously observed in other studies[2,14], this study provides evidence for IL-6 and CRP as potential predictive biomarkers for depression. More specifically, we found that higher levels of IL-6 and CRP are longitudinally associated with incident depressive symptoms over five years follow-up. Even though many large epidemiologic studies reported cross-sectional associations between these inflammatory markers and depression[15-17], we

observed no cross-sectional association in our study. This may be attributed to the small-sample size of the current study. The absence of the cross-sectional association in this study suggests that IL-6 and CRP are measures of a process or a product of key pathophysiological stages that lead to depression (Predictive biomarkers see **Figure 1**).[2]

The fact that we observed a longitudinal but no cross-sectional associations between these inflammatory proteins and depressive symptoms may also indicate that the brain-response with subsequent depression precipitated by inflammatory proteins occurs only if this inflammatory stimulus is chronic/persistent; and does not occur in an acute and transitional response of the inflammatory system.[18] Indeed, ACT is a known acute phase plasma protein. If depressive symptoms occurred as a response to the present acute inflammation, we would expect that elevated levels are associated with more depressive symptoms. However, we observed neither a cross-sectional or a longitudinal association between elevated levels of ACT and depressive symptoms. Earlier clinical studies reported elevated ACT in depression, for example among 21 male depressed patients compared to healthy control subjects.[19] However, this study was based on clinical samples of severe cases of depression (often with other comorbidities). Alternatively, ACT could be a less sensitive inflammation marker in the general population and signal early inflammation.

To conclude, the results of this study confirm the hypothesis of depression occurring rather in response to chronic/persistent and not to acute inflammatory processes. It could have been interesting to investigate whether persistently elevated inflammatory markers contribute the occurrence of depressive symptoms. However, the data to explore this question were not available. Yet, previously published reports show that persistent elevation of inflammatory markers contribute to the development of subsequent common mental disease such as depression.[20]

Depression is a syndrome with a high recurrence rate[21], more than a well-defined disease. Also, depression describes a heterogeneous population including those that experience depressive symptoms once and those that suffer depressive symptoms persistently. The association between inflammatory markers and specific subgroups of depression such as those with persistent depressive symptoms has not been established yet. Therefore, we also tested whether IL-6, CRP, and ACT may also have the power to predict the persistence of depressive symptoms. We found that IL-6 and CRP are not only predictive of the occurrence of depressive symptoms but are also independent predictors of the persistence of depressive symptoms. Therefore, these inflammatory biomarkers may be not only meaningful predictive biomarkers of depression but it should be tested whether they can be used to discriminate subtypes or the course of depression.[2]

Epigenetic findings

Peripherally measured proteins and vitamin D were not the only candidate biomarker of depression I investigated. Variations in a gene, set of genes, and other molecular factors are other potential diagnostic and prognostic biomarkers of depression.[1] Depression is a complex multifactorial disorder substantially affected by both environmental and genetic factors.[22] The genetic contribution to the development of depression has been estimated to be almost 40% and environmental and other factors contribute the additional 60%.[22] However, the knowledge of the interacting processes between the genetic and environmental factors is incomplete. A few studies suggest that epigenetic mechanisms might play an important role in the pathophysiology of depression and capture the genetic environment interection.[22-25] DNA methylation is an epigenetic mechanism in which methyl groups are added to the DNA as a response to an environmental or genetic influence. DNA epigenetic markers may play important role in the pathophysiology of psychiatric diseases but are also interesting targets for clinically translatable biomarker of depression (see **Figure 2**).[1] Yet, the epigenetic method is a novel genomic tool and whether the epigenetic alterations can be used as a biomarker is not yet clear and must be carefully studied. In an intense collaborative effort, we studied possible epigenetic markers of depressive symptoms in a population-based setting.[2]

The aim of **Chapter 4** was to study and map gene-environment interaction in depression by identifying epigenetic alterations associated with depressive symptoms. We performed the first and largest (*N* = 11 256) epigenome-wide association population-based study of depressive symptoms to date. We found three epigenetic markers associated with depressive symptoms (hypermethylation of cg04987734 annotated on the *CDC42BPB* gene; cg12325605 annotated on the *ARHGEF3* gene; and cg14023999 an intergenic epigenetic marker. The genes on which these epigenetic marks are located, are known to be involved in the regulation of serotonin and dopamine levels in the brain[26], sphingolipid metabolism[27], and immune response[22]. Moreover, the gene pathway analyses of our findings have shown that these genes are involved in synaptic plasticity, inflammatory mechanisms, regulation of monoamine levels in the brain, axon guidance, and growth factor signalling. These are all mechanisms previously associated with psychiatric disorders including depression.[22,28-30] Certain expression mechanisms of these genes, RNA transcription, protein function and structures, could also be involved in the pathophysiological mechanisms underlying depression.[22] Indeed, RNA expression analysis within the frame of our study have shown that differentially expressed genes reveal one significant pathway; the immune system (more specifically, expression of these genes was involved in the adaptive immune system).

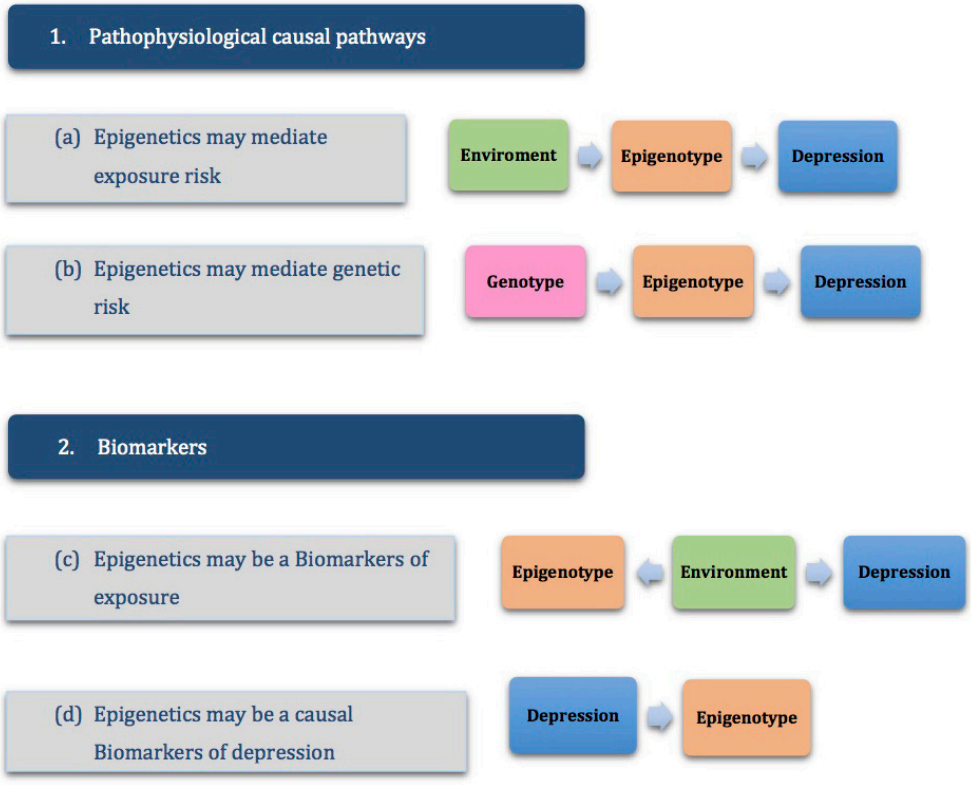
This study was performed within a multi-ethnic population based sample of approximately 11 000 individuals. This epidemiologic design has several strengths and limitations, which were already

discussed in **Chapter 4.** Here I would like to emphasize few considerations. Most importantly, the cross-sectional design of the study limits the interpretation of the associations. I speculate that that there are few possible ways to interpret these findings (Illustrated in **Figure 2**). First (**Figure 2.(a)**), blood DNA hypermethylation of cg04987734, cg12325605, and cg14023999 may occur due to environmental influences, stress events, or early childhood trauma preceding the occurrence of depressive symptomes.[22] Second (**Figure 2.(b)**), the epigenetic alterations may also be mediators of genetic risk for depressive symptoms. In this scenario, the epigenetic mark mediates the causal pathway from genes to depression. Many DNA methylation differences can be partially traced back to genetic variations, suggesting that differentially methylated CpG sites serve as evolutionarily established mediators between the genetic code and phenotypic variability.[31] These epigenetic DNA imprinting's may reflect genetic mechanisms such as silencing of the gene or gene expression[22,32], that lead to depression. Third (**Figure 2.(d)**), depression itself could be the cause of the epigenetic changes. Via endocrinological, immune or other metabolic changes secondary to depression, methylation or demethylation of the DNA may occur.

These proposed pathophysiological mechanisms, however, could not be confirmed by the study presented in this thesis and describing both causality and the direction of causality should be a focus of future research. Finally, these epigenetic marks may not involve in the pathophysiological mechanism underlying or arising from depression and thus present "only" biomarkers of depression. The cross-sectional association between the epigenetic markers and depressive symptoms may also be confounded by large number of environmental factors (**Figure 2.(c)**). The analyses presented in this study were controlled for the effect of factors such as age, sex, smoking habits, and antidepressant medication. Yet, confounding by unmeasured environmental factors is possible. Therefore, if the association between the blood DNA methylation markers and depressive symptoms were explained by these unmeasured environmental factors, then such markers would present biomarker of exposure only.

Since we can only speculate about the directionality, this study is certainly only an indicator that blood methylation of the three identified CpG sites are a potential tool to discriminate individuals suffering from depression from those who do not.

Figure 2. Possible interpretations of the epigenetic epidemiologic findings.



Genetic findings

Depression is a complex, heterogeneous, and common disease with modest heritability.[22] A number of common single nucleotide polymorphisms were found to be associated with depression by DNA microarray-based whole genome-wide association studies (GWAS).[33] However, it has been suggested that common variants have small effects and do not completely account for the heritability. Moreover, GWAS studies are non-informative of rare variants that outnumber the common variants.[34] The next generation sequencing is a new method for identifying rare novel variants that might fulfil the genomic heritability gap.[34] Rare variants in exonic or coding regions that affect protein function, are probably important for depression.[34] Therefore, identifying rare variants opens new hopes for genetic research of depression and psychiatric research in general. In **Chapter 3** we have performed exome sequencing studies to identify rare unknown variants with strong effects contributing to the development of depressive symptoms.

In **Chapter 3.1** we performed the first large-scale population-based high-coverage exome sequencing study. We found a significant association of depressive symptoms with a rare variant (rs77960347-G; MAF = 1%, *N* (carriers) = 60; β = 7.2; *p value* = 5.2×10^{-08}) on chromosome 18. The large estimated effect of this variant on depressive symptoms suggesting that each allele yields an increase of depressive symptoms by 7.2 points. Rs77960347 is a highly conserved missense variant in the *LIPG* gene (*Asn396Ser*). Moreover, it is predicted to be damaging (polyPhen = 1) for the *LIPG* gene. *LIPG* is a gene predicted to be involved in enzymatic function in steroid biosynthesis, cholesterol biosynthesis and thyroid hormone metabolic processes.[35] These results were replicated in three independent samples. Also, within the Rotterdam Study, we observed that carriers of such missense non-synonymous variant have a higher risk of Alzheimer's disease and more white matter lesions.

The heterogeneity and the complexity of depression is a challenge that genetic research constantly faces.[36] To overcome heterogeneity of depression, enlarging the studied sample sizes is widely used as a methodological strategy.[36] Indeed, the number of common variants with small effects identified by GWAS associated with depression increases by expanding the sample size.[37] To identify rare variants associated with complex trait such as depression large sample sizes are required, like in studies of common variants.[34] Using isolated cohorts with less genetic variance, and similar environmental, and cultural characteristics is another strategy to overcome the limited power inherent to this approach.[38] Rare variants accumulated in isolated populations can counterbalance problems of studying heterogeneity. Therefore, in **Chapter 3.2** we performed exome-sequencing and exome-chip genotyping in a genetically isolated European population to identify rare variants associated with depressive symptoms. We found an association of a rare non-synonymous variant in the *NKPD1* gene with depressive symptoms (the missense variant that showed the strongest association within the *NKPD1* gene was rs75291769/exm1479956, *p value* = 2.5×10^{-04} with a frequency of 4.6% in the discovery sample). These variants explained 0.9% of the age- and sex-adjusted variance and 3.8% of heritability of depressive symptoms in the isolated population. Importantly, these results were replicated in the Rotterdam Study. The *NKPD1* gene is involved in the biosynthesis of sphingolipids that was earlier proposed as a therapeutic target for depression.[39]

These two studies are among the first studies to use exome-sequencing on a large scale with a specific aim to discover novel risk genes for depressive symptoms and therefore uncover pathways that lead to the development of depression. The strengths of this approach have already been discussed in the previous chapters. However, the interpretation of the achieved results has to carefully consider several issues.

First, a question that arises when looking in to the results of these studies is whether the observed depression-associated rare variants are really depression specific. Psychiatric disorders have a great overlap in their genetic architecture and it is known that many of the genetic associations are common to all schizophrenia, bipolar disorder and depression.[37,40,41] The studies performed in **Chapter 3** were performed in a population-based setting where the prevalence of diseases such as bipolar disorder and schizophrenia are low (< 2%).[42,43] Thus, we were not able to test the independence of the reported associations from other mental disorders. Yet, in **Chapter 3.1** we tested the statistical difference between carriers of the *Asn396Ser* variant and non-carriers and their risk of suffering Alzheimer's disease. We observed carriers of the *Asn396Ser* variant to be more likely to suffer from Alzheimer's disease compared to non-carriers. Adjusting the association between this rare variant and depressive symptoms for prevalent Alzheimer disease did not change the results. Therefore, we speculate that depressive symptoms could mediate the pathway between this *Asn396Ser* variant and Alzheimer's disease but we cannot reject the possibility that this variant contributes to the development of Alzheimer's as well, i.e. it has pleotropic effects.

Second, the rare variants may be population specific.[34] Rare variants identified in isolated populations such as the one we study in **Chapter 3.2,** could be extremely rare in other populations or absent.[44] We replicated the results of our studies in one population-based sample; however, the samples we used for replication have a similar ancestry. Thus, replication in a more distinct population sample than the discovery sample might be more appropriate.[34] This issue challenges the generalizability of the results to other populations. Yet, we argue that rare variants need not be population specific since rare variants associated with other phenotypes such as Alzheimer's disease were replicated across populations successfully.[45]

Third, even though we used high coverage sequencing in the study presented in **Chapter 3.2** Exome-chip sequencing is not capable in detecting rare variants with a minor allele frequency lower than 0,004%.[46] Exome-sequencing captures only the exomes (~1% of the entire genome) and hence does not capture all rare variation in the genome and that can be achieved using whole genome sequencing. Further other kinds of variants e.g. structural variants cannot be reliably detected using exome-sequencing. Thus, we cannot exclude the possibility that other rare coding variants contribute to depression as well.

Taking all the limitations in to consideration, the results of these two studies present strong evidence that the identified rare variants in the *LIPG* and the *NKPD1* genes are involved in the occurrence of depressive symptoms. Although the variants are rare among populations, the estimated effect on the risk of depression was meaningful. This suggests that we discovered new risk factors for depression providing new insights in the biology of depression.

Physical Health factors for depression

Living with chronic disability and disease such as cardiovascular disease or dementia is difficult, challenging, and decreases the quality of life significantly.[47] These diseases develop slow, they are chronic, require constant treatment, and are perceived by the individual as life threatening.[47] In both, clinical and epidemiological settings high co-occurrence of such chronic diseases with depression was observed.[48,49] Depression also often has a recurrent character and can take a chronic course. The relation between physical health problems and psychiatric disease is complex and the underlying biology is poorly understood. Whether depression is a cause or a consequence of other chronic conditions often remains questionable. Addressing the question of "the chicken or the egg" may help understand why depressed individuals have a substantial increased burden of diseases, physical disability, and mortality. An answer may improve BETTER management and therapy for depression.[50] Moreover, studying the longitudinal relations between depression and diseases such as myocardial infarction and dementia may also elucidate the pathophysiology and provide clues to aetiological mechanisms. The Rotterdam Study setting allowed us to construct advanced longitudinal designs and address some of these questions. Therefore, In **Chapter 5** we took the opportunity to study the impact of depression on neurodegenerative processes such as cognitive decline as well as the impact of chronic diseases such as myocardial infarction on the occurrence of depression.

In **Chapter 5.1** we studied the longitudinal association between depression and cognitive decline. This subject was widely studied and produced contradictory results.[49] Characteristics of depression like the age of first onset, chronicity, duration, number of episodes in the past may predict the longitudinal association between depression and subsequent cognitive decline. Therefore, we followed participants who experienced a single, multiple or no depressive episodes for their cognitive decline over, on average, 12 years to evaluate whether repeated episodes of depression predicted cognitive decline. The simple descriptive analysis showed that most persons experience cognitive decline over time in all the cognitive tasks irrespective of prior depression (See **Table 2, Chapter 5.1** Mini-mental state examination). These results only confirmed the cognitive decline with aging. Moreover, we also observed that at baseline cognitive performance is worse among participants, who had had depression, compared to those with no past episodes of depression (See **Table 2., Chapter 5.1**). Next, we used three repeated measures of four different cognitive domains to construct a latent growth curve of cognitive change. With this advanced approach of modeling repeated measures we evaluated the course of cognitive decline. This approach accounts for the within-person and the between-person change in cognitive functioning over time and was additionally controlled for an extensive list of potential confounders such as age, sex, education and other life-style, demographic factors, and comorbidities that may affect the relation between depression and cognitive decline. As previously reported by others, we observed that depression

and cognitive decline are co-occurring conditions. However, our longitudinal analysis showed that depression contributes significantly to worse long-term cognitive performance only if experienced multiple times. Persons with single or those with no experience episodes of depression had similar cognitive decline.

This study provides evidence that only depression with a recurrent course contributes to a faster cognitive decline than the cognitive decline we normally experience due to aging. This is in line with studies that showed prominent neurological vulnerability among patients suffering recurrent depression.[51,52] A previous study showed that persons with recurrent depressive symptoms have a higher risk of dementia and Alzheimer's disease, reduce hippocampal volume, and are characterized by a cognitive deficit prodrome.[53] We can only speculate that recurrent depression precedes cognitive deficits that eventually evolve into dementia or Alzheimer's disease. Participants with prevalent dementia did not take part in this study. However, I argue that reversed causality could still explain our results. Dementia and Alzheimer's disease are extremes that arise with minor cognitive deficits prior to the occurrence of depression. The baseline cognitive performance was lower among participants with multiple-episode depression compared to those with single-episode or no depression across all cognitive domains. Indeed, there is a limited evidence that milder cognitive performance may precede the occurrence of consequent depression.[54] However, such a scenario should be less likely since the tested models were corrected for baseline general cognitive performance. On the other hand, we were unable to explore whether trajectories with a steeper cognitive decline contributed to the occurrence of or the recurrence of depression. Overall, we conclude that depression and cognitive decline are comorbid conditions. Depression predicts a higher risk of worse long-term cognitive performance only if the experienced depression occurs multiple times.

In **Chapter 5.2** we explored the long-term impact of experiencing health condition such as myocardial infarction (MI) on the occurrence of depression. We have contrasted the impact of a clinically recognized MI to that of a clinically unrecognized MI (without experience of any symptoms and absence of knowledge for the disease) on depression and compared both to the depression risk of persons with no MI experience. Comparing these three groups was an exclusive opportunity to distinguish the biological pathophysiological mechanisms behind the association of MI and depression from the psychological mechanism. The occurrence of depression in someone with an MI with a clear clinical presentation is the result of both psychological/reactive and biological/pathophysiological mechanism. In contrast, those having had a clinically unrecognized MI should develop depression only due to presence of a biological mechanisms linking MI to depression. We observed a discrepancy in the risk for depression among those with clinically recognized MI and those with unrecognized MI if compared to those without MI. We found that men with recognized MI were at a higher risk to develop depression than men without MI. In

contrast, men with unrecognized MI had similar risk to develop depression as those without MI. This suggests that in men the awareness of experiencing a life-threatening condition increased the risk of developing depression and not only the pathology related to MI. Interestingly, we did observe a similar impact of recognized and unrecognized MI on the mortality risk in this population. This confirms the serious pathophysiological impact both conditions had on the general health and life expectancy. Yet, clinically recognized and unrecognized MI seem to affect the mental health rather differently. We have to emphasize that due to low statistical power and possible measurement error we could not draw any conclusions from our data with regard to women. Thus, the evidence we provide should be interpreted carefully and is restricted to men.

In conclusion, **Chapter 5** presents studies in which we see whether physical diseases are risk factors for the occurrence of depression and vice versa, how depression is a risk factor for the occurrence of physical diseases. More specifically, we provide evidence how MI contributes to the development of depression and how multiple-episode depression contributes to poor long-term cognitive performance. The complexity of the interrelations between depression and its comorbidities requires well-designed epidemiological studies including repeated exposure and outcome measures to unravel bi-directional associations and the effect of chronic exposure. We have made an attempt to advance the epidemiological knowledge using epidemiological tools appropriate for the available data. Yet our results can only hint at the possible pathological mechanisms and factors that can help screen specific groups, or target to better manage or treat depressed individuals.

Methodological considerations

Analysing repeated measures in recurrent phenotypes

Traditionally, epidemiologists and statisticians advise that longitudinal population-based studies, even if observational, should be cautiously interpreted when referring to causality. Yet, recently advanced methodological tools have opened the possibility of a deeper and more complex understanding of disease initiation, course and progress. These methods were developed jointly by statisticians, social scientists, and epidemiologists to allow the possibility to disentangle the causal relationship between exposures and disease.[5] As a result, we now have new and more diverse types of approaches to consider when analysing data beyond the traditional epidemiologic analyses.

The Rotterdam Study has a unique way of assessing depression, previously described in detail (see **Chapter 1**). We assessed depression with repeated measures at well-specified examination

rounds (on average every 4 to 5 years). In addition, to learn about the course of the disease and capture events that might have happened between examination rounds, we performed continuous monitoring for depression during 12 years of follow-up. This assessment strategy allows individuals to be followed for the course of the disease, the index occurrence of the observed events; but it also allows individuals to be assessed for the severity of the depressive events. The severity of the experienced depression (from depressive symptoms to MDD) was defined by well-specified and sensitive questionnaires, interviews, and information collected from general practitioners (GP) medical records. This approach is an attempt to assess mental disorders such as depression; disorder with a remitting course across time, and characterized duration and severity, in more detail. The Rotterdam Study depression data allows to conceptualize depression as both episodic (counting episodes of the disorder that occur over time), and persisting (to define whether the disorder persists across time or not), and finally to distinguish between different degrees of severity within a population-based sample. Most longitudinal designs on depression have limited validity[37] due to the fact they conceptualize depression either by severity or by persistence only, but never combine or compare the two aspects. Our complex data avoids such limitation, yet collecting such a rich data is challenging and may introduce a number of limitations and obstacles that should be discussed.

In **Chapters 2.1** and **5.2** two studies with almost identical study design are presented. We test the longitudinal association between an exposure of interest measured at baseline and depression as an outcome. In order to test the change of the depressive symptoms over time I used repeated measures of the outcome at a regular time points (two re-assessments). Such repeated measures of the outcome are correlated with each other and such correlation need to be treated cautiously. I used Generalized Estimating Equations (GEE) model.[55] This model is more advanced than a traditional linear regression model and estimates a marginal population averaged effect that takes into account the within-subject correlation by using a covariance matrix.[55] Our depressive symptoms data are balanced data and therefore such a method is appropriate. However, GEE is a marginal model and no conclusions about individual patterns can be drawn, especially patterns that occur across time such as trajectories.[55] Indeed, depressive symptoms are not a normally distributed outcome variable and this may bias the within-subject correlation. To estimate the individual change of depressive symptoms across time we could have better used conditional model such as the generalized linear mixed model (GLMM).[56] Conditional models are a more fundamental approach and estimate a random effect while holding other covariates constant. Also, from the estimated random effect we can easily draw conclusions about the population effect, which is estimated by the GEE models; but from the GEE model we cannot estimate the individual effect.[56] Moreover, linear mixed models treat loss to follow-up more flexible than GEE models. GLMM are evidently better approach to analyse our data. However, the marginal GEE model is more straight forward and addresses the population-based effect we were particularly interested in.[56]

Longitudinal analyses are considered biased if they do not take into account the duration of follow-up. Thus, in addition to the GEE analyses we estimated the survival risk using a Cox regression model.[57] When compared to the GEE model, the Cox survival model takes into account time but does not fully address the course of the depressive disorder. Moreover, the survival model was used to estimate the effect of a one-time measure of the exposure at the baseline of the study. Such an approach is an important limitation given the type of exposure that changes over time such as vitamin D (**Chapter 2.1**).[58,59] Therefore, I argue that longitudinal associations between exposure and outcome, both changeable across time, and with complex interrelations; should be tested by advanced models that would address this issue. Joint models where the exposure using GLMM and the outcome using Cox model with time varying exposure are possible solutions.[60]

In **Chapter 2.2** we estimated the 5 year risk of increased depressive symptoms in persons with inflammatory measures such as CRP. Traditional linear and logistic models were used to analyse the longitudinal association between an exposure assessed at baseline and an outcome assessed at a fixed follow-up time. In order to test whether persistence of depressive symptoms can explain this longitudinal association we combined the baseline measure of the outcome depressive symptoms and defined persistence, as clinically relevant depressive symptoms at baseline and at follow-up. We tested both whether inflammatory factors were related to worsening depressive symptoms and persistent depressive symptoms. However, such model is incomplete and does not answer the question whether more depressive symptoms also increase the risk of increased inflammatory markers at follow-up. Prior depression and reverse causality may bias the reported results. Ideally, if data of two inflammatory markers assessments and depressive symptoms were available, we could have used cross-lagged structural equation modelling and address the directionality of the association and reduced the potential bias of reverse causality.[61]

Chapter 5.1 of this thesis uses a contemporary statistical method to estimate the contribution of depression to the long-term course of cognitive functioning using the latent growth curve modeling.[62] First, in order to conceptualize depression as episodic disorder, the exposure of interest, we used repeated measures of depression. Second, in order to estimate individual curves of cognitive decline over 12 years' time (the outcome of interest) we used repeated measures of four tests of cognitive functioning. Finally, to estimate the risk of a faster cognitive decline among persons with multiple episodes, a single episode compared to no episodes of depression, we regressed the exposure on the estimated latent growth curve of cognitive decline. This design appears ideal, however it was limited by a small overlap of the depression and the cognitive functioning assessments (see **Figure 1, Chapter 5.1**). Thus, our results only partially explain how the course of depression could predict the course of cognitive functioning. Most of the studies addressing similar hypothesis assessed depression in parallel to cognitive functioning assessment

(i.e. the periods of depressive symptoms and cognitive decline assessment overlap). The limitation of such a design increases the likelihood of reverse causality. Mixed modelling with depression assessment prior to the cognitive curve would be a possible solution to avoid this limitation of our design. One study assessed a trajectory of persistent depressive symptoms prior to the cognitive functioning assessments to avoid this limitation.[63] However, such a design also has limitations as it might introduce bias by immortal time.[64]

The above-mentioned more advanced statistical methods would ideally fill the gaps and limitations of the traditional analysis we performed in this thesis. However, such analyses require data that were not yet available within the Rotterdam Study. Longitudinal cohort studies are time consuming and expensive designs. Financial and other methodological obstacles (such as data cleaning) sometimes result in incomplete measurements in a particular examination round. My decisions in terms of statistical designs were mostly based on the available data. Finally, we should also consider that advanced methodologies have difficult computations (require certain fits, are dependent on number of assumptions), require big sample sizes, and are mostly difficult to interpret. They are also a formidable challenge for PhD students; however, mastering these techniques certainly improves academic career prospects.

Time varying confounding

Population-based studies use observational data to study associations while trying to address causal questions. One of the fundamental problems of these studies is confounding bias. Confounding bias is a bias introduced by all factors considered to be common causes of both the exposure and the outcome of the study.[65] Defining which are the factors that may bias the association between the exposure and outcome, to be addressed as potential confounders, requires prior knowledge on the causal assumption. In this thesis we control for confounding by using baseline adjustment. Based on prior knowledge we typically included all factors that may be related to both exposure and outcome in the tested model. This approach constitutes the most basic statistical method for treating confounding and it is called analysis of covariance.[66]

The Rotterdam Study has extensive measures of demographic, life-style and health factors across time. Therefore, this prospective cohort design allows us to adjust for numerous potential confounders and minimize the risk of residual confounding by unknown or unmeasured confounding. Yet, what remains a serious concern in using this methodology is omitting the effect that time-varying factors have on the observed association. For example, in the tested vitamin D depression association in **Chapter 2.1** vitamin D is an exposure that is time varying and change of the exposure over time may have been caused by confounder factors. In the case that the change due to cofounders may have occurred independent of baseline vitamin D and prior to

the outcome the association may be biased. Due to limited data (repeated measures of vitamin D were not available in the Rotterdam Study) we could not account for time-varying exposure and confounding.

Ideally, we could have avoided time-varying confounding by using the methodology of inverse probability weighting and the Marginal Mixed models.[66] As previously suggested, using Joint models are another strategy to treat both repeated measures of time-varying exposure and outcome under one model.[60] These methodologies are more advanced and allow to test the interrelation and interaction between exposure and time. Such methods are also appropriate for recurrent outcomes and may answer potential causal relations better than traditional designs. However, as already discussed it is not always possible to implement such methods. Moreover, using such methods may also produce biased results. For example, similarly to traditional models they do not account for unmeasured confounding.

Broad phenotype of depression

Depression is a heterogeneous disease, and no established mechanism, risk factor, or features of the disease can explain all the aspects of this disorder. In my opinion, the biggest challenge of the work presented in this thesis and in psychiatric research in general is the definition and the characterization of the depressive phenotype. The question whether depression should be studied as a broad continuous phenotype (depressive symptoms) or as the narrow DSM-classification based phenotype of depression is what I have been asking myself most often. Indeed, the scientific world knows an endless discussion over the most appropriate depression phenotype to be used in research. This dilemma has not been solved yet and the decision to use one or the other depression phenotype brings advantages and disadvantages that should be discussed.

In epidemiologic research, both depressive symptoms phenotype but also DSM-categorizations phenotypes are widely used. The Rotterdam Study, as already discussed, performs assessment of both phenotypes; screening for depressive symptoms and establishment of DSM-criterion based depression diagnosis among those with clinically relevant depressive symptoms. Using the broad phenotype of depressive symptoms has helped identify numerous associations. Depressive symptoms were found to be associated with number of demographic factors[67], comorbidities[68], life style factors[69], biological markers[70], increased risk of disabilities[71], chronic diseases[72], and mortality[73]. Moreover, the recently developed neuroimaging technology has enabled us to relate depressive symptoms to various structural and functional brain measures.[74-76] Obviously, epidemiologic research profited from using such a broad phenotype. Depressive symptoms are meaningful on the population-based level and for public health purposes in general. However, results from studies using depressive symptoms are not easily translatable to clinical practice.

Therefore, many population-based studies have attempted to study DSM-categorized diagnosis as their depression phenotype of interest. This work showed success as well and replicated almost all associations reported with depressive symptoms.[77] For example, structural brain changes associated with MDD were also found to be associated with even mild depressive symptomes.[75,77] In two of the studies we also observed consistent results regardless of the used depression phenotype (**Chapter 2.1** and **5.2**). However, we have to emphasize that assessment of DSM-categorized depression phenotypes within population-based studies is challenging. It requires more complicated semi-structured psychiatric interviews or/and retrieval of GP records data. Compared to this approach, screening questionnaires are easily assessable, cheap and more simple. Moreover, the prevalence of cases with depressive symptoms is much higher (around 10 - 29%) than the low prevalence of MDD cases (around 6%) in population-based studies.[77,78] Therefore, the likelihood of detecting associations with low effect decreases when using DSM-categorizes depression phenotypes and requires larger sample sizes although this may be partly counterbalanced by smaller effects if clinically relevant depressive symptoms are used as a categorical outcome.

In genetic-epidemiology research, however, we are not so successful in identifying associations and we are certainly far from replicating results using different phenotypes of depression as in non-genetic epidemiologic research. Early Genome wide association study (GWAS) failed to identify single nucleotide polymorphisms in association with MDD in almost 9 000 participants.[79] It was suggested, that the potential for discoveries in genetics is likely to increase if we use more broad and heterogeneous samples of depressive phenotypes without exclusions based on the DSM-diagnostic criteria and without constraining their measurement to consensus based categorize.[37] Therefore, following GWAS investigated the depressive symptoms phenotype in 30 000 individuals.[80] However, this study also failed to identify genome-wide significant genetic associations. Genetic studies are based on the assumption that studying heterogenic, polygenic, complex, and common phenotypes such as depression requires studying bigger sample sizes or lowering the heterogeneity in order to find significant genetic associations.[33,77] Indeed, restricting the definition of depression to a more severe type of recurrent disorder helped in identifying two genetic loci.[33] Also, combining the phenotypes MDD and depressive symptoms to increase power at the expense of heterogeneity to detect loci associated with depressive symptoms.[81] However, such an approach may not unravel the effects of SNPs associated only with one phenotype, and therefore produce biased results towards the genetics of a broad phenotype.[33] Finally, the most successful published GWAS so far, by 23 and Me, identified 15 loci by using mild MDD phenotyping from a self-report of diagnosis.[82] Although this phenotype strategy brought success in identifying depression associated SNPs, the possibility of misclassification bias is likely.[33] Therefore, using the most appropriate phenotype of depression in genetic research is crucial and again the same question emerges: "What is the most appropriate phenotype?"

It is sometimes assumed that a smaller number of genes will be involved in the causality of more severe depressive illnesses than in depressive symptoms. Also, some genotype variations should be related to a specific subtype of depressive illness only.[37] Thus, we may expect that the identified genetic loci of studying depressive symptoms and MDD will differ but a considerable overlap. Indeed, such overlap is likely and a recent GWAS has proven that by replicating genetic association between depressive symptoms and the *FHIT* gene in an independent sample of MDD cases.[83] Yet, the rare variants located on the *LIPG* and *NKPD1* gene (**Chapter 3**) we identify as associated with depressive symptoms may not be relevant to MDD or other more specific depression phenotypes. But that does not mean our results are biased or wrong.

In conclusion, epidemiology benefited from using broad depression phenotypes but these findings typically replicate when DSM-criterion based phenotypes were used. Whether we choose to use one or the other phenotype should mainly depend on the research question. Genetic research on the other hand, still attempts to learn which phenotype choice will bring success and, even more important, accurate results. So far it seems that combining phenotypes and using more flexible measures allows identifying more genetic associations. However, whether such findings have important theoretical meaning for an in-depth understanding of depression is questionable.

The depression phenotype issue was probably a good motivation for The National Institute of Mental Health (NIMH) to develop new criteria for research; the research domain criteria (RDoC).[84] The RDoC is a matrix based approach structured by specified functional constructs characterized by genes, molecules, cells, circuits, physiology, self-report, and paradigms. The constructs are finally grouped into five higher-level domains of functioning (negative valence systems, positive valence systems, cognitive systems, systems for social processes, and arousal/regulatory systems). The main aim of the RDoC is to provide an attempt for a deeper understanding of the biological and psychosocial basis of psychiatric disorders but also to help to improve current classification systems.[77] Such an approach opens new possibilities for psychiatric research in general and may solve the current phenotyping issues.

Future perspectives and implications

Most of the suggestions for future research perspectives for and clinical implications of the studies included in this thesis were already discussed in the earlier sections of this discussion. However, here I would like to summarize and emphasize some of the most important points and draw a brief conclusion of this thesis.

The projects included in **Chapter 2**, **Chapter 3**, and **Chapter 4** all provide evidence on biological factors related to depression. Vitamin D, inflammatory markers, blood DNA epigenetic markers, and the rare variants of the *LIPG* and *NKPD1* gene are biological factors that may be used as biomarkers for depression. The reported studies suggest these biological factors as potential diagnostic markers of depression that might help differentiation between depressed and non-depressed persons. However, future epidemiologic studies should prove their clinical validity and translate results to clinical practice. Moreover, some of these factors (for example the studied inflammatory markers IL-6 and CRP) could be used to identify specific subgroups of depressed individuals, for example individuals with persistent depression. They may therefore provide better screening of depressed individuals for specific subtypes and help to target more successful treatment for depression.

The studies presented in **Chapter 2.1** and **5.2** have limited clinical implications. Both vitamin D serum levels and unrecognized MI are not related to the occurrence of depression. They may help targeting specific groups of individuals suffering from chronic diseases and conditions; for example, detecting depressed patients with lower vitamin D serum levels for osteoporosis preventing strategies. However, I doubt their clinical meaning as depression specific biomarkers.

The presented evidence is population-based; therefore, it is more relevant to populations than clinical samples. Policy makers can use these reports as direction for making future public health policies and strategies to improve people's general health and quality of life. Yet, these studies also provide evidence for biological associations (for example between certain genes and depression) that have theoretical importance and may guide future basic research, which will hopefully confirm their clinical relevance and practical use.

In terms of future research perspectives, the literature lacks prospective studies with repeated measures of both depression and its determinants. Such designs will allow studying interrelations between depression and exposures of interest as well as address direction of the relations. Also, repeated assessment data will allow using more advance statistical methodology and thus yield more valid results.

Moreover, future psychiatric research should certainly address poly-causation by considering multiple genetic and environmental measures. Depression is a common disease, often with an adolescent first onset, and with a high prevalence through the entire life span. Elderly cohorts studying aging such as the Rotterdam Study are limited in studying first onset, childhood exposures, and life-course of depression. Exploring such a retrospective hypothesis within these setting is inaccurate and would produce biased results. Therefore, future epidemiological research requires prospective designs that start in childhood, at birth or even pregnancy and pre-conception period.[37]

Genetics is a field that has evolved and developed rapidly in the last years and scientists had high hopes that genetics will disentangle many mysteries. However, psychiatric genetics faced several challenges when studying depression: the definition of depressive phenotypes appeared to be the greatest obstacle for the scientists. In the future we will have to agree on concepts and definitions of the depressive phenotype and be open for implementing new concepts in psychiatric research such as RDoC.[84] I believe overcoming the issue of defining depression will help genetic research but also the scientific knowledge of depression in general.

Finally, depression is a multifactorial disorder and in order to deeply understand the aetiology we need to address gene-environment interactions. Analyses of poly-causation including gene-environment and environment-environment interactions can help in learning more about causal types and mechanisms that are also likely to generate new methods of prevention and therapy.[37] Yet, current studies of gene-environment interactions are flawed due to the lack of power and biological understanding.

In conclusion, this thesis uses epidemiologic and genetic tools to explore the associations between potential aetiological factors and depression. We have managed to partly fill a few of the many gaps in the knowledge of the pathophysiological mechanisms underlying depression and the biological markers for depression. Further, I have discussed the clinical importance of the findings in this thesis, as well as the methodological challenges I have faced during this work. Most of the evidence presented in this thesis will not directly influence clinical practice and the life of those suffering depression. However, I believe that this body of research will definitely be a good guidance for future research and will help the management of depression.

REFERENCES

1. Venkatasubramanian G, Keshavan MS. Biomarkers in Psychiatry - A Critique. *Ann Neurosci.* 2016;23(1):3-5.
2. Gadad BS, Jha MK, Czysz A, et al. Peripheral biomarkers of major depression and antidepressant treatment response: Current knowledge and future outlooks. *J Affect Disord.* 2017.
3. Scirica BM. Use of Biomarkers in Predicting the Onset, Monitoring the Progression, and Risk Stratification for Patients with Type 2 Diabetes Mellitus. *Clin Chem.* 2017;63(1):186-195.
4. WHO International Programme on Chemical Safety Biomarkers in Risk Assessment: Validity and Validation. 2001.
5. McGrath JJ. Vitamin D and mental health - the scrutiny of science delivers a sober message. *Acta Psychiatr Scand.* 2017;135(3):183-184.
6. Lapid MI, Cha SS, Takahashi PY. Vitamin D and depression in geriatric primary care patients. *Clin Interv Aging.* 2013;8:509-514.
7. Ju SY, Lee YJ, Jeong SN. Serum 25-hydroxyvitamin D levels and the risk of depression: a systematic review and meta-analysis. *J Nutr Health Aging.* 2013;17(5):447-455.
8. Anglin RE, Samaan Z, Walter SD, McDonald SD. Vitamin D deficiency and depression in adults: systematic review and meta-analysis. *Br J Psychiatry.* 2013;202:100-107.
9. Scarr E, Millan MJ, Bahn S, et al. Biomarkers for Psychiatry: The Journey from Fantasy to Fact, a Report of the 2013 CINP Think Tank. *Int J Neuropsychopharmacol.* 2015;18(10):pyv042.
10. Hutcheon JA, Chiolero A, Hanley JA. Random measurement error and regression dilution bias. *BMJ.* 2010;340:c2289.
11. Clarke R, Shipley M, Lewington S, et al. Underestimation of risk associations due to regression dilution in long-term follow-up of prospective studies. *Am J Epidemiol.* 1999;150(4):341-353.
12. Uher R, Payne JL, Pavlova B, Perlis RH. Major depressive disorder in DSM-5: implications for clinical practice and research of changes from DSM-IV. *Depress Anxiety.* 2014;31(6):459-471.
13. Holick MF. Vitamin D deficiency. *N Engl J Med.* 2007;357(3):266-281.
14. Lopresti AL, Maker GL, Hood SD, Drummond PD. A review of peripheral biomarkers in major depression: the potential of inflammatory and oxidative stress biomarkers. *Prog Neuropsychopharmacol Biol Psychiatry.* 2014;48:102-111.
15. Liu Y, Al-Sayegh H, Jabrah R, Wang W, Yan F, Zhang J. Association between C-reactive protein and depression: modulated by gender and mediated by body weight. *Psychiatry Res.* 2014;219(1):103-108.
16. Matthews KA, Schott LL, Bromberger JT, Cyranowski JM, Everson-Rose SA, Sowers M. Are there bi-directional associations between depressive symptoms and C-reactive protein in mid-life women? *Brain Behav Immun.* 2010;24(1):96-101.
17. van den Biggelaar AH, Gussekloo J, de Craen AJ, et al. Inflammation and interleukin-1 signaling network contribute to depressive symptoms but not cognitive decline in old age. *Exp Gerontol.* 2007;42(7):693-701.
18. Dantzer R, O'Connor JC, Freund GG, Johnson RW, Kelley KW. From inflammation to sickness and depression: when the immune system subjugates the brain. *Nat Rev Neurosci.* 2008;9(1):46-56.
19. Joyce PR, Hawes CR, Mulder RT, Sellman JD, Wilson DA, Boswell DR. Elevated levels of acute phase plasma proteins in major depression. *Biol Psychiatry.* 1992;32(11):1035-1041.
20. Kivimaki M, Shipley MJ, Batty GD, et al. Long-term inflammation increases risk of common mental disorder: a cohort study. *Mol Psychiatry.* 2014;19(2):149-150.
21. Hoertel N, Blanco C, Oquendo MA, et al. A comprehensive model of predictors of persistence and recurrence in adults with major depression: Results from a national 3-year prospective study. *J Psychiatr Res.* 2017;95:19-27.

22. Saveanu RV, Nemeroff CB. Etiology of depression: genetic and environmental factors. *Psychiatr Clin North Am.* 2012;35(1):51-71.
23. Hoffmann A, Sportelli V, Ziller M, Spengler D. Epigenomics of Major Depressive Disorders and Schizophrenia: Early Life Decides. *Int J Mol Sci.* 2017;18(8).
24. Nemoda Z, Szyf M. Epigenetic Alterations and Prenatal Maternal Depression. *Birth Defects Res.* 2017;109(12):888-897.
25. Pishva E, Rutten BPF, van den Hove D. DNA Methylation in Major Depressive Disorder. *Adv Exp Med Biol.* 2017;978:185-196.
26. Scriver CR, Eisensmith RC, Woo SL, Kaufman S. The hyperphenylalaninemias of man and mouse. *Annu Rev Genet.* 1994;28:141-165.
27. Demirkan A, Isaacs A, Ugocsai P, et al. Plasma phosphatidylcholine and sphingomyelin concentrations are associated with depression and anxiety symptoms in a Dutch family-based lipidomics study. *J Psychiatr Res.* 2013;47(3):357-362.
28. Burkhardt C, Zacharias M. Modelling ion binding to AA platform motifs in RNA: a continuum solvent study including conformational adaptation. *Nucleic Acids Res.* 2001;29(19):3910-3918.
29. Fiore R, Puschel AW. The function of semaphorins during nervous system development. *Front Biosci.* 2003;8:s484-499.
30. Giordano S, Corso S, Conrotto P, et al. The semaphorin 4D receptor controls invasive growth by coupling with Met. *Nat Cell Biol.* 2002;4(9):720-724.
31. Heyn H, Moran S, Hernando-Herraez I, et al. DNA methylation contributes to natural human variation. *Genome Res.* 2013;23(9):1363-1372.
32. Egger G, Liang G, Aparicio A, Jones PA. Epigenetics in human disease and prospects for epigenetic therapy. *Nature.* 2004;429(6990):457-463.
33. Mullins N, Lewis CM. Genetics of Depression: Progress at Last. *Curr Psychiatry Rep.* 2017;19(8):43.
34. Tennessen JA, Bigham AW, O'Connor TD, et al. Evolution and functional impact of rare coding variation from deep sequencing of human exomes. *Science.* 2012;337(6090):64-69.
35. Fehrmann RS, Karjalainen JM, Krajewska M, et al. Gene expression analysis identifies global gene dosage sensitivity in cancer. *Nat Genet.* 2015;47(2):115-125.
36. van der Sluis S, Posthuma D, Nivard MG, Verhage M, Dolan CV. Power in GWAS: lifting the curse of the clinical cut-off. *Mol Psychiatry.* 2013;18(1):2-3.
37. Uher R, Zwicker A. Etiology in psychiatry: embracing the reality of poly-gene-environmental causation of mental illness. *World Psychiatry.* 2017;16(2):121-129.
38. Hatzikotoulas K, Gilly A, Zeggini E. Using population isolates in genetic association studies. *Brief Funct Genomics.* 2014;13(5):371-377.
39. Gulbins E, Palmada M, Reichel M, et al. Acid sphingomyelinase-ceramide system mediates effects of antidepressant drugs. *Nat Med.* 2013;19(7):934-938.
40. Cross-Disorder Group of the Psychiatric Genomics C. Identification of risk loci with shared effects on five major psychiatric disorders: a genome-wide analysis. *Lancet.* 2013;381(9875):1371-1379.
41. Witt SH, Streit F, Jungkunz M, et al. Genome-wide association study of borderline personality disorder reveals genetic overlap with bipolar disorder, major depression and schizophrenia. *Transl Psychiatry.* 2017;7(6):e1155.
42. Saha S, Chant D, Welham J, McGrath J. A systematic review of the prevalence of schizophrenia. *PLoS Med.* 2005;2(5):e141.
43. Merikangas KR, Akiskal HS, Angst J, et al. Lifetime and 12-month prevalence of bipolar spectrum disorder in the National Comorbidity Survey replication. *Arch Gen Psychiatry.* 2007;64(5):543-552.
44. Lescai F, Als TD, Li Q, et al. Whole-exome sequencing of individuals from an isolated population implicates rare risk variants in bipolar disorder. *Transl Psychiatry.* 2017;7(2):e1034.

45. Sirkis DW, Bonham LW, Aparicio RE, et al. Rare TREM2 variants associated with Alzheimer's disease display reduced cell surface expression. *Acta Neuropathol Commun.* 2016;4(1):98.
46. Marouli E, Graff M, Medina-Gomez C, et al. Rare and low-frequency coding variants alter human adult height. *Nature.* 2017;542(7640):186-190.
47. Megari K. Quality of Life in Chronic Disease Patients. *Health Psychol Res.* 2013;1(3):e27.
48. Hare DL, Toukhsati SR, Johansson P, Jaarsma T. Depression and cardiovascular disease: a clinical review. *Eur Heart J.* 2014;35(21):1365-1372.
49. Papazacharias A, Nardini M. The relationship between depression and cognitive deficits. *Psychiatr Danub.* 2012;24 Suppl 1:S179-182.
50. Gallo JJ, Hwang S, Joo JH, et al. Multimorbidity, Depression, and Mortality in Primary Care: Randomized Clinical Trial of an Evidence-Based Depression Care Management Program on Mortality Risk. *J Gen Intern Med.* 2016;31(4):380-386.
51. Sheline YI, Wang PW, Gado MH, Csernansky JG, Vannier MW. Hippocampal atrophy in recurrent major depression. *Proc Natl Acad Sci U S A.* 1996;93(9):3908-3913.
52. McIntyre RS, Cha DS, Soczynska JK, et al. Cognitive deficits and functional outcomes in major depressive disorder: determinants, substrates, and treatment interventions. *Depress Anxiety.* 2013;30(6):515-527.
53. Diniz BS, Butters MA, Albert SM, Dew MA, Reynolds CF, 3rd. Late-life depression and risk of vascular dementia and Alzheimer's disease: systematic review and meta-analysis of community-based cohort studies. *Br J Psychiatry.* 2013;202(5):329-335.
54. Panza F, Frisardi V, Capurso C, et al. Late-life depression, mild cognitive impairment, and dementia: possible continuum? *Am J Geriatr Psychiatry.* 2010;18(2):98-116.
55. Ziegler A, Vens M. Generalized estimating equations. *Methods of information in medicine.* 2010;49(5):421-425.
56. Muff S, Held L, Keller LF. Marginal or conditional regression models for correlated non-normal data? *Methods in Ecology and Evolution.* 2016;7(12):1514-1524.
57. Cox DR, Oakes D. *Analysis of survival data.* Vol 21: CRC Press; 1984.
58. Landel V, Wion D. Vitamin D and mental health: optimizing in the midst of the complexity. *Acta Psychiatr Scand.* 2017;136(2):228-229.
59. Griffen TC. Comment on Vitamin D serum levels are cross-sectionally but not prospectively associated with late-life depression. *Acta Psychiatr Scand.* 2017.
60. Yang L. *Joint models for longitudinal and survival data.* Indiana University; 2013.
61. Cacioppo JT, Hawkley LC, Thisted RA. Perceived social isolation makes me sad: 5-year cross-lagged analyses of loneliness and depressive symptomatology in the Chicago Health, Aging, and Social Relations Study. *Psychology and aging.* 2010;25(2):453.
62. Preacher KJ. *Latent growth curve modeling.* Sage; 2008.
63. Zeki Al Hazzouri A, Vittinghoff E, Byers A, et al. Long-term Cumulative Depressive Symptom Burden and Risk of Cognitive Decline and Dementia Among Very Old Women. *The Journals of Gerontology: Series A.* 2014;69(5):595-601.
64. Lévesque LE, Hanley JA, Kezouh A, Suissa S. Problem of immortal time bias in cohort studies: example using statins for preventing progression of diabetes. *BMJ.* 2010;340.
65. Hernán MA. Confounding. *Encyclopedia of Quantitative Risk Analysis and Assessment*: John Wiley & Sons, Ltd; 2008.
66. Barber JS, Murphy SA, Verbitsky N. Adjusting for Time-Varying Confounding in Survival Analysis. *Sociological Methodology.* 2004;34(1):163-192.
67. Chen J, Yu J. Sex Differences in Genetic and Environmental Influences on Adolescent Depressive Symptoms: A Meta-Analytic Review. *Depress Res Treat.* 2015;2015:476238.

68. Danna SM, Graham E, Burns RJ, Deschenes SS, Schmitz N. Association between Depressive Symptoms and Cognitive Function in Persons with Diabetes Mellitus: A Systematic Review. *PLoS One.* 2016;11(8):e0160809.
69. Bursnall P. The relationship between physical activity and depressive symptoms in adolescents: a systematic review. *Worldviews Evid Based Nurs.* 2014;11(6):376-382.
70. Jovanova O, Aarts N, Noordam R, Carola-Zillikens M, Hofman A, Tiemeier H. Vitamin D serum levels are cross-sectionally but not prospectively associated with late-life depression. *Acta Psychiatr Scand.* 2017;135(3):185-194.
71. Bruce ML, Seeman TE, Merrill SS, Blazer DG. The impact of depressive symptomatology on physical disability: MacArthur Studies of Successful Aging. *American Journal of Public Health.* 1994;84(11):1796-1799.
72. Mourao RJ, Mansur G, Malloy-Diniz LF, Castro Costa E, Diniz BS. Depressive symptoms increase the risk of progression to dementia in subjects with mild cognitive impairment: systematic review and meta-analysis. *Int J Geriatr Psychiatry.* 2016;31(8):905-911.
73. Kozela M, Bobak M, Besala A, et al. The association of depressive symptoms with cardiovascular and all-cause mortality in Central and Eastern Europe: Prospective results of the HAPIEE study. *Eur J Prev Cardiol.* 2016;23(17):1839-1847.
74. Szymkowicz SM, McLaren ME, O'Shea A, Woods AJ, Anton SD, Dotson VM. Depressive symptoms modify age effects on hippocampal subfields in older adults. *Geriatr Gerontol Int.* 2016.
75. Allan CL, Sexton CE, Filippini N, et al. Sub-threshold depressive symptoms and brain structure: A magnetic resonance imaging study within the Whitehall II cohort. *J Affect Disord.* 2016;204:219-225.
76. Zetzsche T, Frodl T, Preuss UW, et al. Amygdala volume and depressive symptoms in patients with borderline personality disorder. *Biol Psychiatry.* 2006;60(3):302-310.
77. Otte C, Gold SM, Penninx BW, et al. Major depressive disorder. *Nat Rev Dis Primers.* 2016;2:16065.
78. Glaesmer H, Riedel-Heller S, Braehler E, Spangenberg L, Luppa M. Age- and gender-specific prevalence and risk factors for depressive symptoms in the elderly: a population-based study. *Int Psychogeriatr.* 2011;23(8):1294-1300.
79. Major Depressive Disorder Working Group of the Psychiatric GC, Ripke S, Wray NR, et al. A mega-analysis of genome-wide association studies for major depressive disorder. *Mol Psychiatry.* 2013;18(4):497-511.
80. Hek K, Demirkan A, Lahti J, et al. A genome-wide association study of depressive symptoms. *Biol Psychiatry.* 2013;73(7):667-678.
81. Okbay A, Baselmans BM, De Neve JE, et al. Genetic variants associated with subjective well-being, depressive symptoms, and neuroticism identified through genome-wide analyses. *Nat Genet.* 2016;48(6):624-633.
82. Hyde CL, Nagle MW, Tian C, et al. Identification of 15 genetic loci associated with risk of major depression in individuals of European descent. *Nat Genet.* 2016;48(9):1031-1036.
83. Direk N, Williams S, Smith JA, et al. An Analysis of Two Genome-wide Association Meta-analyses Identifies a New Locus for Broad Depression Phenotype. *Biol Psychiatry.* 2017;82(5):322-329.
84. Insel TR. The NIMH Research Domain Criteria (RDoC) Project: precision medicine for psychiatry. *Am J Psychiatry.* 2014;171(4):395-397.

CHAPTER 7

Summary/Samenvatting

R1
R2
R3
R4
R5
R6
R7
R8
R9
R10
R11
R12
R13
R14
R15
R16
R17
R18
R19
R20
R21
R22
R23
R24
R25
R26
R27
R28
R29
R30
R31
R32
R33
R34
R35
R36
R37
R38
R39

SUMMARY

Depression is a complex multifactorial disease, it is the result of an interaction and accumulation of various different psychosocial, biological and environmental risk factors. Evidence on the aetiology of depression is mixed and important elements are still missing. In this thesis, we attempted to fill some of these knowledge gaps by studying the association of several biomarkers, genetic, epigenetic, and physical health factors with depression. Here we summarize the main findings that have been described in detail in chapters 2 to 5.

Chapter 2 presents two studies exploring the relation between serum extracted biomarkers and depression. In **Chapter 2.1**, we studied the cross-sectional and longitudinal relation between vitamin D serum levels and depression. All analyses were controlled for life-style factors, including sunlight exposure, socio-demographic factors, and general health factors. Although vitamin D deficiency co-occurred with depression, we found no evidence for an effect of vitamin D serum on the change in depressive symptoms or the long-term (12 year) occurrence of depression. In **Chapter 2.2**, we explored serum levels of inflammatory biomarkers in relation to depressive symptoms with a specific focus on the effect of inflammation on the persistence of depressive symptoms. Markers of inflammation, such as higher levels of interleukin-6 and C reactive protein, predicted both increase and persistence of depressive symptoms over 5 years. In contrast, markers of inflammation were not associated with depressive symptoms cross-sectionally. These results suggest that depression does not occur as an acute response to inflammation but rather as a reaction to chronic inflammation.

Chapter 3 employed exome-sequencing and exome-chip genotyping methodology to explore rare genetic variants that potentially had large effects on the risk of depression. In **Chapter 3.1** we demonstrated that depressive symptoms were associated with missense *Asn396Ser* mutation (rs77960347) located in the endothelial lipase (*LIPG*) gene in a general population setting. **Chapter 3.2** identified rare non-synonymous variants located in the *NKPD1* gene associated with more depressive symptoms in a genetically isolated European population. The *LIPG* and the *NKPD1* genes are candidate genes for depression that require further genetic and functional investigation.

Chapter 4 presents the largest (*N* = 11,256) cross-ethnic epigenome-wide association study (EWAS) of depressive symptoms. Hypermethylation of cg04987734 (annotated to the *CDC42BPB* gene), cg12325605 (annotated to the *ARHGEF3* gene), and cg14023999 (having an intergenic position) were all associated with depressive symptoms. The identified genes are involved in mechanisms such as axon guidance, nerve growth factor, regulation of serotonin and dopamine levels in the brain, and inflammation, yielding new insights into the molecular mechanisms potentially underlying the pathophysiology of depression.

In **Chapter 5** we explored the relation between depression and comorbid physical conditions such as cognitive decline and myocardial infarction. In **Chapter 5.1** we tested whether cognitive function declines faster in older adults who had single or multiple depressive episodes than in older adults with no depression. We found that persons with multiple episodes of depression had a faster cognitive decline over a 12 years' period than persons with a single or no depressive episode. This study suggests that the recurrence of depression, opposed to a single episode, contributes to subsequent cognitive decline. **Chapter 5.2** aimed to distinguish the physiological and psychological impacts from myocardial infarction on depression. Therefore, we explored the risk of recognized (RMI) and unrecognized myocardial infarction (UMI) pose for depression. The results demonstrated that men with RMI had an increased risk to develop depression, while men with UMI were at similar risk to develop depression as persons without MI. However, RMI and UMI increased mortality risk similarly compared to those with no MI. We thus concluded that both RMI and UMI have evident physiological effects but that the psychological burden of experiencing MI contributes to the long-term risk of depression.

Finally, **Chapter 6** of this thesis discusses the main findings and elaborates on the main methodological limitations of the presented work. In addition, this chapter gives an overview of some possible clinical implications and directions for future research.

SAMENVATTING

Depressie is een complexe, multifactoriële ziekte die ontstaat door interactie tussen, diverse psychosociale, biologische en omgevingsfactoren. Hoewel er meer en meer bekend is over depressie, is de etiologie van depressie nog altijd onduidelijk. In dit proefschrift probeer ik enkele lacunes in kennis over de etiologie van depressie in te vullen door het bestuderen van de relatie van depressie met verschillende biomarkers, genetische, epigenetische en fysieke gezondheidsfactoren. De belangrijkste onderzoeksbevindingen van dit proefschrift zullen hier worden samengevat, een gedetailleerde beschrijving kan gevonden worden in de hoofdstukken 2 tot en met 5.

Hoofdstuk 2 omvat twee studies die de relatie tussen depressie en serum biomarkers beschrijven. In **hoofdstuk 2.1** bestudeerden we de cross-sectionele en longitudinale relatie tussen de concentratie van vitamin D in serum en depressie. We corrigeerden daarbij voor een uitgebreid aantal lifestyle factoren (waaronder blootstelling aan zonlicht), socio-economische en demografische factoren, en algemene gezondheidsfactoren. Alhoewel vitamine D deficiëntie vaak samen gaat met depressie, vonden we geen bewijs dat de vitamine D concentratie de verandering in depressieve symptomen beïnvloedt, noch dat vitamine D de ontwikkeling van depressieve episoden op de lange termijn (12 jaar) beïnvloedt. In **hoofdstuk 2.2** onderzochten we de relatie van verschillende biomarkers voor inflammatie met depressieve symptomen, we richtten ons daarbij specifiek op de persistentie van de depressieve symptomen. Biomarkers van inflammatie, zoals verhoogde interleukine-6 en C reactieve proteïne waarden, voorspellen zowel de toename als de persistentie van depressieve symptomen over een periode van 5 jaar. De toename van inflammatoire markers was echter niet gerelateerd aan depressieve symptomen in cross-sectionele analyses, wat suggereert dat depressie waarschijnlijk niet een acute reactie is op inflammatie, maar een reactie op chronische inflammatie.

In **hoofdstuk 3** kijken we met exome sequencing en exoom-chip genotypering naar zeldzame genetische variaties, waarvan bekend is dat ze samenhangen met een verhoogd risico op depressie. In **hoofdstuk 3.1** demonstreerden we dat depressieve symptomen gerelateerd zijn aan een missense *Asn396Ser* mutatie (rs77960347), gelegen in het endothelial lipase *(LIPG)* gen in de algemene bevolking. In **hoofdstuk 3.2** identificeren we zeldzame 'niet-synonieme' varianten, gelegen in het *NKPD1* gen. Het *NKPD1* gen was geassocieerd met verhoogde depressieve symptomen in een genetisch geïsoleerde Europese populatie. Onze resultaten geven aan dat zowel het *LIPG* als het *NKPD1* gen mogelijk belangrijke kandidaat genen zijn voor depressie. Verder genetisch en functioneel onderzoek zal uit moeten wijzen hoe deze genen depressie beïnvloeden.

In **hoofdstuk 4** presenteren we de grootste, multiculturele, epigenoom associatiestudie ('EWAS') van depressieve symptomen (*N* = 11,256). We identificeerden dat hypermethylatie van cg04987734, behorend bij het *CDC42BPB* gen, cg12325605, behorend bij het *ARHGEF3* gen, en cg14023999, in een intergenic positie, allen gerelateerd zijn aan depressieve symptomen. Deze genen zijn betrokken bij mechanismen zoals axongeleiding, groeifactoren voor zenuwcellen, regulatie van serotonine en dopamine in het brein, en ontsteking. Mogelijk duiden deze bevindingen op moleculaire mechanismen die de ontwikkeling van depressie onderliggen.

In **hoofdstuk 5** onderzochten we de complexe relatie tussen depressie en comorbide fysieke factoren zoals cognitie en hartfalen. In **hoofdstuk 5.1** bestudeerden we of de cognitie van ouderen met één of meerdere depressieve episode sneller achteruit gaat dan ouderen zonder een historie van depressieve episoden. Ouderen met meerdere depressieve episoden hadden inderdaad een snellere cognitieve achteruitgang over een periode van 12 jaar dan ouderen met een historie zonder depressieve episoden of een enkele depressie episode. Dit suggereert dat niet een enkele episode, maar de terugkeer van depressieve episoden, een rol speelt in de cognitieve achteruitgang in ouderen. **Hoofdstuk 5.2** probeert te achterhalen in hoeverre de fysiologische dan wel de psychologische gevolgen van een hartinfarct van belang zijn voor het ontwikkelen van depressie. Hiervoor onderzochten wij depressie in relatie tot gediagnosticeerde hartinfarcten, bekend bij de deelnemer, en zogenaamde 'stille' hartinfarcten, onbekend bij de deelnemer. Mannen met een gediagnosticeerd hartinfarct hadden een verhoogd risico op het ontwikkelen van depressie in vergelijking met mannen zonder hartinfarct. In contrast, mannen met 'stille' hartinfarcten hadden een vergelijkbaar risico op depressie als mannen zonder hartinfarct. Ter vergelijking, het risico op overlijden is verhoogd voor zowel mannen met gediagnosticeerde als 'stille' hartinfarcten ten opzichte van mannen zonder hartinfarct. Zowel gediagnosticeerde als 'stille' hartinfarcten hebben dus nadelige gevolgen voor de fysieke gezondheid, maar alleen de psychologische ervaring van een gediagnosticeerd hartinfarct draagt bij aan het ontwikkelen van een depressie.

Als laatste worden in **hoofdstuk 6** de belangrijkste bevindingen en conclusie van dit proefschrift besproken. Daarnaast wordt verder ingegaan op de methodologische beperkingen, de betekenis van dit proefschrift voor de klinische praktijk en suggesties voor toekomstig onderzoek op het gebied van de etiologie van depressie.

CHAPTER 8

Addendum

CHAPTER 8.1

PhD Portfolio

Summary of PhD training and teaching activity

Name PhD student:	Olivera Story-Jovanova
Erasmus Medical Center department:	Epidemiology
PhD period:	August 2012 – August 2017
Research School:	NIHES
Promotor:	Prof.dr. H. Tiemeier
Copromotor:	Dr. N. Amin

PhD Training	**Year**	**Workload (ECTS)**
General courses		
Research integrity	2014	2.0
Biomedical writing	2015	4.0
Specific courses (Research School)		
Master of science in Health Sciences, Epidemiology	2012-2013	75
Study Design	2012	4.3
Biostatistical Methods I: Basic Principles	2012	5.7
Clinical Epidemiology	2012	5.7
Methodologic Topics in Epidemiologic Research	2012	1.4
Biostatistical Methods II: Classical Regression Models	2012	4.3
Principles of Research in Medicine	2012	0.7
Methods of Public Health Research	2012	0.7
Cohort Studies	2012	0.7
Introduction to Global Public Health	2012	0.7
Primary and Secondary Prevention Research	2012	0.7
Social Epidemiology	2012	0.7
Courses for the Quantitative Researcher	2012	1.4
English Language	2012	1.1
Introduction to Medical Writing	2013	1.4
Causal Inference	2013	0.7
History of Epidemiologic Ideas	2013	0.7
Markers and Prognostic Research	2013	0.7
Advances in Epidemiologic Analysis	2013	0.4
Logistic Regression	2013	1.4
Psychiatric Epidemiology	2013	1.1
Women's Health	2013	0.9
Planning and Evaluation of Screening	2013	1.4
Chronic Disease Epidemiology, Institute of Public Health, University of Cambridge, United Kingdom	2013	1.1
Maternal and Child Health	2013	0.9
International conferences		
29th European College of Neuropsychopharmacology Congress Vienna, Austria (Poster presentation)	2016	1.0

18th meeting of the Section of Epidemiology and Social Psychiatry of the European Psychiatric Association EPA Gothenburg, Sweden (Oral presentation)	2016	1.0
Workshops, Meetings and symposia		
Research meetings, Psychiatric Rotterdam Study Group	2012-2017	1.0
Epidemiology research seminars	2012-2017	2.0
Epigenetic CHARGE research meetings	2015-2017	2.0
Teaching and Supervision		
Teaching Assistant,		
Erasmus Summer Program – Principles of Research in Medicine	2014	0.4
Supervision of medical students,		
- Supervising Cienne Poolmans, medical student, Department of Psychiatry, Erasmus MC. *Project title*: A sad heart : depression and cardiovascular disease a virtuous cycle.	2014	2.0
- Supervising Amin Achari, medical student, Department of Psychiatry, Erasmus MC. *Project title*: A sad heart : depression and cardiovascular disease a virtuous cycle.	2014	2.0
Supervision of master students,		
- Supervising Mayara Auler, master student, Department of Epidemiology, Erasmus MC. *Project title*: Cardiovascular, metabolic and renal biomarkers and their association with depression: a longitudinal population-based study of older adults.	2014-2015	2.0
Other activities		
Peer review Acta Scandinavica Psychiatrica	2017	0.2

1 ECTS (European Credit Transfer System) is equal to a workload of 28 hours.

CHAPTER 8.2

Publications and manuscripts

R1
R2
R3
R4
R5
R6
R7
R8
R9
R10
R11
R12
R13
R14
R15
R16
R17
R18
R19
R20
R21
R22
R23
R24
R25
R26
R27
R28
R29
R30
R31
R32
R33
R34
R35
R36
R37
R38
R39

List of publications

1. Zalli A, **Jovanova O,** Hoogendijk WJ, Tiemeier H, Carvalho LA. Low-grade inflammation predicts persistence of depressive symptoms. *Psychopharmacology*. 2016 May; 233(9):1669-78.

2. **Jovanova O**, Luik AI, Leening MJG, Noordam R, Aarts N, Hofman A, Franco OH, Dehghan A, Tiemeier H. The long-term risk of recognized and unrecognized myocardial infarction for depression in older men. *Psychological Medicine*. 2016 Mar; 46(9):1951-60.

3. Amin N, **Jovanova O,** Adams HH, Dehghan A, Kavousi M, Vernooji MW, Peeters RP, de Vrij FM, van der Lee SJ, van Rooij JG, van Leeuwen EM, Chaker L, Demirkan A, Hofman A, Brouwer RW, Kraaij R, Willems van Dijk K, Hankemeier T, van Ijcken WF, Uitterlinden AG, Niessen WJ, Franco OH, Kushner SA, Ikram MA, Tiemeier H, van Duijn CM. Exome-sequencing in a large population-based study reveals a rare Asn396Ser variant in the LIPG gene associated with depressive symptoms. *Molecular Psychiatry*. 2017 Apr; 22(4):537-543.

4. Amin N, Belonogova NM, **Jovanova O**, Brouwer RW, van Rooij JG, van den Hout MC, Svishcheva GR, Kraaij R, Zorkoltseva IV, Kirichenko AV, Hofman A, Uitterlinden AG, van IJcken WF, Tiemeier H, Axenovich TI, van Duijn CM. Nonsynonymous Variation in NKPD1 Increases Depressive Symptoms in the European Populations. *Biological Psychiatry*. 2017 Apr; 81(8):702-707.

5. **Jovanova O**, Aarts N, Noordam R, Carola-Zillikens M, Hofman A, Tiemeier H. Vitamin D serum levels are cross-sectionally but not prospectively associated with late-life depression. *ACTA Psychiatrica Scandinavica*. 2017 Mar; 135(3):185-194.

6. Amin N, de Vrij FMS, Baghdadi M, Brouwer RWW, van Rooij JGJ, **Jovanova O**, Uitterlinden AG, Hofman A, Janssen HLA, Darwish Murad S, Kraaij R, Stedehouder J, van den Hout MCGN, Kros JM, van IJcken WFJ, Tiemeier H, Kushner SA, van Duijn CM. A rare missense variant in RCL1 segregates with depression in extended families. *Molecular Psychiatry*. 2017 Mar; doi: 10.1038/mp.2017.49.

CHAPTER 8.3

Word of thanks

R1
R2
R3
R4
R5
R6
R7
R8
R9
R10
R11
R12
R13
R14
R15
R16
R17
R18
R19
R20
R21
R22
R23
R24
R25
R26
R27
R28
R29
R30
R31
R32
R33
R34
R35
R36
R37
R38
R39

Oh what a journey this was ☺! I am awaiting the birth of my son Aaron (who was supposed to arrive in this world 4 days ago) and completing this thesis by writing probably the most read section: the words of thanks. All of you who know me good enough, you know how important family and friends are to me. Private life is and it will be number one on my priority list. Without the love and support of the people I am surrounded with, I would have never achieved any of what I did today. You can imagine that the list of those who I have to thank will be a very long one; but I have to start somewhere. I will do it in an unconventional chronological way not to miss someone ☺.

Taking risks and grabbing what life brings was somehow my expertise since I know myself. Coming in Rotterdam and doing my PhD training was more of an adventure to me rather than living my dreams. To be honest, I have never even dreamt that I will get a Master's or a PhD Degree. Academic career was always a plan B for the old days. Yet, during my clinical work at Heliomedika, I realized I might need a change and I might have to try doing something new. I needed more knowledge, I simply wanted to study more. Thanks to my dear colleagues Dr. Vani, nurse Lile, Natka, and my ex-director and biggest motivator till this day Dr. Slavica; I decided to look for opportunities and challenged myself. I applied for Master's study at the University of Erasmus in Rotterdam supported by the ERAWEB grant. Thank you Heliomedika ladies, the working experience I had with you and the support you gave me through the whole application procedure means a lot to me. You all have a big part in all the decisions I have made then and without you, I am not sure I would have discovered myself and what my needs were with such an ease.

My application was accepted and I got admitted not only to a Master's program, but also got an opportunity to continue with PhD training. In August 2012, I was here in front of the Erasmus MC building yearning for all what this adventure would bring. The start was very difficult but I somehow managed it. For that, I would first like to thank the Ullmann Family. Mor, if it was not for you, I would have never came to the Netherlands at first. Thank you for that!! You were my start here, helping me integrate with the speed of light, supporting me during my beginner's troubles and whenever it was needed. You are the best ex-boyfriend someone can have. Bella what can I say, your Dutch Integration expertise helped me go through the last years smoothly. Thanks to you, I make the best soup in the world and not even one word of this book would have been written if it was not for the laptop I got from you. In short family Ullmann; you are the kick off to this PhD adventure, I own one to you.

Prof. Henning Tiemeier was one of the first people I met at Erasmus MC. The first things I learnt from him was not to shake in front of people who have a title professor. Therefore, Henning I always say if you can choose your mentor than choose for Henning ☺. You are a brilliant scientist

and a great teacher but for the most you are so human. Always down to earth, with a beautiful harsh German humor, it was a pleasure to have you as a leader in the working environment. I really appreciate that you always took in to account my personal circumstances when leading me through this PhD trajectory. You had that deep understanding and compassion for both work and private life and for that I give you all my credits. We are very different characters but somehow we managed working well together. Thank you for everything! I am happy to know you and even more happy to say that I have finish my PhD under your supervision. This thesis is a mutual product of both; your enthusiasm and stubbornness, and my patience and dedication.

Then, my co-promoter and mentor Najaf Amin. Dear Najaf, you are a role model to me. What Henning and I were missing to complete this project, was you. You were our balance and our tornado at the same time. Why tornado? I remember when you proposed to work together I thought: "Genetics, nooooo way I cannot do this." So yes your ideas came to me as a tornado. But you also had another skill, the skill of reading my mind all the time and you motivated me by saying "Don't worry genetics is not that difficult." You brought the balance in this PhD project. Learning and working with genetic data was not easy at all but with your systematic, organized, efficient and friendly way of working, everything became possible. Thank you, you are an example of a strong women and we women scientists should all follow your example.

I will never forget the first course I took that first summer with prof. Albert Hofman. Prof. Hofman you approved my application and I have learned the basics of epidemiology from you. Thank you for that. But Oh g-d during your course, everything was so new to me that I did not understood a word of what you were teaching. However, with joint efforts with my ERAWEB-NIHES friends, we managed to learn all and pass the Master's courses easily. Bibi, Dina, Kate, can you remember our long hours in the library? It was worth it. We have all managed to accomplish what we came here for and even better we have built friendships that will last forever. Thank you girls I am happy that these five years brought a lot of professional success but also extended family which you certainly are. Ana Maksimovic, Olja, Nasos, Gustavo, Ana Vitezova, and Janko you were my partners in crime from the very beginning. Owing to your presence, I did not forgot to smile, party and have fun during the first year of this PhD trajectory. Hope to meet you all once more and revive the memories we have built together.

During this PhD story, I was part of a great group: the PSYCH-EPI group. I had an opportunity to work with great colleagues, some of which I still work with even today and some will be marked as great friends whom I meet on regular basis. Annemarie you are such a friend! I have no words to explain it, but I can say you crashed my stereotype for Dutch friendship ☺. I love to talk with you and even more to gossip with you. I am even more glad that you are coming back to the NL so we can continue seeing each other again regularly. Also, I have to mention that thanks to you and Tom I have nice Dutch summary of this thesis. Dank je wel jongens.

Lisette, even though I have lost you somewhere on the road of motherhood, ☺ I can never forget our lunch moments in the park and the way you stress out every time I kiss you and hug you. Nese, my soulmate sister, you are the mini older me. Miss you a lot curly. Деси, сепак да останеме достоинствени на мајчниот Македонски јазик ☺. Ти и Кате сте ми моја Македонска банда. ☺ кога на балканот ќе кажеме крвта не ☺ вода, не ☺ за џабе кажано. Вие ми бевте најголема подршка и луѓе од доверба, другарки за се и сешто. Маките не споиа засекогаш ☺. Ве обожавам секоја на свој начин најискрено, и ви благодарам за се.

Ayesha my paranymph, what are you "women"? Are you my group colleague, my office-mate, my soulmate, my pregnancy/motherhood trouble partner, my friend, my shrink? I would say you are all of that. Henning was testing if a Jew and a Muslim can share an office, but he did not know he matched the only two strong crazy women in this world that can actually adore and love each other for real. Thank you Henning you have connected me with someone I know I will have near me for a life-time! Dear Ayesha two years ago, I was supporting you during this moment of glory and now there you are supporting me in the same. But that is what we do for each other. No matter what or where we are, we will always be there one for another. Thank you my princess Jasmine, you know how much I love you and appreciate you and sorry that I am so useless at the moment because of living so far from you.

The colleagues' circle is not small and there are some more people I have to thank. Saira, Heidi, Jelena, Marina, Akhghar, Mayara, Tom, Andrea, Else, Raluca, Nina, Ank, Rosanne, Jolien S, Irene, Dragan, Gogo, Ryan, Alex, Laura, Maartje B, Chantal, Roza, and Ingrid some of you I have worked with, some of you were my office-mates, some of you are just friends. What we have shared was Henning as a mentor, and thanks to all of you for making this period easier for me, since we have also shared our struggles that we faced during our scientific journeys. Fadila pa šta smo to nas dve preživele ☺? Kafa sa tobom mi je najdrža. Gde si sad našla da ideš tako daleko?! Nedostaješ jako jako. Frank, Vincent, Hieab, Hazel, Marileen, Rens, Renee, Natalie T., Trudy, Taulant, Olta, Jelena P., Clodian, Rene Sch., Emina, Jana N, Kozeta, Maarten B, Fjorda, and Sigi, sharing the elevator and the space with you which was a pleasure. Even better were all the chit-chat moments and all the taarten we shared these years.

This thesis is a product of joint work of many amazing people I have worked with. Prof. Oscar Franco, prof. Arfan Ikram, prof. Albert Hofman, prof. Cornelia van Duijn, prof. Wite Hoogendijk, prof. Norbert Schmitz, prof Miriam Fornage and the epigenetic group within CHARGE Consortium, Abbas, Maarten Leening, Wolters Frank, Carvalho Livia, Argita Zallili, Raymond Noordam, Nikkie Aarts, thank you for collaborating with me, I have learned a lot from all of you. Finally, I would like to thank Ivana Nedeljkovic. Ivana with joint forces we have finished my biggest and dearest project of this thesis. Thank you very much Ivana. I really appreciate your work, and want to let you know that you are very dear to me.

*

All these years, a large number of people helped me grow in this process in a different way. Most of them were coming in my life and going out of my life; a real dynamic cohort. However, there is a cluster of people that were always by my side and I know they will stay that way. In last part of this word of thanks, I would like to especially thank THEM; my family and my friends.

Ana Tasevska, Dafina, Marija, Ive, Rushka, you are certainly the footprint of my Skopje life. But, even though we are separated physically (viva la Viber) you are still a big part of my daily life. I miss you always. Chris bedankt for de cover and the funny discussions that we have. Helena, Marlies, Rachel you are my only friends in Alphen thank you for all the nice coffee/tea moments.

My Jewish friends: Lela, Dina, Mina, Dunja, Mavro, Sale, Urosh, Vase, Jana Nichota, Ruben, Liza, Andrea Steiner, Beca, Gabor, Pidja, Bojana, Sandra, Poli, Deborah, and Martin Levi; writing your names reminds me that I am the richest person on this world. Some of you I know since I am 10, that is long ☺. We all live in different parts of the world. Yet, our Jewish life connected us and made us who we all are today. You have shaped a big part of me and I am who I am also because of all of you. !תודה רבה

Лелка ти си ми другарица par-excellence! Толку сме исти, но спак толку различни, фала ти за сите долгочасовни разговори и што си дел од мојот живот. Дина и Мина другарице најдраже, па ви сте самном данас ☺. Нама је заједно увек тако добро. Нема тих речи које могу описати колко сам захвална да сте данас самном и што сте тако добре праве другарице. Јако сам поносна што вас имам.

A special word in English for my second paranymph, Dina. In 1996 I met you in Szarvas, Hungary, during a camp. You were from Serbia, me from Macedonia, both from freshly separated Yugoslavia. I am sure our parents did not have a clue what a big favor they did to us by sending us to that camp. Here we are now, in 2018 and still together, sharing the best and the worst in our lives. Living the past, the present and the future. Dina dear, I don't need you to be my paranymph to give you special place in my life, because you are so special to me. No one else in this world can be proud of such a friendship. Nothing can separate us because our lives are deeply connected. Thank you so much for being my paranymph and supporting me in this time of my life as you supported me in everything else so far (we should send JDC a copy of this book as a proof that their projects are not failures ☺: Lela I will send you a second copy highlighting these sentences and so that you can present it to the board of directors LoL). Дики љубим те!

My family Dedo Pero, Baba Vera, Dedo Ljupcho, Baba Tinche, Teta Lide, Tetin, Tea, Martin, Family Stojanovi, Family Shumkovski, Family Gligorovi, you are my roots and I am grateful that I have the opportunity to grow with so much love and as part of such a huge family. Дедо Перо биди горд на Оличка и сите твои внуци, ние сме твојата иднина и таа е светла. Баба Тинче ми фалиш најмногу и нема ден да не помислам на тебе. Хуморот и желбата за патувања ги наследив од тебе а тие нешта го чинат животот пресладок. Дедо Љупчо ти си непреболен, ама ете ја имам таа чест да те гледам повторно секој ден во очите на Арон. Баба Вера, фала ти на се а највеќе на тврдоглавоста што ја наследив од тебе. Да не беше таа немаше да постигнам волку многу. Тети, ти благодарам на сите моменти поминати со мене, знам дека ми се радуваш многу и ми е зал што не можеш да присуствуваш на одбраната на оваа теза. Бибе ти си ми единствена во Холандија и многу сум благодарна што те имам.

Mijn schoonfamilie: Opa Gerard, Oma Riek, Michael, Mariska, Sam, Nick en Jill Story. Ik ben trots dat ik een Story ben en dat ik deel mag uitmaken van jullie familie. Jullie zijn een grote steun in mijn dagelijks leven. Dank je wel voor alles. Opa, dank voor je onnavolgbare humor, dank je dat je zoveel van ons houdt. Oma dank je wel voor alle zorgen voor Ruth en alle hulp die je ons hebt geboden.

*

My Sister Eli. The biggest pain is that I am so far from you. Whatever I am busy with, I am constantly thinking: wish Eli would have been here. As I write this, I have tears in my eyes. Sorry if I was bothering you with my 100 daily calls. It is simply because I miss you so much. Thank you for being my sister, thank you for believing in me and always supporting me. Kufi thank you for taking care for my sister so much. Taking care for her means the same as if you have taken care for me. I am so sorry you are both not here for this day but it is for a good cause, awaiting another member of the Jovanov family.

Mama and Tato, I can write a whole book for you and it would still not be enough. You should be proud. Henning told me that he performed a research once, in which he was studying the offspring of couples that married in early age. The results of this research work were that no one that married at the age of 17 stayed happily married 33 years after with children that became doctors, master's in science and certainly not PhD. Conclusion, you are an exception. Thank you for all that you have done and you still do for me. Thank you for teaching me to be human, to be a good friend, wife, doctor, and most importantly a good mother! I am a mirror of your wishes, expectations, and ambitions. И малку на македонски; ви се извинувам што сум неподнослива кога ми е тешко. Ве сакам најмногу и сигурна сум дека не постојат други родители како вас!

*

And who shell I be grateful the most for all? My husband Egbert, of course. ☺ No matter how much I think, I just do not get to the right words to thank you my love. You are the biggest surprise in my life. I know that If someone could have told me before I met you, that I will be so loved by a man I would not have believed it. Our story deserves another book, and this thesis is just a small chapter of it. Schat, if it was not you there would have certainly not be a PhD either. That unconditional love we share makes us both so strong that we can take every challenge and solve every problem. That small paradise we live in, our home, is reason enough for me to move mountains. We have mastered depression for sure. Ik hou van je! I believe in karma and I am sure I must have done something very good in my life previously to get the opportunity to share the rest of my life with you. Dank je wel!

Finally, this is the end of my PhD fairytale. I must admit it took me 4 months to write these words of thanks ☺. It started when I was still pregnant with Aaron and here I am today 6th February,2018, and still writing it. It's time to finish it for good!

My life is an adventure that I am super proud of and happy to be writing about it. On 4th April I will become a PhD most probably but to be honest the title I am the most happy with is being a MOTHER. During this period, I was pregnant two times and I have become a mother of two kids. I got Ruth in the second year of my PhD and Aaron 4 months ago. They have marked this period by large. Therefore, I must write a few words to my children: Dear Ruth and Aaron, you are now so little and you can't read yet (Ruth knows the letter Н in Macedoninan only), but in a few years you can find these words and be even more proud of your mother. You are the best thing I have ever created. I am so proud of both of you. You are beautiful, sweet, funny, adorable and the biggest source of inspiration for me. You suck my energy out every day. I must admit that taking care of you is not easy. But those smiles and kisses gives me tremendous power to go further every day. Thank you for making my live meaningful and so happy. Ве сака мама!

Olivera Story Jovanova; 6th February 2018; Alphen

CHAPTER 8.4

About the author

®portrait Olivera Story-Jovanova made by Irena Mila

ABOUT THE AUTHOR

Olivera Story-Jovanova was born on April 23th 1986 in the city of Skopje, the Socialist Federal Republic of Yugoslavia, nowadays Republic of Macedonia. Her interest in Medical Sciences and Health Care began at a very early age. At the age of 14, she enrolled a special Nursing Medical High School "Dr. Pance Karagozov", in Skopje. After graduating from high school in 2004, she enrolled the Faculty of Medicine at the University of st. Cyril and Methodius also in Skopje. She graduated from the University in January 2011 and seven months later she obtained her licence to practice as a Medical Doctor in The Republic of Macedonia. She practiced medicine in the private psychiatric clinic for addictions and family concealing "Heliomedika" from October 2011 till July 2012.

In July 2012, she was admitted to a Master of Science program at the Netherland Institute for Health Sciences in Rotterdam, and moved to the Netherlands. She obtained her Master's degree in Epidemiology from the Erasmus University in August 2013. In August 2013, she began her PhD research work presented in this thesis under the supervision of Dr. Henning Tiemeier and Dr. Najaf Amin as a part of the Department of Epidemiology, the Psychiatric Epidemiologic group.

Her ambition is to return to clinical practice and to specialize in Psychiatry. In parallel efforts, she would like to proceed with her academic career in the field of Psychiatry and Psychiatric Epidemiology.

She is proud to be happily married to Egbertus Bernardus Gerardus Story, with who she has two beautiful children Ruth Hendrika Christine and Aaron Eli Michael Story.