(Epi) genetics and twins

van Dongen, J.

2015

Introduction

To date, the leading causes of disease burden in moderate and high income countries (based on the disability-adjusted-life-year metric [1]) are complex diseases, with cardiovascular disease being the number one contributor followed by mental disorders, in particular depression. Complex diseases often cluster in families yet they show no clear pattern of inheritance and are thought to result from the interplay of environmental influences and genetic susceptibility acting on multiple biological pathways. The underlying disease mechanisms of such traits may be best understood if they are approached from a variety of different angles. In this thesis, multiple approaches are applied to study the etiology of human complex traits, with an emphasis on metabolic syndrome traits and inflammation.

The metabolic syndrome is a condition that involves various metabolic disturbances that are risk factors for cardiovascular disease (CVD) and type 2 diabetes (T2D), including central obesity, insulin resistance, dyslipidemia, and hypertension [2-4]. It is well established that the metabolic syndrome is associated with a state of chronic, low-grade inflammation that is thought to contribute to the pathogenesis of the disease and to the risk for CVD and T2D, although inflammation itself is generally not included in diagnostic guidelines [5, 6]. The most prevalent common mental disorder is major depressive disorder. Cardiovascular disease and depressive disorders at least to some extent touch upon common biological pathways as evidenced by the fact that biological risk factors for cardiovascular disease are also connected to depression, including BMI (either high or low), abdominal obesity, and dysregulated levels of lipids and inflammation biomarkers in blood (including pro-inflammatory cytokines, cytokine receptors and acute phase reactive proteins) [7-9]. Therefore, a better understanding of the biological mechanisms that contribute to individual differences in the level of metabolic and inflammation biomarkers in blood[1], and in particular the role of genomic variation, could ultimately be of benefit to both physical and mental health.

Known risk factors for the metabolic syndrome (e.g. physical inactivity, smoking and obesity) clearly emphasize the role of lifestyle factors in the etiology [10-13]. In fact, it is thought that interventions targeted at these risk factors are largely responsible for a decline in the prevalence of cardiovascular events that has been observed in several countries in the past decades, although improved treatment may also have contributed [14-16]. The high prevalence of

[1]In this thesis, I will use the term metabolic biomarkers to refer to the concentration of lipids, glucose and insulin in blood. The term inflammation biomarkers is used to refer to inflammation components measured in blood (e.g. blood level of cytokines and cytokine receptors).

major depressive disorder in wealthy societies is less well understood, and depression has been estimated to become the world leading cause of disease burden in 2030[17].

The research described in this thesis is characterized by two major themes: the first relates to the question to which extent differences between individuals in inflammation biomarkers and metabolic syndrome traits are caused by the impacts of genetic and environmental differences between people. The second important focus of this thesis is on examining the genetic and non-genetic sources of variation underlying individual differences in DNA methylation; an epigenetic mechanism that receives increasing attention as it may provide novel insights into human disease. These research themes are the subject of a series of chapters in which I analyze phenotypic, genetic and epigenetic information that was collected in twin families who participate in longitudinal research of the Netherlands Twin Register. The etiology of individual differences is examined with approaches from genetic epidemiology, including the classical twin design and the extended twin-family design, and by approaches that make use of measured genetic variants. Chapter 3 of this thesis illustrates how these different approaches can be combined to unravel the contribution of known genetic (or candidate) variants versus unidentified genetic variants to a complex trait (soluble interleukin-6 receptor levels in blood). In this introduction, I present an outline of the topics covered in this thesis. I shortly introduce the current knowledge regarding metabolic syndrome traits and inflammation biomarkers and especially focus on epigenetics. It is relatively recent that large scale epigenetic studies (genome-wide) have become feasible and have been applied to twins. In this chapter I therefore pay detailed attention to epigenetic mechanisms and their potential involvement in human complex disease.

The metabolic syndrome

A diagnosis of the metabolic syndrome is often given to a person if at least three of the following traits exceed a certain clinical threshold: waist circumference, body mass index, waist-to-hip-ratio, fasting plasma glucose, fasting insulin level, systolic blood pressure (SBP), diastolic blood pressure (DBP), HDL cholesterol, and/or triglycerides. There are multiple diagnostic guidelines for the metabolic syndrome that include slightly different combinations of traits. These guidelines were established with the primary aim to identify people at high risk of developing CVD and T2D. The metabolic syndrome is associated with a doubled risk of developing CVD and a more than 5-fold increased risk of T2D, but also with increased risk for many other diseases, including nonalcoholic fatty liver disease, reproductive disorders, depression and sleeping disorders[18, 19]. The pathophysiological mechanisms that characterize the metabolic syndrome and are thought to contribute to the comorbid conditions include excess adipose tissue mass, ectopic fat deposition, excessive flux of fatty acids, and inflammation [3, 18, 20]. Although

there is ongoing debate about the constituents and primary causal mechanism underlying the metabolic syndrome, it is generally acknowledged that obesity (particularly in the abdominal area) and insulin resistance are the major underlying risk factors for the metabolic syndrome. Insulin resistance is often secondary to obesity but may also present in individuals with a normal weight [21].

Overall, as much as 14% of individuals who live in the Netherlands [22] and 37% of individuals in the United States [23] are affected by the metabolic syndrome, but the prevalence varies greatly with sex and age and depends on the guideline used to characterize the metabolic syndrome [3]. In the Netherlands, the metabolic syndrome was reported to be more prevalent in males according to most but not all guidelines [24], and to be more prevalent in older people [22, 25]. Of Dutch individuals of 65 years or older 37% of individuals meet the metabolic syndrome criteria [25]. Although traditionally viewed as an age-related disease, the metabolic syndrome is becoming increasingly prevalent in all age groups and can even be present in childhood [26-28]. The rising prevalence of the metabolic syndrome is thought to be the consequence of the rising prevalence of obesity[29].

Global obesity trends and changes in lifestyle

In the period between 1980 and 2008, the worldwide prevalence of obesity almost doubled; from 6.4% to 12% [30], although there is considerable variation between countries[31]. In the Netherlands, the prevalence of overweight and obesity was 47.8% and 16.2% in adults aged 20+, respectively, in 2008[31]. There is no doubt that the outbreak of overweight- and obesity-related disease is related to modern day lifestyle. Worldwide, the average caloric intake per person increased with 450 kcal per day between 1960 and the late 1990s [32], and 31.1% of adult individuals (age 15+) worldwide were “physically inactive” according to a publication from 2012 [33]. It has been suggested that obesity-related disease and a lot of other modern day complex diseases as well are relate to a mismatch between our current lifestyle in comparison with our evolutionary history, a topic I will we discuss in more detail in chapter 8. Yet, while we are all surrounded by weight-promoting influences in wealthy countries nowadays (e.g. environments and jobs that limit the need to be physically active and high calorie food available to everyone), not all individuals develop obesity and not every obese individual develops obesity-related diseases. It has been reported that 65% of obese males and 56% of obese females meet the criteria for the metabolic syndrome [34], while the rest has been characterized as displaying “metabolically healthy obesity” [35]. By contrast, “metabolically unhealthy non-obese” individuals also exist [21, 36]. In the United States 23.% of individuals (age 20+) with a normal BMI show two or more metabolic abnormalities (e.g. insulin resistance, dyslipidemia, hypertension) [37]. Thus, individuals differ in their vulnerability to develop overweight, obesity and metabolic disease.

Relevant metabolic and inflammation biomarkers in the population

Of note, the metabolic syndrome characterizes individuals at the extreme end of the distribution of variation in metabolic traits. In the following chapters of my thesis, I study variation in individual metabolic syndrome traits in a population-based sample of individuals, who were unselected with respect to disease status or health. The aim was to characterize the causes of individual differences underlying population variation in metabolic syndrome traits. Of note, several other tools exist to assess the risk of developing coronary heart disease or cardiovascular disease within the 'normal population'. The most widely applied Framingham score for coronary heart disease is based on traits that are commonly referred to as "traditional risk factors": sex, age, total cholesterol or LDL cholesterol, systolic blood pressure, smoking and self-reported diabetes [38, 39]. In addition to the 'classical lipids' (i.e. total cholesterol, HDL, LDL, triglycerides), many other types of circulating lipid particles as well as other molecules (e.g. amino acids) may be informative to metabolic health: such markers may be studied through metabolomics technology[40].

Although metabolic syndrome criteria and other risk scores do not commonly include inflammation biomarkers, growing evidence suggests that a key mechanism behind the pathogenesis of the consequences of obesity and associated conditions (e.g. type II diabetes and cardiovascular disease) involves chronic over-activation of cellular stress signalling and inflammatory pathways in metabolic cells in response to excessive energy intake[41, 42]. In obese individuals, adipose tissue secretes a range of a pro-inflammatory cytokines such as TNF-alpha and IL-6 [43, 44]. Pro-inflammatory cytokines stimulate the release of so-called acute phase reactants (for example fibrinogen and CRP) [45]. Other molecules in the inflammation cascade that are elevated in obese individuals include so-called sensors of the innate immune system including the inflammosome and Toll-like receptors [46, 47]. It is now well-known that in addition to adipose tissue, the liver, pancreas and brain also respond to obesity or excess energy intake by increased levels of inflammation, and this inflammation is highly important in the development of obesity-related disease[48-50]. Systemic low-grade inflammation is also known as the pro-inflammatory state.

In the Netherlands Twin Register, the levels of inflammation biomarkers in blood and traits included in metabolic syndrome criteria have been assessed in a representative population-based sample. This study sample allowed us to study the entire distribution of these traits to examine the causes of individual differences in these traits at the population level. In this thesis, I study the following inflammation biomarkers: tumor necrosis factor-alpha (TNF-α), C-reactive protein (CRP), fibrinogen, interleukin-6 (IL-6), and the soluble IL-6 receptor (sIL-6R). Metabolic syndrome traits included systolic and diastolic blood pressure, blood levels of glucose, insulin and lipids, and the

anthropometric traits weight, BMI, waist circumference and waist-to-hip ratio (WHR).

Genetics of BMI and biomarkers

Feeding experiments conducted on initially lean male prisoners in the 1960s illustrated the existence of individual differences in the effects of over-eating (~10,000 kcal per day)[51]. While most of the 20 men who managed to complete this feeding regime for a period of 200 days quickly lost the weight they gained as soon as the feeding regime turned back to normal, two men had great difficulties losing the weight [52]. These men were initially the fastest to gain weight and had a family history of obesity, although they had not previously been overweight themselves.

In societies in which the majority of the population has access to sufficient calories, individual differences in BMI are to a large extent explained by genetic factors. Adoption studies show that the BMIs of adopted individuals at middle age are more similar to their biological parents' BMI than to their adoptive parents' [53]. Monozygotic (MZ) twins usually have very similar BMIs, irrespective of whether the twins grow up together or are adopted by different families [54]. A number of twin and family studies have estimated the proportion of variation in BMI between individuals that can be attributed to genetic effects (the heritability); these estimates vary between approximately 24% and 90% [55]. Similar results hold for inflammation biomarkers and for metabolic biomarkers including blood lipid levels, fasting insulin and glucose and blood pressure. A large body of research shows that variation in these phenotypes is also largely influenced by genes [56]. In the past decade or so, a number of variants in the DNA sequence (risk alleles) have been identified that contribute to individual differences in these traits in humans [57], many of which are located in regulatory regions of the genome [58]. For example, the Genetic Investigation of ANthropometric Traits (GIANT) consortium has reported 32 independent genetic variants for BMI [59] and the Gobal Lipids Genetics Consortium has published 157 loci for blood lipid levels (including LDL cholesterol, HDL cholesterol, triglycerides and total cholesterol)[60]. Sixteen loci that are involved in fasting glucose homeostasis have been identified by the meta-analyses conducted by the meta-analyses of Glucose and Insulin-related traits Consortium (MAGIC)[61].

Beyond investigations into the DNA sequence itself, large cohort studies are now also starting to address at a genome-wide resolution other types of molecular variation by which variation in human complex traits is created. Thus, recent studies have reported associations of BMI [62], T2D [63, 64] and vascular disease [65] with epigenetic variation, that is, variation in molecular mechanisms that regulate to which extent genes need to be expressed depending on the internal (such as the degree of adiposity) and external environment of the person.

Epigenetic mechanisms

Nearly every human cell contains the same DNA sequence (genome) inherited from the parents, although there is increasing recognition that mosaicism may occur (meaning that an individual has cell populations with distinct genotypes, due to a postzygotic mutation that arose in one cell lineage)[66]. Yet many different cell types and organs are formed with the same sequence information, requiring different genes to be activated or inactivated (switched on and off) in each cell at the right time. This regulation of gene activity is coordinated in each cell by numerous molecular mechanisms that collectively control chromatin structure; including chemical tags attached to the DNA molecule itself and to the histone proteins it is wrapped around (i.e. DNA methylation and histone modifications), and molecules interacting with the DNA or RNA transcripts (e.g. non-coding RNAs, transcription factors, and methyl-CpG-binding proteins) [67]. While the field of genetics is traditionally concerned with the study of the DNA sequence, epigenetics refers to the study of those molecular mechanisms that influence gene expression without changing the DNA sequence and that are transmitted from one cell generation to the next through cell division (mitosis and possibly meiosis) [68]. In practice, the term epigenetic regulation is often applied more broadly, to refer to DNA methylation plus all histone modifications, although it is expected that histone modifications are not fully transmitted during cell division [69].

Whereas a person's DNA sequence remains the same during the entire lifetime (except for occasionally occurring *de novo* mutations), DNA methylation and other epigenetic marks are dynamic; they may change during the lifetime as part of developmentally regulated process (related to e.g. tissue differentiation) [70] and aging [71], and may change in response to specific environmental exposures [72]. For example, multiple studies have reported altered methylation levels of the *AHRR* gene and several other genes in blood cells of smokers (see for example [73]), and in babies of mothers who smoked during pregnancy [74]. A study of middle-aged individuals who were in the womb when their mother experienced severe famine during the Dutch Hunger Winter demonstrated that this exposure had left persistent changes in the methylation patterns at a diversity of genes, illustrating that environmental exposures can have long-term effects on DNA methylation [75]. Because DNA methylation may change over the life time and can respond to environmental exposures, studying this molecular layer of information may shed light on disease mechanisms that would remain hidden when focusing on the DNA sequence only. This mechanism also may shed light on the frequent discordance of monozygotic twin pairs for many complex traits and disorders.

Chromatin

If the DNA molecule would not be condensed it would not fit into the cell nucleus. Chromatin refers to the complex of the chromosomal DNA molecule and all attached histone proteins that facilitate to package the entire human genome (approximately 2 meters of DNA) into the cell nucleus (which has an

average diameter of 6 micrometers [76]). When viewed under an electron microscope, the 3D structure of chromatin in regions where DNA transcription occurs resembles “a string of beads”: a DNA string with so-called nucleosomes (‘beads’), which are units of about 146 base pairs of DNA that are wrapped 1.65 times around a complex of histone proteins [77-79]. Between each nucleosome are little pieces of DNA (stretching on average 20 base pairs) called linker DNA. Because the histone proteins are positively charged while the DNA molecule is negatively charged, histones help to fold the DNA into a smaller volume through electrostatic interactions. At genomic regions that are not transcribed, chromatin is further compacted into a highly condensed conformation through higher order structuring of the nucleosomes. The exact structure of chromatin, also called ‘chromatin state’ is fundamental to the expression potential of a region and is intimately linked to covalent post-translational modifications of the histone proteins.

Histone modifications

The histone protein complex within nucleosomes, also called “histone octamer”, consists of eight histone proteins: two copies of each of the histone types H2A, H2B, H3 and H4. A fifth type of histone (H1) can bind to the linker DNA between nucleosomes. Each histone protein is subject to numerous modifications at specific amino acids, including methylation, acetylation, phosphorylation and ubiquitination and many more. Most of these histone modifications occur at the “tails” (amino termini) of histones that stick out from the nucleosome complex. In total, there are more than 100 different histone modifications [80], most not well understood, which together control chromatin structure. For example, acetylation of the amino acid lysine is universally seen in transcriptionally active, accessible chromatin. Other modifications are associated with specific DNA elements. For example, histone H3 lysine 4 trimethylation (H3K4me3) is associated with promoters of actively transcribed genes [81]. It has been hypothesized that the exact chromatin state of a region lies written in the specific combination of histone modifications within that region (“the histone code”)[82] .

Histone modifications may exert their influence on chromatin structure by changing the electric charge of the histone (thereby modifying the strength of the interaction between the histone protein and the DNA molecule), and by recruiting other molecules to the DNA. Histone modifications and chromatin structure are mediated by a number of proteins that fall in one or more of the following categories: writers (responsible for depositing histone modifications, for example histone acetyl transferases; HATs [83]), erasers (which can remove modifications, for example histone deacetylases; HDACs), and readers (proteins that ‘interpret’ the histone code by recruiting additional molecules that can stabilize or remodel specific chromatin states, upon recognition of specific histone modifications).

Chromatin states: Insight into the genome-wide expression landscape
Transcription of DNA requires that transcription factors can bind to specific target sequences in the DNA (i.e. promoters). Whether a transcription factor can bind to its target depends on whether interaction partners such as co-activators and repressors are present, and on whether the local chromatin structure permits such interactions to take place, in other words, whether the DNA molecule is accessible [84]. In each cell, the structure of chromatin varies along the genome. In genomic regions with 'inactive chromatin' or 'heterochromatin', the chromatin is highly condensed (nucleosomes are tightly packaged) and transcription is silenced due to DNA and histone methylation, while other regions are characterized by an open chromatin state ('active chromatin' or 'euchromatin'), where the distance between nucleosomes is larger, thereby permitting transcription to occur [80]. It is thought that, in addition to active and inactive chromatin, chromatin comes in a number of intermediate states. For example, in permissive or repressed states, DNA is accessible to become de-repressed or activated through interaction with for example transcription factors. A recent study described fifteen distinct chromatin states that were observed in nine human cell types [85].

In addition to the study of histone modifications, important insight into gene regulation has come from mapping DNAse hypersensitive sites (DHSs) in the genome; DNA that is sensitive for being cut by DNAse enzymes. DNAse enzymes can only cut DNA in accessible chromatin, i.e. DNA that is situated between nucleosomes ("linker DNA"). DNAse experiments have revealed that this nucleosome-free DNA contains regulatory elements including promoters, enhancers, silencer, insulators and locus control regions. Importantly, only ~5% of accessible chromatin occurs within 2.5kb of transcriptional start sites, whereas ~95% of accessible chromatin is located in distal intronic and intergenic regions in human cells[86].

DNA methylation
DNA methylation, the covalent attachment of a methyl-group (CH_3) to the DNA molecule, is one of the best studied epigenetic mechanisms in humans, and is currently the only epigenetic mark that is suited for measuring in large-scale human epidemiological studies. In vertebrates, DNA methylation occurs mostly at cytosines located next to guanines (CpG sites). Non-CpG methylation is quite prevalent in human embryonic stem cells, but is very rare or absent in differentiated somatic cells [87-89]. DNA methylation is established and maintained by enzymes from the methyltransferase family, including DNMT1, DNMT3a and DNMT3a [90]. Some studies have suggested that *de novo* DNA methylation represents a passive process that targets regulatory sequences in the DNA that are not occupied by transcription factors [86, 91]. In total, an estimated 70-80% of CpG sites in the genome are methylated in mammalian cells [92]. Many CpGs occur in clusters called CpG islands (CGIs). The promoters of approximately 70% of human genes overlap with a CpG island.

The classic view is that methylation of CpGs in the promoter area of genes is associated with repression of gene expression. It has now become clear that the effect of methylation on expression may vary depending on the exact location that is methylated [93] and there are indications that methylation at enhancers is more strongly related to the expression level of genes than methylation at promoters [94].

It has been postulated that methylation may impact on regulation of transcription through two main mechanisms. Firstly, methylation of specific sequences may prevent the binding of regulator proteins to the DNA (e.g. transcription factors [95] and insulators [96]). Secondly, methylated CpGs may attract methyl-CpG-binding domain proteins, which are regulatory proteins that recruit chromatin regulators such as histone deacetylases and chromatin remodeling complexes to the site [97, 98]. Thus, rather than acting as independent mechanisms, it is thought that the different layers of epigenetic information (e.g. DNA methylation, histone modifications) generally work together to regulate transcription. Although the presence of DNA methylation correlates with the presence of repressive histone marks [99], there are exceptions. For example, most CpGs within promoter CGIs are unmethylated, but genes can be repressed through repressive histone marks even if their CGIs are unmethylated. Methylation of promoter CGIs is thought to contribute to long-term repression, for example at inactive X-chromosome genes in females [100] and imprinted genes [101]. In comparison with other epigenetic marks, it has been suggested that DNA methylation may be best described as a "memory signal for the long-term maintenance of gene silencing"[102]. Thus, it has been shown in colon cancer cells that while drugs that target histone modifications(histone deacetylase inhibitors (HDACi)) can lead to transient re-activation of loci silenced by DNA methylation, permanent re-activation can only be induced by DNA-demethylating drugs [102].

Although multiple techniques exist to measure DNA methylation, the Infinium humanMethylation 450 array (Illumina 450k) has become a popular platform in recent years for assessing DNA methylation at a genome-wide scale in human cohorts [103, 104]. This array assesses methylation level at ~485,000 CpG sites across a variety of regions in the human genome, including regulatory with genes and in intergenic regions.

Epigenetic gene regulation: link between environmental exposures and disease?

The question is to what extent epigenetic variation, by itself or in interaction with genetic variation, influences the metabolic and inflammatory biomarkers mentioned above and thereby the risk of metabolic diseases. Convincing evidence that epigenetic dysregulation can cause obesity in humans comes from rare neurodevelopmental disorders that are associated with obesity and result from imprinting defects, including Prader Willi syndrome [105] and Beckwith-Wiedemann Syndrome [106]. Importantly, epigenetic changes that arise

as a result of obesity may be involved in the development of obesity-related disease. A recent study examined the relationship between BMI and a group of CpGs where DNA methylation is strongly related to age [71] (also known as "the epigenetic clock") in multiple tissues, and found that people with a higher BMI show accelerated 'epigenetic aging' of liver tissue and 'epigenetically older livers' showed differential expression of a number of genes [107]. The findings suggest that epigenetic dysregulation in the liver may connect obesity to the development of age- and liver-related disease such as insulin resistance, although this hypothesis remains to be examined. Many studies have indicated the importance of epigenetic regulation within relevant disease tissues for metabolic syndrome traits and inflammation [63, 64, 108]. For example, cholesterol homeostasis was shown to be epigenetically regulated by a microRNA (miR-33) that controls the expression of genes involved in cholesterol transport including *ABCA1* [109]. The expression of the TNF-alpha locus is regulated through development and in response to e.g. lipopolysaccharide stimulation (LPS) by various epigenetic marks including DNA methylation and histone modifications [110]. To identify novel loci where epigenetic variation is related to disease risk, a lot of researchers at the moment are working on Epigenome-Wide Association Studies (EWAS) that test a large number of epigenetic marks along the genome for association with a (disease-) phenotype.

Epigenetic regulation of gene expression thus is as an important candidate mechanism that may mediate the effects of external influences (e.g. environmental exposures) on disease development. However, the other way other around, DNA methylation may also be influenced by the DNA sequence itself [111], and this heritable epigenetic variation between individuals may represent another pathway contributing to human disease susceptibility. In addition to EWAS, important understanding of how the epigenome may mediate variation in disease susceptibility may come from studies that are able to delineate the extent to which epigenetic variation between individuals in accessible tissues such as blood and buccal cells can be attributed to heritable mechanisms versus environmental exposures and stochastic effects.

Outline of this thesis

In the first part of this thesis, I study the genetics of inflammation biomarkers circulating in blood. Chapter 2 describes a study of the heritability of various components of the pro-inflammatory state based on an extended twin-family design. Chapter 3 focuses on variation in soluble IL-6 receptor levels. This chapter investigates the total heritability, the heritability explained by measured DNA sequence variants (SNPs), and characterizes the relationship between such SNPs and gene expression as a possible mode of action of the genetic variants.

In the second part of this thesis, the importance of genetic and environmental influences on BMI and other metabolic syndrome traits is examined. In chapter 4, the extended twin-family design is applied to estimate

the heritability of metabolic syndrome traits. In chapter 5, I examine the prevalence and development of BMI discordance over time in MZ twins, and examine whether discordance for body composition (BMI) in genetically identical subjects is associated with differences in metabolic and inflammation biomarkers and gene expression.

The third part of this thesis is devoted to studies that look beyond the information that lies within the primary DNA sequence, by exploring variation in DNA methylation. In chapter 6, I examine the etiology of individual differences in DNA methylation in peripheral blood, based on data from MZ and dizygotic (DZ) twins and their parents. In chapter 7, I examine variation in DNA methylation in buccal cells from MZ twins.

The fourth part of this thesis consists of two reviews covering theoretical background and future directions for studies on human complex traits. Chapter 8 introduces a topic that is crucial for understanding the origin of common human disease; the topic of evolution. In this chapter, I review evolutionary perspectives of schizophrenia, a common mental disorder. This chapter explains the links between the evolutionary history of traits and their genetic architecture, and the implications thereof for gene finding studies. Chapter 9 explains the value of twins for studying the sources of variation in human traits and discusses the role that twin studies may play in future . Two designs based on data from twins that are used in this thesis; the estimation of heritability and the discordant MZ twin design, and their application in the context of ‘omic’ data, including studies of biomarkers, gene expression and epigenetics are reviewed in chapter 9.

In chapter 10, I summarize the most important results from all chapters and discusses these findings in the broader context of the current state and future directions of research on complex trait genetics.

Reference List

1. Mathers,C., Fat,D.M., & Boerma,J.T. *The global burden of disease: 2004 update*(World Health Organization,2008).
2. Alberti,K.G., Zimmet,P., & Shaw,J. The metabolic syndrome--a new worldwide definition. *Lancet* **366**, 1059-1062 (2005).
3. Eckel,R.H., Grundy,S.M., & Zimmet,P.Z. The metabolic syndrome. *The Lancet* **365**, 1415-1428 (2005).
4. Eckel,R.H., Alberti,K.G., Grundy,S.M., & Zimmet,P.Z. The metabolic syndrome. *Lancet* **375**, 181-183 (2010).
5. Haffner,S.M. The metabolic syndrome: inflammation, diabetes mellitus, and cardiovascular disease. *Am. J. Cardiol.* **97**, 3A-11A (2006).
6. Kassi,E., Pervanidou,P., Kaltsas,G., & Chrousos,G. Metabolic syndrome: definitions and controversies. *BMC. Med.* **9**, 48 (2011).
7. Lasserre,A.M. *et al.* Depression with atypical features and increase in obesity, body mass index, waist circumference, and fat mass: a prospective, population-based study. *JAMA Psychiatry* **71**, 880-888 (2014).
8. Penninx,B.W., Milaneschi,Y., Lamers,F., & Vogelzangs,N. Understanding the somatic consequences of depression: biological mechanisms and the role of depression symptom profile. *BMC. Med.* **11**, 129 (2013).
9. Vogelzangs,N., Comijs,H.C., Oude Voshaar,R.C., Stek,M.L., & Penninx,B.W. Late-life depression symptom profiles are differentially associated with immunometabolic functioning. *Brain Behav. Immun.* **41**, 109-115 (2014).
10. Dhingra,R. *et al.* Soft drink consumption and risk of developing cardiometabolic risk factors and the metabolic syndrome in middle-aged adults in the community. *Circulation* **116**, 480-488 (2007).
11. Lakka,T.A. *et al.* Sedentary lifestyle, poor cardiorespiratory fitness, and the metabolic syndrome. *Med. Sci. Sports Exerc.* **35**, 1279-1286 (2003).
12. McKeown,N.M. *et al.* Carbohydrate nutrition, insulin resistance, and the prevalence of the metabolic syndrome in the Framingham Offspring Cohort. *Diabetes Care* **27**, 538-546 (2004).
13. Park,Y.W. *et al.* The metabolic syndrome: prevalence and associated risk factor findings in the US population from the Third National Health and Nutrition Examination Survey, 1988-1994. *Arch. Intern. Med.* **163**, 427-436 (2003).
14. Berry,J.D. *et al.* Lifetime risks of cardiovascular disease. *N. Engl. J. Med.* **366**, 321-329 (2012).
15. Ford,E.S. *et al.* Explaining the decrease in U.S. deaths from coronary disease, 1980-2000. *N. Engl. J. Med.* **356**, 2388-2398 (2007).
16. Wijeysundera,H.C. *et al.* Association of temporal trends in risk factors and treatment uptake with coronary heart disease mortality, 1994-2005. *JAMA* **303**, 1841-1847 (2010).
17. Mathers,C.D. & Loncar,D. Projections of global mortality and burden of disease from 2002 to 2030. *PLoS. Med.* **3**, e442 (2006).
18. Cornier,M.A. *et al.* The metabolic syndrome. *Endocr. Rev.* **29**, 777-822 (2008).
19. Vogelzangs,N. *et al.* Metabolic depression: a chronic depressive subtype? Findings from the InCHIANTI study of older persons. *J Clin. Psychiatry* **72**, 598-604 (2011).
20. Grundy,S.M. The Metabolic Syndrome in *Atlas of Atherosclerosis and Metabolic Syndrome* (ed. Grundy,S.M.) 1-26 (Springer, New York, 2011).

21. Ruderman,N., Chisholm,D., Pi-Sunyer,X., & Schneider,S. The metabolically obese, normal-weight individual revisited. *Diabetes* **47**, 699-713 (1998).
22. Bos,M.B. *et al.* The prevalence of the metabolic syndrome in the Netherlands: increased risk of cardiovascular diseases and diabetes mellitus type 2 in one quarter of persons under 60. *Nederlands tijdschrift voor geneeskunde* **151**, 2382-2388 (2007).
23. Cameron,A.J., Shaw,J.E., & Zimmet,P.Z. The metabolic syndrome: prevalence in worldwide populations. *Endocrinology and metabolism clinics of North America* **33**, 351-375 (2004).
24. Dekker,J.M. *et al.* Metabolic syndrome and 10-year cardiovascular disease risk in the Hoorn Study. *Circulation* **112**, 666-673 (2005).
25. Oosterwerff,M.M., van Schoor,N.M., Lips,P., & Eekhoff,E.M. Osteocalcin as a predictor of the metabolic syndrome in older persons: a population-based study. *Clin. Endocrinol. (Oxf)* **78**, 242-247 (2013).
26. Weiss,R. *et al.* Obesity and the metabolic syndrome in children and adolescents. *N. Engl. J. Med.* **350**, 2362-2374 (2004).
27. Friend,A., Craig,L., & Turner,S. The prevalence of metabolic syndrome in children: a systematic review of the literature. *Metab Syndr. Relat Disord.* **11**, 71-80 (2013).
28. Fazeli,F.S., van der Aa,M.P., van der Vorst,M.M., Knibbe,C.A., & de,B.A. Global trends in the incidence and prevalence of type 2 diabetes in children and adolescents: a systematic review and evaluation of methodological approaches. *Diabetologia* **56**, 1471-1488 (2013).
29. National Institutes of Health. http://www.nhlbi.nih.gov/health/health-topics/topics/ms/, Accessed on 03/11/. 2014.
30. Stevens,G.A. *et al.* National, regional, and global trends in adult overweight and obesity prevalences. *Popul. Health Metr.* **10**, 22 (2012).
31. World Health Organization. http://www.who.int/gho/ncd/risk_factors/overweight/en/, Accessed on 12/06/. 2014.
32. Food and Agriculture Organization of the United Nations. World agriculture: towards 2015/2030 (Summary report). 2012. Rome.
33. Hallal,P.C. *et al.* Global physical activity levels: surveillance progress, pitfalls, and prospects. *Lancet* **380**, 247-257 (2012).
34. Ervin,R.B. Prevalence of metabolic syndrome among adults 20 years of age and over, by sex, age, race and ethnicity, and body mass index: United States, 2003-2006. *Natl. Health Stat. Report.*1-7 (2009).
35. Hamer,M. & Stamatakis,E. Metabolically healthy obesity and risk of all-cause and cardiovascular disease mortality. *J. Clin. Endocrinol. Metab* **97**, 2482-2488 (2012).
36. Conus,F., Rabasa-Lhoret,R., & Peronnet,F. Characteristics of metabolically obese normal-weight (MONW) subjects. *Appl. Physiol Nutr. Metab* **32**, 4-12 (2007).
37. Wildman,R.P. *et al.* The obese without cardiometabolic risk factor clustering and the normal weight with cardiometabolic risk factor clustering: prevalence and correlates of 2 phenotypes among the US population (NHANES 1999-2004). *Arch. Intern. Med.* **168**, 1617-1624 (2008).
38. Anderson,K.M., Wilson,P.W., Odell,P.M., & Kannel,W.B. An updated coronary risk profile. A statement for health professionals. *Circulation* **83**, 356-362 (1991).
39. Wilson,P.W. *et al.* Prediction of coronary heart disease using risk factor categories. *Circulation* **97**, 1837-1847 (1998).
40. Vaarhorst,A.A. *et al.* A metabolomic profile is associated with the risk of incident coronary heart disease. *Am. Heart J.* **168**, 45-52 (2014).

41. Ozcan,U. *et al.* Endoplasmic reticulum stress links obesity, insulin action, and type 2 diabetes. *Science* **306**, 457-461 (2004).
42. Stienstra,R., Tack,C.J., Kanneganti,T.D., Joosten,L.A., & Netea,M.G. The inflammasome puts obesity in the danger zone. *Cell Metab* **15**, 10-18 (2012).
43. Berg,A.H. & Scherer,P.E. Adipose tissue, inflammation, and cardiovascular disease. *Circ. Res.* **96**, 939-949 (2005).
44. Hotamisligil,G.S., Arner,P., Caro,J.F., Atkinson,R.L., & Spiegelman,B.M. Increased adipose tissue expression of tumor necrosis factor-alpha in human obesity and insulin resistance. *J. Clin. Invest* **95**, 2409-2415 (1995).
45. Shoelson,S.E., Lee,J., & Goldfine,A.B. Inflammation and insulin resistance. *J. Clin. Invest* **116**, 1793-1801 (2006).
46. Schroder,K., Zhou,R., & Tschopp,J. The NLRP3 inflammasome: a sensor for metabolic danger? *Science* **327**, 296-300 (2010).
47. Shi,H. *et al.* TLR4 links innate immunity and fatty acid-induced insulin resistance. *J. Clin. Invest* **116**, 3015-3025 (2006).
48. Cai,D. *et al.* Local and systemic insulin resistance resulting from hepatic activation of IKK-beta and NF-kappaB. *Nat. Med.* **11**, 183-190 (2005).
49. De Souza,C.T. *et al.* Consumption of a fat-rich diet activates a proinflammatory response and induces insulin resistance in the hypothalamus. *Endocrinology* **146**, 4192-4199 (2005).
50. Ehses,J.A. *et al.* Increased number of islet-associated macrophages in type 2 diabetes. *Diabetes* **56**, 2356-2370 (2007).
51. Sims,E.A. *et al.* Experimental obesity in man. *Trans. Assoc. Am. Physicians* **81**, 153-170 (1968).
52. Shell E.R. Hunger in *The hungry gene: the inside story of the obesity industry.* (Grove Press, New York, 2003).
53. Stunkard,A.J. *et al.* An adoption study of human obesity. *N. Engl. J. Med.* **314**, 193-198 (1986).
54. Stunkard,A.J., Harris,J.R., Pedersen,N.L., & McClearn,G.E. The body-mass index of twins who have been reared apart. *N. Engl. J. Med.* **322**, 1483-1487 (1990).
55. Elks,C.E. *et al.* Variability in the heritability of body mass index: a systematic review and meta-regression. *Front Endocrinol. (Lausanne)* **3**, (2012).
56. Pilia,G. *et al.* Heritability of cardiovascular and personality traits in 6,148 Sardinians. *PLoS Genet* **2**, e132 (2006).
57. Visscher,P.M., Brown,M.A., McCarthy,M.I., & Yang,J. Five years of GWAS discovery. *Am. J. Hum. Genet.* **90**, 7-24 (2012).
58. Hindorff,L.A. *et al.* Potential etiologic and functional implications of genome-wide association loci for human diseases and traits. *Proc. Natl. Acad. Sci. U. S. A* **106**, 9362-9367 (2009).
59. Speliotes,E.K. *et al.* Association analyses of 249,796 individuals reveal 18 new loci associated with body mass index. *Nat. Genet.* **42**, 937-948 (2010).
60. Global Lipids Genetics Consortium Discovery and refinement of loci associated with lipid levels. *Nat. Genet.* **45**, 1274-1283 (2013).
61. Dupuis,J. *et al.* New genetic loci implicated in fasting glucose homeostasis and their impact on type 2 diabetes risk. *Nat. Genet.* **42**, 105-116 (2010).
62. Dick,K.J. *et al.* DNA methylation and body-mass index: a genome-wide analysis. *Lancet* **383**, 1990-1998 (2014).

63. Dayeh,T. *et al.* Genome-wide DNA methylation analysis of human pancreatic islets from type 2 diabetic and non-diabetic donors identifies candidate genes that influence insulin secretion. *PLoS. Genet.* **10**, e1004160 (2014).
64. Nilsson,E. *et al.* Altered DNA methylation and differential expression of genes influencing metabolism and inflammation in adipose tissue from subjects with type 2 diabetes. *Diabetes* **63**, 2962-2976 (2014).
65. Connelly,J.J. *et al.* Epigenetic regulation of COL15A1 in smooth muscle cell replicative aging and atherosclerosis. *Hum. Mol. Genet.* **22**, 5107-5120 (2013).
66. Biesecker,L.G. & Spinner,N.B. A genomic view of mosaicism and human disease. *Nat. Rev. Genet.* **14**, 307-320 (2013).
67. Goldberg,A.D., Allis,C.D., & Bernstein,E. Epigenetics: a landscape takes shape. *Cell* **128**, 635-638 (2007).
68. Bird,A. Perceptions of epigenetics. *Nature* **447**, 396-398 (2007).
69. Bernstein,B.E., Meissner,A., & Lander,E.S. The mammalian epigenome. *Cell* **128**, 669-681 (2007).
70. Reik,W. Stability and flexibility of epigenetic gene regulation in mammalian development. *Nature* **447**, 425-432 (2007).
71. Horvath,S. DNA methylation age of human tissues and cell types. *Genome Biol.* **14**, R115 (2013).
72. Feil,R. & Fraga,M.F. Epigenetics and the environment: emerging patterns and implications. *Nat. Rev. Genet.* **13**, 97-109 (2011).
73. Philibert,R.A., Beach,S.R., Lei,M.K., & Brody,G.H. Changes in DNA methylation at the aryl hydrocarbon receptor repressor may be a new biomarker for smoking. *Clin. Epigenetics.* **5**, 19 (2013).
74. Joubert,B.R. *et al.* 450K epigenome-wide scan identifies differential DNA methylation in newborns related to maternal smoking during pregnancy. *Environ. Health Perspect.* **120**, 1425-1431 (2012).
75. Heijmans,B.T. *et al.* Persistent epigenetic differences associated with prenatal exposure to famine in humans. *Proc. Natl. Acad. Sci. U. S. A* **105**, 17046-17049 (2008).
76. Alberts,B. *et al. Molecular Biology of the Cell 4th Edition: International Student Edition*(Routledge,2002).
77. Luger,K., Mader,A.W., Richmond,R.K., Sargent,D.F., & Richmond,T.J. Crystal structure of the nucleosome core particle at 2.8 A resolution. *Nature* **389**, 251-260 (1997).
78. Olins,A.L. & Olins,D.E. Spheroid chromatin units (v bodies). *Science* **183**, 330-332 (1974).
79. Olins,D.E. & Olins,A.L. Chromatin history: our view from the bridge. *Nat. Rev. Mol. Cell Biol.* **4**, 809-814 (2003).
80. Kouzarides,T. Chromatin modifications and their function. *Cell* **128**, 693-705 (2007).
81. Santos-Rosa,H. *et al.* Active genes are tri-methylated at K4 of histone H3. *Nature* **419**, 407-411 (2002).
82. Jenuwein,T. & Allis,C.D. Translating the histone code. *Science* **293**, 1074-1080 (2001).
83. Brown,C.E., Lechner,T., Howe,L., & Workman,J.L. The many HATs of transcription coactivators. *Trends Biochem. Sci.* **25**, 15-19 (2000).
84. Li,B., Carey,M., & Workman,J.L. The role of chromatin during transcription. *Cell* **128**, 707-719 (2007).

85. Ernst,J. *et al.* Mapping and analysis of chromatin state dynamics in nine human cell types. *Nature* **473**, 43-49 (2011).
86. Thurman,R.E. *et al.* The accessible chromatin landscape of the human genome. *Nature* **489**, 75-82 (2012).
87. Laurent,L. *et al.* Dynamic changes in the human methylome during differentiation. *Genome Res.* **20**, 320-331 (2010).
88. Lister,R. *et al.* Human DNA methylomes at base resolution show widespread epigenomic differences. *Nature* **462**, 315-322 (2009).
89. Ramsahoye,B.H. *et al.* Non-CpG methylation is prevalent in embryonic stem cells and may be mediated by DNA methyltransferase 3a. *Proc. Natl. Acad. Sci. U. S. A* **97**, 5237-5242 (2000).
90. Goll,M.G. & Bestor,T.H. Eukaryotic cytosine methyltransferases. *Annu. Rev. Biochem.* **74**, 481-514 (2005).
91. Stadler,M.B. *et al.* DNA-binding factors shape the mouse methylome at distal regulatory regions. *Nature* **480**, 490-495 (2011).
92. Bird,A., Taggart,M., Frommer,M., Miller,O.J., & Macleod,D. A fraction of the mouse genome that is derived from islands of nonmethylated, CpG-rich DNA. *Cell* **40**, 91-99 (1985).
93. Jones,P.A. Functions of DNA methylation: islands, start sites, gene bodies and beyond. *Nat. Rev. Genet.* **13**, 484-492 (2012).
94. Aran,D., Sabato,S., & Hellman,A. DNA methylation of distal regulatory sites characterizes dysregulation of cancer genes. *Genome Biol.* **14**, R21 (2013).
95. Watt,F. & Molloy,P.L. Cytosine methylation prevents binding to DNA of a HeLa cell transcription factor required for optimal expression of the adenovirus major late promoter. *Genes Dev.* **2**, 1136-1143 (1988).
96. Hark,A.T. *et al.* CTCF mediates methylation-sensitive enhancer-blocking activity at the H19/Igf2 locus. *Nature* **405**, 486-489 (2000).
97. Nan,X. *et al.* Transcriptional repression by the methyl-CpG-binding protein MeCP2 involves a histone deacetylase complex. *Nature* **393**, 386-389 (1998).
98. Wade,P.A. *et al.* Mi-2 complex couples DNA methylation to chromatin remodelling and histone deacetylation. *Nat. Genet.* **23**, 62-66 (1999).
99. Fuks,F. DNA methylation and histone modifications: teaming up to silence genes. *Curr. Opin. Genet. Dev.* **15**, 490-495 (2005).
100. Yasukochi,Y. *et al.* X chromosome-wide analyses of genomic DNA methylation states and gene expression in male and female neutrophils. *Proc. Natl. Acad. Sci. U. S. A* **107**, 3704-3709 (2010).
101. Choufani,S. *et al.* A novel approach identifies new differentially methylated regions (DMRs) associated with imprinted genes. *Genome Res.* **21**, 465-476 (2011).
102. Raynal,N.J. *et al.* DNA methylation does not stably lock gene expression but instead serves as a molecular mark for gene silencing memory. *Cancer Res.* **72**, 1170-1181 (2012).
103. Bibikova,M. *et al.* High density DNA methylation array with single CpG site resolution. *Genomics* **98**, 288-295 (2011).
104. Sandoval,J. *et al.* Validation of a DNA methylation microarray for 450,000 CpG sites in the human genome. *Epigenetics.* **6**, 692-702 (2011).
105. Ohta,T. *et al.* Imprinting-mutation mechanisms in Prader-Willi syndrome. *Am. J. Hum. Genet.* **64**, 397-413 (1999).

106. Weksberg,R. *et al.* Discordant KCNQ1OT1 imprinting in sets of monozygotic twins discordant for Beckwith-Wiedemann syndrome. *Hum. Mol. Genet.* **11**, 1317-1325 (2002).
107. Horvath,S. *et al.* Obesity accelerates epigenetic aging of human liver. *Proc. Natl. Acad. Sci. U. S. A*(2014).
108. Sinclair,K.D. *et al.* DNA methylation, insulin resistance, and blood pressure in offspring determined by maternal periconceptional B vitamin and methionine status. *Proc. Natl. Acad. Sci. U. S. A* **104**, 19351-19356 (2007).
109. Rayner,K.J. *et al.* MiR-33 contributes to the regulation of cholesterol homeostasis. *Science* **328**, 1570-1573 (2010).
110. Sullivan,K.E. *et al.* Epigenetic regulation of tumor necrosis factor alpha. *Mol. Cell Biol.* **27**, 5147-5160 (2007).
111. Bell,J.T. *et al.* DNA methylation patterns associate with genetic and gene expression variation in HapMap cell lines. *Genome Biol.* **12**, R10 (2011).