Molecular genetic research on personality

van den Berg, Stephanie M.; de Moor, Marleen H.M.

published in
Behavior Genetics of Temperament and Personality
2020

DOI (link to publisher)
10.1007/978-1-0716-0933-0_4

document version
Publisher's PDF, also known as Version of record

Chapter 4
Molecular Genetic Research on Personality

Stéphanie M. van den Berg and Marleen H. M. de Moor

Introduction

The view that personality has a biological origin traces back to the ancient Greeks. Hippocrates postulated as early as 400 years B.C. that humans can be classified into four different personality types: sanguine (optimistic, hopeful), melancholic (sad, depressed), choleric (irascible), and phlegmatic (apathetic) (Merenda, 1987). Hippocrates stated that these different personality types are linked to different bodily fluid systems, also called humors: the blood for the sanguine type, black bile for the melancholic type, yellow bile for the choleric type, and phlegm for the phlegmatic type. A balance between these bodily fluids was thought to be needed for a good health.

Although the personality typology of Hippocrates has undoubtedly influenced contemporary scientific ideas about the *structure* of personality, the idea of bodily fluids as the *biological origin* of personality has clearly been discarded. Instead, modern science has turned toward the brain as the seat of our personality. In order to understand individual *differences* in personality, and in particular, the origin of these personality differences, studies on the biological, hereditary basis of human individual differences in personality have expanded to include, besides twin and family studies, molecular genetic studies. These molecular genetic studies try to link variation in the human DNA to individual differences in personality traits. In this chapter, we will provide an overview of this molecular genetic work.

S. M. van den Berg (✉)
Department of Research Methodology, Measurement and Data Analysis, University of Twente, Enschede, The Netherlands
e-mail: stephanie.vandenberg@utwente.nl

M. H. M. de Moor
Section of Clinical Child and Family Studies, Amsterdam Public Health Research Institute, Vrije Universiteit Amsterdam, Amsterdam, The Netherlands

K. J. Saudino, J. M. Ganiban (eds.), *Behavior Genetics of Temperament and Personality*, Advances in Behavior Genetics,
https://doi.org/10.1007/978-1-0716-0933-0_4

The chapter is structured as follows. First, we will describe how historical breakthroughs in our knowledge about the human genome, including novel scientific techniques to study the genome, have enabled molecular genetic research on personality. Second, we will explain in more detail the methodology behind the most commonly used types of molecular genetic studies. Next, we will proceed with reviewing the findings from molecular genetic studies on personality. Did we succeed in finding genes that can explain the individual differences in personality across humans? We will close with a general conclusion about the current state of our knowledge and a discussion about future research avenues regarding molecular genetic research on personality.

History

Molecular genetic research on personality aims to investigate whether there is a relationship between measured variations in human DNA and personality traits. Variation in DNA comes in different forms, and techniques to measure this variation have not always been available in the past. In this section, we provide a brief historical account of molecular genetic research applied to the study of personality (Table 4.1).

The seminal work of Gregory Mendel published in 1866 provides the foundation of modern genetic research (Mendel, 1866). As a botanist, Mendel conducted a series of experiments in his garden and discovered that the inheritance of certain characteristics of pea plants, such as whether they were wrinkled or smooth, followed a specific pattern. He concluded that there must be some "heredity units" that are passed on from one generation to the next in a random manner, which became known as the Law of Segregation, one of Mendel's three laws on the principles of

Table 4.1 History of molecular genetic research on personality

Year	Event
1865	Gregor Mendel discovered that heredity is transmitted in "units"
1909	Wilhelm Johannsen coined the word "gene" to describe Mendel's heredity units
1918	Ronald Fisher proposed polygenic model of inheritance
1953	James Watson and Francis Crick discovered the double helix structure of DNA
1980	Genetic linkage analysis using anonymous DNA polymorphisms was proposed
1983	First gene discovery for a Mendelian disease (Huntington's disease)
1996	First candidate gene studies to link genes to neuroticism and novelty seeking
1998	First genetic linkage study for neuroticism
2000	First working draft of the human genome project was presented
2003	Human genome project was completed
2008	The first two genome-wide association studies for neuroticism were published
2012	First SNP-heritability studies and first meta-analytic genome-wide association study for five personality traits

inheritance. In the decades that followed (roughly between 1865 and 1950), scientists have made many important discoveries that collectively constitute our current vital knowledge about these "heredity units" that we nowadays call genes, a term first coined by Wilhelm Johannsen in 1909 (Johanssen, 1909). Genes are specific segments of DNA molecules that are functional and code for proteins via the processes of transcription and translation (gene expression). Humans have 46 of such DNA molecules: two copies of 22 autosomes and 2 sex-specific chromosomes (XX for females and XY for males). Together our chromosomes carry approximately 20,000–25,000 genes. The chromosomes, and hence the genes located on them, are found in each cell nucleus of the human body. Each chromosome consists of two interconnected strands of nucleotides, also known as the double helix, which was first described by Watson and Crick in 1953 (Watson & Crick, 1953). The nucleotides on these strands contain different chemical groups: a sugar molecule, a phosphate group, and one of four nitrogenous bases: adenine (A), cytosine (C), guanine (G), and thymine (T). Together with the mitochondrial DNA, the chromosomal DNA constitutes the human genome and contains all the genetic information in a person. The human genome is approximately 99.9% identical across all humans. The remaining 0.01% contains the genetic variation that is responsible for individual variation in human characteristics, such as disease status, physical appearance, and complex traits like personality. To summarize, most of our basic knowledge about DNA derives from a series of discoveries made in the first half of the twentieth century. It was only in the second half of this century that it became possible to study more systematically the links between DNA variation and numerous phenotypes, starting with hereditary diseases but soon also more complex traits like personality.

In 1980, researchers for the first time discovered actual polymorphisms in the DNA, which are pieces of DNA that show variation within a population, which led to an explosion of two types of studies linking DNA variation to a particular phenotype between the 1980s and 2000: genetic linkage studies and candidate gene studies (explained in more detail in the next section). In 1983, a human disease, Huntington disease, was first linked to such a polymorphic DNA marker (Gusella et al., 1983). With regard to personality, it was 13 years later that the first three candidate gene association studies were published (Benjamin et al., 1996; Ebstein et al., 1996; Lesch et al., 1996). Two years after that, the first genome-wide linkage study for personality appeared (Cloninger et al., 1998).

In the meantime, the Human Genome Project (HGP) was launched in 1990, which aimed to sequence the entire human genome in 15 years of time. A first draft of the human genome was presented by the HGP in 2000 (Lander et al., 2001). Three years later, in 2003, the HGP was considered complete, having sequenced 99% of the human genome with 99.9% accuracy (Consortium, 2004). This important accomplishment paved the way for genome-wide association (GWA) studies, in which a large number of genetic variants across the genome, typically in the order of hundreds of thousands or even millions, can be tested for association with a phenotype of interest. Five years after completion of the HGP, the first two GWA studies for personality appeared (Shifman et al., 2008; van den Oord et al., 2008). Since

then, gene-finding studies for personality have increased in terms of sample size and number of variants studied and have also included some novel techniques estimating the heritability of personality traits based on measured genomic variation. In the next section, we will explain in somewhat more detail the methodological principles behind the molecular genetic methods currently available (candidate gene studies, linkage studies, GWA studies and beyond). In the section thereafter, we will discuss what the application of these methods to the study of personality has yielded in terms of gene-discovery results.

Methodology

The object of molecular genetics is to identify the regions on the genome that are responsible for the *variation* in a trait. That means that researchers look for variations in the four nitrogenous bases indicated by the letters A, C, G, and T in genomic regions. All of the methods are based on the phenomena of *segregation* and *linkage*. Segregation refers to the separation of pairs of genetic variants in the process of reproduction. This means that for each variant, offspring receive one copy of the genetic variant from one parent and another copy of the genetic variant from the other parent. Linkage refers to the phenomenon that DNA sequences that are located closely together on a chromosome are more likely to be inherited together during the meiotic phase of reproduction. In other words, they are said to be linked (i.e., correlated). Genetic variants on different chromosomes are unlinked. In the next paragraphs we will describe the methods that use these principles of segregation and linkage to investigate the links between genetic variation and phenotypic variation in more detail, focusing on the basic logic of these methods.

Candidate Gene Association Studies

In a candidate gene association study, researchers typically test the association between a single or a few specific genetic variants located within genes and levels of a particular trait (phenotype). Briefly, such a study involves four steps: (1) selecting a putative gene based on prior knowledge about the function of this gene, (2) selecting one polymorphism (or sometimes a few) on the gene that might affect gene regulation or its protein product, (3) genotyping individuals for this polymorphism and phenotyping the trait, and followed by (4) a statistical analysis of the relationship between the polymorphism and the phenotypic score.

As an example, suppose earlier work on mice has shown that a strain of mice that has two *nonfunctional* (i.e., non-protein coding) copies of a gene A shows a particular pathological behavior that we might consider relevant for human personality trait Z. Mice that have one or two *functional* gene copies (i.e., in this case gene A codes for a particular protein) do not show this particular pathological behavior. Our

interest therefore lies in gene A that also occurs in a human form. Gene A would be selected as our candidate gene (step 1). This selection could be based on information from online catalogues about genes in different species, more specifically on information about the gene structure, expression and function, its variation in humans, and previously found links between the gene and phenotypes of relevance for the particular behavior. Some of the genetic variation across humans occurs in protein coding regions of the gene, and some of this variation occurs elsewhere. Based on several criteria, we select regions in which a nitrogenous base sequence varies across individuals (step 2: selecting a polymorphism in a gene). For instance, we find in some gene copies the base series AC and in other gene copies the base series ACACAC, so the sequence AC is repeated three times. Thus, we have a gene copy with a long series and a gene copy with a short series. This would be an example of a variable number tandem repeat (VNTR), sometimes also called micro- or minisatellites, depending on the number of repeated nucleotides that occur in the population. In the third step, we *genotype* a sample of individuals: we find that some individuals have two short copies (one inherited from their mother, the other from their father, resulting in a homozygous short genotype), other individuals have two long copies (homozygous long genotype), and the rest have one long copy and one short copy (heterozygous genotype). In the same study, we can *phenotype* the individuals: we ask them to fill in a personality questionnaire and measure their score on trait Z. In the fourth step we can correlate the number of long copies with the personality trait score (representing an additive gene model, but note that non-additive models could also be used). For more on candidate gene studies and their limitations, see Patnala, Clements, and Batra (2013).

Linkage Studies

Candidate gene studies are only likely to succeed if there are good reasons to suspect an association between variants of the gene and the trait of interest, a priori. If there are no prior ideas about which genes would be associated with a trait, one can conduct a *genome-wide linkage study* to search for regions along the genome where genes associated with the trait of interest might be located. *Genetic linkage studies* investigate how a genetic marker co-segregates with a trait (or disease) in related individuals (e.g., nuclear families or sibling pairs). Linkage studies use the principle of genetic linkage: loci close together are said to be "linked," because they tend to be inherited together (Ferreira, 2004).

Similar to candidate gene association studies, polymorphisms are selected to measure genetic variation. However, these markers are not selected based on a priori knowledge about their gene function, but they are selected because they are known to be highly polymorphic. In addition, a much larger number of markers is used, typically a few thousands, spread evenly across the genome to maximize genomic coverage. Typically, microsatellites are used in linkage studies, but other variants such as single nucleotide polymorphisms (SNPs) have been successfully used (Ball

et al., 2010). Another important difference between linkage and candidate gene studies is that in linkage studies, family data are needed, while this is not required for a candidate gene study.

Linkage studies are conducted with the following steps. In a first step, for every polymorphism or marker, the genotypes for all individuals in the sample are determined. In a second step, a linkage statistic is calculated that measures the evidence for an association between the marker and the trait of interest. Which linkage statistic is used depends on the design of the study, such as the inclusion of large family pedigrees, nuclear families, or sibling pairs. Based on family data, for each location on the genome, the identical by descent (IBD) status is estimated. IBD means that at that particular location, the genomic letters were inherited from the same ancestor. For example, imagine two sisters who at the same location each have the same two copies ACCT and ACCCT (heterozygous genotype). The copies of the two sisters seem identical. However, it might be that sister 1 inherited the ACCT from her father and the ACCCT from the mother, and sister 2 vice versa. In that case, at that location on the genome, they are identical by state (IBS) but not identical by descent (IBD). In the Haseman-Elston algorithm (Haseman & Elston, 1972), the squared difference between trait scores of the two sisters is regressed on the estimated number of IBD markers. The reason behind this is that if siblings who are IBD 2 (i.e., inherited the same two polymorphisms from the same parents) in a genomic region tend to have a very small difference in trait scores, while other sibling pairs who are IBD 1 or 0 tend to have larger differences in trait scores, there is likely a genetic variant in the same region associated with the trait. This is so because if they are IBD for the marker, they are very likely to share the same genetic variant in the same region that is responsible for the trait of interest (the causal variant). Many extensions and variations of the Haseman-Elston algorithm have been developed for complex pedigrees, including variance components modeling methods (Almasy & Blangero, 1998; Amos, 1994).

In a third step, the linkage statistics obtained for the different regions are compared, and it is determined whether at some point across the genome one or more regions show a significant deviation from what would be expected if there would be no linkage. Based on the principles of segregation and linkage, the closer the causal variant is to the marker, the larger the probability that the test statistic will pick up the signal. However, it will also very much depend on the frequencies of both the causal variant and the marker. It has turned out that genetic linkage is only a suitable technique to pick up regions of interest that harbor causal genetic variants if the effect sizes are relatively large.

Genome-Wide Association (GWA) Studies

GWA studies are usually based on unrelated individuals, just like in candidate gene association studies. An important difference, however, is that while candidate gene studies are viewed as inherently hypothesis-driven, GWA studies are considered hypothesis-free or hypothesis-generating at most. In that sense, GWA studies are

more comparable to linkage studies. Their combined aim is to map genes to traits, that is, to find genomic locations that harbor genetic variations that are responsible for observed variations in a phenotype and to do this without using a priori knowledge about the functions of specific genes or genetic variants. There are however also some notable differences between GWA studies and linkage studies. Linkage studies are conducted in families using a few thousands microsatellite markers, while GWA studies are mostly conducted in unrelated individuals measuring hundreds of thousands or even millions of SNPs.

The method of GWA is deceptively simple: (1) genotype a sample of individuals for a large number of SNPs using microarrays specially developed for this purpose, and phenotype the trait; (2) possibly also impute SNPs that are known to be polymorphic in the population of interest, but not present on the microarray, using a reference set that is suitable for that population, for example, 1000G or HAPMAP reference sets; (3) for every genotyped SNP, determine the number of risk alleles (0, 1, or 2) (or for imputed SNPs: weigh according to the probabilities of the genotypes); (4) for every SNP, perform a regression of the phenotype on the number of risk alleles (linear regression for a quantitative trait, logistic regression for a disease phenotype). Control variables are typically added to the regression model, for example, age, gender, and principal components to correct for population stratification. Next, for every SNP, it is tested whether the regression coefficient is statistically significant, using a predefined multiple-testing corrected significance level. Given that a typical GWA array contains about 500,000 SNPs, a Bonferroni corrected significance level of $5*10^{-8}$ is generally used. If the *p*-value of the association between a SNP and the trait is found below this level, this is regarded as a genome-wide significant "hit." The genomic region surrounding the SNP with this *p*-value is then studied for interesting genes that might be meaningful with regard to the trait studied.

Recent years have seen many meta-analyses of such GWA studies. In a meta-analysis, the statistical results of several studies are combined into one to identify a robust pattern of results. In the case of a GWA, for each SNP, the estimated regression coefficients for the phenotype from each individual study are pooled to form one estimated effect. It is a weighted average effect, weighted by the respective sample sizes of the individual studies. In addition, the standard error of this pooled regression coefficient is also adjusted, to reflect the increased sample size due to the pooling. This can lead to an enormous increase in statistical power, making it more likely, as compared to linkage studies, that meta-analytic GWA studies detect genetic variants of small effect size.

SNP-Based Heritability Studies

In 2011, a new method was proposed that uses genome-wide SNP data to estimate heritability (Yang, Lee, Goddard, & Visscher, 2011). The method is also known as genome-wide complex trait analysis (GCTA). Studies that apply this method are often called SNP-based heritability studies. The method briefly works as follows. It

is based on the notion that even unrelated individuals are to some degree genetically related, because they share common ancestors. The GCTA method estimates a genetic relationship matrix among all unrelated individuals in a sample, based on the available genome-wide SNP data. Typically, the genetic data are pruned to include a subset of SNPs that are largely independent (in low linkage disequilibrium (LD)). In a next step, the genetic relationship matrix is used to estimate the proportion of phenotypic variance explained by (additive) genetic variance, the so-called SNP-based heritability of a trait. This technique has more recently been extended to allow to use the information from all SNPs (without pruning) by modeling the LD structure among SNPs (Vilhjálmsson et al., 2015) and multivariate extensions using LD regression techniques to compute the SNP-based genetic correlations among phenotypes (Bulik-Sullivan et al., 2015).

In the next section, we will review what candidate gene studies, linkage studies, and GWA studies, including SNP-based heritability techniques, have yielded in terms of our knowledge about which and how many genes are related to explaining individual differences in personality traits.

Findings

Molecular genetic studies on personality have assessed personality traits with different instruments, grounded in different theories, mainly Cloninger's theory on temperament and characters (Cloninger, 1987; Cloninger, Svrakic, & Przybeck, 1993), Eysenck's personality theory (Eysenck, 1967; Eysenck & Eysenck, 1985), and the five-factor model (FFM) of personality (McCrae & Costa, 1987). The theories of Cloninger and Eysenck both propose that personality has a genetic, (neuro) biological basis but distinguish different higher-order traits. The FFM of personality is derived from factor analytic techniques and does not theorize about any biological origin.

Cloninger's theory distinguishes between temperament and character. Temperament is thought to be strongly heritable and develop early in life, while character is less heritable and rather develops in adulthood. Molecular genetic studies have therefore focused on temperament. Four dimensions of temperament have been proposed: novelty seeking, harm avoidance (sometimes labeled anxiety proneness), reward dependence, and persistence. Novelty seeking reflects the behavioral activation system (seeking out novelty and stimulation) and is hypothesized to be linked to the dopaminergic pathway. Harm avoidance reflects the behavioral inhibition system and is suggested to be linked to serotonergic brain activity. Reward dependence, encompassing characteristics like warm communication, sensitivity to social cues, sympathy, and in extreme form social dependency, is proposed to be linked to low noradrenergic neurotransmitter activity. Lastly, persistence is defined as the tendency to persevere a task, especially when challenged by frustration or fatigue. This dimension has suggestively been related to prefrontal brain activity rather than to any particular neurotransmitter system.

Eysenck's theory posits three main dimensions: neuroticism, extraversion, and psychoticism. Persons high on neuroticism are emotionally reactive in response to external stimuli, which has been linked to the limbic brain system as well as the closely connected endocrine and autonomic nervous systems. Persons high on extraversion are underaroused, therefore searching for more stimulation through (social) activities, whereas persons low on extraversion (introverts) are inherently overaroused and therefore tend to avoid and withdraw from stimulating activities and environments. Extraversion is proposed to be related to cortical brain structures. Psychoticism includes characteristics such as recklessness, non-conformity, impulsivity, hostility, and anger and has been linked to testosterone levels and the brain's monoamine oxidase pathway.

The FFM consists of five personality dimensions: neuroticism, extraversion, openness to experience, agreeableness, and conscientiousness. Neuroticism and extraversion are conceptually similar, although certainly not identical, to the traits with the same labels in Eysenck's theory. Empirical work has shown that there is substantial overlap in the personality traits from the three personality theories (Markon, Krueger, & Watson, 2005). Neuroticism and harm avoidance cluster into a negative emotionality factor. Agreeableness, conscientiousness, novelty seeking, and persistence cluster into a disinhibition factor, while reward dependence and extraversion cluster into a positive emotionality factor.

Candidate Gene Association Studies

The first three candidate gene association studies appeared in the same issue of *Nature Genetics* in 1996. They focused on a serotonin gene and on a dopamine gene. Lesch et al. tested the hypothesis that anxiety-related personality traits are correlated with genetically influenced serotonergic brain functioning. In a sample of 505 individuals, higher neuroticism was associated with one or two copies of the short allele of a VNTR (5-HTTLPR) in a promoter region near the serotonin transporter gene (*5-HTT*) on chromosome 17. In addition, 5-HTTLPR was also significantly associated with measures of anxiety and harm avoidance. It was a landmark study, cited over 5,000 times, that many others tried to replicate. These replication efforts have been meta-analyzed a couple of times (Munafo, Freimer, et al., 2009; Schinka, Busch, & Robichaux-Keene, 2004; Sen, Burmeister, & Ghosh, 2004). The meta-analysis of Sen, Burmese, and Gosh (2004) included 23 studies on 5-HTTLPR and anxiety-related personality traits in adults and found a significant association between 5-HTTLPR and neuroticism ($p = 0.000016$), but not harm avoidance ($p = 0.166$) or other anxiety-related personality traits ($p = 0.944$). A similar conclusion was reached by Schinka, Busch, and Robichaux-Keene (2004) who meta-analyzed 26 studies, showing that the effect size for the association between 5-HTTLPR and anxiety-related personality overall was small ($d = 0.10$), though medium for neuroticism ($d = 0.23$) and absent for harm avoidance ($d = 0.04$). Five years later, an even larger third meta-analysis, based on 42 studies, confirmed that

5-HTTLPR is not significantly associated with harm avoidance ($d = 0.02$, $p = 0.37$) or Eysenckian neuroticism ($d = 0.01$, $p = 0.71$) but showed consistent evidence for an association with FFM-based neuroticism ($d = 0.18$, $p < 0.001$) (Munafo, Freimer, et al., 2009). Taken together, these outcomes lend only partial support for Cloninger's hypothesis that serotonergic brain activity is involved in anxiety-related personality traits. If a certain genetic effect is only relevant for a five-factor measure of neuroticism, one might question its meaning (Munafo, Freimer, et al., 2009).

The other two candidate gene studies that appeared in the 1996 *Nature Genetics* issue reported on the dopamine receptor D4 (*DRD4*) gene and its association with novelty seeking (Benjamin et al., 1996; Ebstein et al., 1996). Based on a sample of 124 adults, Ebstein et al. (1996) showed an association between the most common long allele (7-repeat) of the *DRD4* exon III polymorphism and higher scores on novelty seeking. Benjamin et al. (1996), using a sample of 315 subjects, replicated these findings by showing that the *DRD4* exon III long alleles (6 to 8 repeats) were associated with higher scores on extraversion and with lower scores on conscientiousness compared to short alleles (2 to 5 repeats). In addition, the *DRD4* exon III polymorphism was significantly associated with novelty seeking. Since then, many studies have tried to replicate these findings regarding *DRD4* and *5-HTT*. However, as a meta-analysis by Munafo et al. (2003) showed, this was with mixed results. The meta-analysis included all studies on the association between candidate genes in the serotonin and dopamine neurotransmitter pathways (including, but not restricted to, *5-HTT* and *DRD4*). They concluded that only the association between 5HTTLPR and neuroticism remained significant.

After the disappointment of failing to replicate initially promising associations between serotonin-related and dopamine-related genes and personality traits, researchers have broadly taken two different paths in the search for genetic variants for personality. One is *gene x environment interaction* (GxE) and the other is *GWA studies*. Regarding the first strategy, some have argued that the genetic effects in the dopamine and serotonin pathways may be masked by individual differences in exposure to environments and hypothesized that we can only detect genetic effects if we consider their interaction with the environment (Belsky et al., 2009; Caspi, Hariri, Holmes, Uher, & Moffitt, 2010). However, similar problems arose in these GxE interaction studies like in the candidate gene studies: initially promising studies could not be consistently replicated by other research groups. This was seen, for example, for the interaction effect between 5-HTTLPR and stressful life events on a depression measure (Munafo, Durrant, Lewis, & Flint, 2009; Risch et al., 2009). The interaction between 5-HTTLPR and stressful life events was also not found for the depression-related personality risk trait neuroticism (Middeldorp et al., 2010).

While some focused their attention on GxE interaction, others considered GWA studies more promising. Instead of focusing on one specific gene based on existing personality theory, they were interested in finding new loci to develop new theories about personality and how it relates to its biological underpinnings (Cichon et al., 2009). This new route was based on the successes seen in the GWA studies for some disease phenotypes in medicine. Researchers thought that once such interesting new loci were identified, any interaction pattern with environmental exposure could be

studied next. Before turning to GWA studies, still during the years of trying to replicate the candidate gene studies, yet another route was taken to find genes for personality, which were linkage studies.

Linkage Studies

The first genome-wide linkage study for personality was conducted by Cloninger et al. (1998), examining harm avoidance, reward dependence, and novelty seeking. They reported a genome-wide significant location for harm avoidance on chromosome 8p. The results for reward dependence and novelty seeking were not significant. The hit for harm avoidance was replicated by Zohar et al. (2003). None of the other locations could be replicated.

Most other linkage studies for personality focused on neuroticism (Fullerton et al., 2003; Kuo et al., 2007; Nash et al., 2004; Neale, Sullivan, & Kendler, 2005). Fullerton et al. (2003) found a region on chromosome 12q that showed suggestive linkage in two other studies (Gillespie et al., 2008; Kuo et al., 2007). They also found two regions on chromosome 1. The 1p region was close to a region reported by Nash et al. (2004), and the 1q region was close to a region reported by Neale et al. (2005). However, these locations were not genome-wide significant. None of the other locations for neuroticism could be replicated across studies. Given the evidence from twin-family studies that neuroticism and harm avoidance are strongly phenotypically and genotypically correlated (Gillespie, Johnstone, Boyce, Heath, & Martin, 2001), it is remarkable that the linkage studies fail to consistently point to specific locations on the genome for these phenotypes.

In sum, genome-wide linkage studies have not been very successful in identifying genomic regions that are replicated. With sample sizes of around 200–1300 sibling pairs, these studies might not have had enough statistical power. Furthermore, with linkage studies it is quite difficult to pinpoint regions with enough precision (Roberts, MacLean, Neale, Eaves, & Kendler, 1999). Nowadays, genome-wide linkage studies are considered to be a suboptimal technique to identify genes for complex traits (i.e., phenotypes that are influenced by many common genetic variants with each small effect size) (McCarthy et al., 2008; Stranger, Stahl, & Raj, 2011), and researchers have turned to GWA studies instead.

Genome-Wide Association Studies

In 2008, the first two GWA studies on personality were reported, both examining neuroticism. Shifman et al. (2008), using a sample size of 2000, found no genome-wide significant locations but did report a suggestive association between neuroticism and the *PDE4D* gene (chromosome 5). The other 2008 study used a sample size of 1227 and found a suggestive association between neuroticism and the

MAMDC1 gene (also labeled *MDGA2*) (van den Oord et al., 2008). This location was also significant in a German replication sample of size 1880, as reported in the same study. Interestingly, the same gene associated with neuroticism was found to be associated with harm avoidance in a study comparing the genotypes of 199 patients with depression and 541 healthy controls (Heck et al., 2011). Unfortunately, the locus was not associated with neuroticism in another study with sufficient power (Hettema, van den Oord, An, Kendler, & Chen, 2009).

After these initial studies on neuroticism, the first study on all five personality traits from the FFM was reported by Terracciano et al. (2010). It was based on 3972 individuals from a genetically isolated sample in Sardinia, Italy. Unfortunately, no genome-wide hits were found. There were a few suggestive associations for brain-related genes, for example, between neuroticism and the *SNAP25* gene, between extraversion and the *BDNF* and two cadherin genes, and between agreeableness and the *CLOCK* gene (involved in circadian rhythms). The *CLOCK* gene's association with agreeableness was present in two out of three replication samples, but none of the other associations could be replicated. The first study on the four Cloninger temperament dimensions was published in the same year by Verweij et al. (2010), using a sample size of 5117 individuals. Similarly, no genome-wide significant locations were reported, though the study had a statistical power of 90% to find a locus that explains 1% of the phenotypic variation. Both these studies show that personality traits are probably highly polygenic in nature, with many variants of each very small effect size. It therefore became clear that if one were to find genetic variants for personality, much larger sample sizes would be required. One way to achieve at such large sample sizes would be to pool results through meta-analysis.

Meta-Analytic Genome-Wide Association Studies

Using the strategy of having multiple research groups carrying out a GWA study on the same five FFM personality traits, and then combining the statistical results of ten samples in a meta-analysis, de Moor et al. (2012) carried out an analysis with a total sample size of 17,375 subjects. There were an extra five independent replication samples with a total sample size of 3294 subjects. These 15 cohorts were the start of the Genetics of Personality Consortium (GPC). In this first study, significant loci were found for openness to experience near the *RASA1* gene on chromosome 5 and for conscientiousness in the brain-expressed *KATNAL2* gene on chromosome 18. However, these associations were very small (explained variances below 0.25%) and not clearly replicated in the five replication samples. Contrary to expectation, no genome-wide significant loci were found for neuroticism, the personality trait that had been so focused on in earlier molecular genetic studies.

In an attempt to further increase sample size, the GPC sought to include other research groups that in addition to genetic data had personality data, not just FFM traits but traits from other personality theories that are known to be highly correlated. Using phenotypic harmonization, data from various personality questionnaires were transformed in such a way that scores were representative of

either neuroticism/harm avoidance or extraversion/reward dependence. This procedure was based on *test linking* and *test score equating*, methods applied in educational measurement settings that in turn makes use of models from item response theory (van den Berg et al., 2014). This harmonization attempt resulted in a sample size of well over 63,000 subjects from 29 cohorts (de Moor et al., 2015). For neuroticism/harm avoidance, a genome-wide significant result was found in the *MAGI1* gene (chromosome 3). Unfortunately, this was not significant in the replication cohort consisting of 10,000 subjects, though it remained significant in a meta-analysis including all 30 cohorts (de Moor et al., 2015). The *MAGI1* is an interesting gene, as it is expressed in the hippocampus and had previously been linked to major depression, schizophrenia, and bipolar disorder. These are disorders that we know are genetically correlated with neuroticism. For extraversion/reward dependence, the harmonization attempt with the resulting increase in sample size did not lead to any genome-wide significant hits, which could be partly explained by the larger phenotypic differences in the scales used to assess extraversion/reward dependence, which made it somewhat more difficult to harmonize the traits as compared to neuroticism/harm avoidance. Another explanation could be that extraversion is even more genetically complex (i.e., more polygenic, more rare variants, non-additive gene action, etc.) compared with neuroticism.

These first meta-analyses were followed by a larger meta-analysis for neuroticism with a sample size of 106,716 subjects (Smith et al., 2016). The large sample size was mainly due to the inclusion of data from the UK Biobank project with a sample size of 91,370. The UK Biobank contains health, medical, and genetic data for over 500,000 individuals between the ages of 39 and 73 years from the UK, assessed between 2006 and 2010. Smith et al. (2016) reported nine new loci for neuroticism. Two of these loci were in very large LD blocks, which make it very hard to determine the exact locus of interest. Among the genes that could be of potential interest in one of these regions is *GRIK3* (a glutamate receptor gene). The gene identified by de Moor et al. (2015), the *MAGI1*, showed the same direction of effect, but was not significant in this larger study.

When the UK Biobank data were combined with data from the GPC, 11 significant loci for neuroticism were reported (Okbay et al., 2016). Five of these had not been found in either the GPC study or the UK Biobank study, alone or any other previous GWA study. Of interest is the significant finding close to the *DRD2*, a dopamine receptor gene on 11q. This gene had been identified earlier in candidate gene studies (Munafo et al., 2003) but had been dismissed on the basis of heterogeneity in findings across samples. Lo et al. (2017) combined the data from the GPC with data from 23andMe and conducted a meta-analysis for agreeableness, openness to experience, conscientiousness, neuroticism, and extraversion. Two loci for neuroticism were found, one of which in the same large LD block on chromosome 8p23.1 as already identified by Smith et al. (2016) and Okbay et al. (2016). Subsequently, Lo et al. (2017) combined the data from the GPC with data from 23andMe, an independent replication sample from 23andMe, deCODE, and the UK Biobank. This resulted in a total of 260,861 phenotyped and genotyped individuals. Six genome-wide significant loci were found, one for conscientiousness, four for extraversion, and two variants for neuroticism, each variant explaining between 1

and 4% of the phenotypic variance. One of the variants for neuroticism was located in the same block on 8p23.1 that was previously associated with neuroticism and showed the largest effect size with 3.95% explained phenotypic variance. To date, it is the largest reported effect size for a personality trait for a single SNP. Most likely the effect size, if it is a true effect, is highly overestimated, since we already saw that a sample size of 5117 is sufficient to have a probability of 90% to detect a true effect that explains 1% of the variance (Verweij et al., 2010). If the true effect size were really this large, it should have been detected by earlier meta-analytic GWA studies like the GPC or even single-sample GWA studies.

The latest meta-analysis using the UK Biobank, 23andMe, and the Genetics of Personality Consortium data sets had a sample size of 449,484 subjects and studied neuroticism only (Nagel et al., 2018). There were 136 loci that were genome-wide significant; the effect sizes were not clearly reported. These loci could be linked to many genes involved in specific brain cell types, including dopaminergic neuroblasts, medium spiny neurons, and serotonergic neurons. Gene set analyses seemed to implicate pathways involved in neurogenesis, behavioral response to cocaine processes, and axon part. Nagel et al. (2018) also showed first molecular genetic evidence of some genetic heterogeneity in neuroticism: two genetically distinguishable subclusters of neuroticism labeled "depressed affect" and "worry." A meta-analysis focusing on facets of neuroticism rather than the full scale was reported by Kim et al. (2017). Notwithstanding the relatively small sample size, 5584, they found a significant locus for the angry-hostility facet within the *MIR548H3* gene. A non-meta-GWA analysis using only UK Biobank data, with a sample size of 329,821 individuals, was reported by Luciano et al. (2018): 116 independent loci were associated with neuroticism, of which only 15 replicated in 23andMe and the Genetics of Personality Consortium.

The general pattern that emerges is that as sample size increases, the number of significant hits increases as well. This seems promising. However, an important problem is that replication is no longer possible: the increase in sample size is in large part due to the pooling of independent samples, and the overlap in samples across published studies is large. Only the UK Biobank data set is large enough to detect a large number of hits, but there is (yet) no possibility to check these results in equally sized independent samples. Another problem involves harmonization of personality phenotypes. Only the GPC used harmonized measures, so that the effect sizes of SNPs could be compared across cohorts. For all other subsequent meta-analyses, this was not the case: UK Biobank data is based on EPQ short-form measures, 23andME uses the Big Five Inventory, and the deCode data are based on NEO-FFI measures. For instance, we see the detection of the 18p23.1 locus for neuroticism in a couple of studies and although the direction of effect is in all cases the same, the effect sizes cannot be meaningfully compared since the studies all use different scales for the neuroticism phenotype. As it is more meaningful to compare effect sizes across studies rather than *p*-values (Amrhein, Greenland, & McShane, 2019), this is a major hurdle in the interpretation of results across studies.

Setting aside these problems of replication and effect sizes, the loci that so far appear to have been found for personality traits explain only a very small percentage of the variance, indicating that a large proportion of the genetic variance as estimated in twin and adoption studies yet remains to be explained. What could SNP-based heritability studies tell us about this unexplained genetic variance, also known as the "the missing heritability problem" (Maher, 2008)?

SNP-Based Heritability Studies

SNP-based heritability studies estimate the genetic variance in a trait explained not by a single SNP but by all SNPs in a study taken together. This type of study can answer the question whether it is realistic to find loci related to personality using SNPs as in a GWA study. The reasoning is that if the variance accounted for by SNPs is much less than the genetic variance found in twin and adoption studies, much of the variance is hidden and will not be found in GWA studies with the current set of SNPs (or, alternatively, the estimates from twin and adoption studies could be biased). If estimates from twin and adoption studies are not biased, the variance may be hidden in regions not covered by SNPs, low frequency alleles, structural variations, or in complex interactions among genes, or gene-environment interactions. Most GWA studies now report SNP-heritability results; estimates range from 0% to 18% (de Moor et al., 2015; Lo et al., 2017; van den Berg et al., 2016; Verweij et al., 2012; Vinkhuyzen et al., 2012). The sample size used by Lo et al. (2017) was the largest (60,000), and they reported SNP-heritabilities between 9% (agreeableness) and 18% (extraversion). The SNP-based heritabilities were significant for all five personality traits, which prove that common genetic variants do influence personality, though the individual genetic variants are difficult to detect, most likely because each variant accounts for only a small proportion of variance (small effect size and/or low allele frequency). Nevertheless, the reported percentages of variance explained by SNPs are clearly less than the heritability estimates from twin and adoption studies, suggesting that common genetic variants with additive gene action do not fully explain all genetic variation in personality scores.

Taken together the evidence from the studies discussed above, it seems indeed possible to trace back the genetic variance from twin/adoption studies in the actual DNA, which is an important outcome from the recent meta-analytic whole-genome studies on personality, in spite of the difficulties still encountered to robustly detect single genetic variants. Yet, still an important portion of the genetic variance remains to be explained. To gain further insight into the relevance of the polygenetic variance of personality traits, GWA studies for personality have additionally started to investigate the polygenetic overlap among personality traits and with related psychiatric phenotypes.

Polygenic Associations Between Personality Traits and Psychiatric Phenotypes

Although some theories of personality, such as the FFM, posited that the main dimensions of personality are orthogonal, it is well-known that most personality traits are correlated. For example, the phenotypic correlations between the Big Five personality traits in the large 23andMe sample ranged between |0.01| to |0.34| and were significant for all pairs of traits except for the correlation between conscientiousness and openness to experience (Lo et al., 2017). The correlation between neuroticism and extraversion was −0.24 and highly similar to previous reports (e.g. Gillespie et al., 2001). Multivariate twin-family studies have shown that personality traits are also *genetically* correlated (Franic, Borsboom, Dolan, & Boomsma, 2014; Jang et al., 2006). Using genome-wide SNP data to estimate genetic correlations among personality traits yields a similar picture: genetic correlations among Big Five personality traits range between |0.11| and |0.40| and were significant for all pairs of traits with the exception of the correlation between openness to experience and agreeableness. This analysis indicates that the personality traits are not just correlated at the phenotypic level but their genetic influences are associated as well (Lo et al., 2017), which could indicate that the same genes impact on multiple personality traits (Carey, 1988).

Personality traits are viewed as important predictors for a variety of life outcomes, including mental and physical health outcomes. Molecular genetic studies on personality in this regard are often conducted in the hope to not only increase our understanding of the underlying (neuro)biological etiology of personality traits themselves but also of the risk to develop psychiatric disorders. Middeldorp et al. (2011) showed that polygenic risk scores, based on meta-analytic GWA results for the five FFM personality traits (de Moor et al., 2012), uniquely predicted diagnoses of major depressive disorder (MDD) and bipolar disorder (BD) in independent samples. Polygenic risk scores for neuroticism predicted MDD, while polygenic risk scores for extraversion predicted BD. The prediction of MDD from polygenic scores for neuroticism was confirmed in the larger study using harmonized neuroticism as a phenotype by De Moor et al. (2015). Okbay et al. (2016) estimated the genetic correlations of neuroticism with depressive symptoms, subjective well-being, and a range of neuropsychiatric and physical health phenotypes, by investigating the overlap of the GWA results of all these phenotypes. Neuroticism was strongly genetically correlated to both depressive symptoms ($r_g = 0.70$) and subjective well-being ($r_g = -0.75$). The genetic correlation of neuroticism with anxiety disorders was also large. Smaller, yet significant, genetic correlations were found for neuroticism with BD, schizophrenia, coronary heart disease, and smoking. Of note, neuroticism was not genetically correlated with Alzheimer disease or autism spectrum disorder (ASD), body-mass index, fasting glucose, or triglycerides (Okbay et al., 2016). Lo et al. (2017) computed the genetic correlations among the Big Five personality traits and six psychiatric disorders: MDD, BD, schizophrenia, attention-deficit hyperactivity disorder (ADHD), ASD, and anorexia nervosa. Neuroticism was highly genetically

correlated with MDD (r_g = 0.56); extraversion showed a moderate genetic correlation with ADHD (r_g = 0.30). Openness to experience was moderately correlated with schizophrenia, BD, and MDD (r_g from 0.28 to 0.36). MDD was further not only genetically correlated with neuroticism and openness to experience but also with agreeableness and conscientiousness (moderate negative correlations). Principal component analysis of the genetic correlation matrix further showed that the five personality traits and six psychiatric phenotypes can be grouped along two genetic dimensions: the first dimension mainly separating personality traits from psychiatric phenotypes and the second dimension reflecting an externalizing-internalizing distinction (Lo et al., 2017). Taken together, these studies show clear overlap in the SNP-based genetic variance that associates with personality traits on the one hand and psychiatric phenotypes on the other hand. Moreover, neuroticism is most prominently genetically linked to internalizing disorders such as MDD and anxiety disorders, while extraversion is rather linked to ADHD and BD. In a previous section of this chapter, it was shown that GWA studies have just begun to identify some specific loci for personality. An important next step would be to further replicate these loci in terms of effect sizes using the same personality scales, to characterize the function of such loci in human (brain) tissues, and to relate this functional knowledge to insights into the (neuro)biological etiology of psychiatric phenotypes.

Conclusion

The landscape of molecular genetic research has changed dramatically over the last three decades, as we have reviewed in this chapter, and this also holds more specifically for the molecular genetic work on personality. Only over two decades ago, the first candidate gene studies for personality were published in which single genetic variants in a gene were associated with neuroticism and novelty seeking in just a few hundred participants (Benjamin et al., 1996; Ebstein et al., 1996; Lesch et al., 1996). This was followed by family-based genome-wide linkage studies conducted in at most a few thousand subjects and relying on a few thousands of microsatellite markers to cover the genome. Both types of studies are now considered to have been largely underpowered. Furthermore, these studies did not cover or failed to cover most of what we now know constitutes the common variation across the entire human genome. Hence, these studies did not yield any notably consistent results.

Now contrast this with the latest molecular genetic study on personality, in which more than eight million different genetic variants across the entire human genome were investigated in almost half a million of subjects (Nagel et al., 2018). This study has detected 136 genome-wide significant genetic variants for neuroticism. In addition, such recent large-scale GWA studies have demonstrated that part of the heritability of personality as established in twin, family, and adoption studies can indeed be traced back to actual variation in the DNA. Furthermore, it has become clear that the variation in the DNA that is associated with personality

traits is also linked to psychiatric phenotypes and even health-related phenotypes (Lo et al., 2017; Okbay et al., 2016).

All in all, molecular genetic studies for personality have shown that likely a very large number of genes are implicated in explaining human variation in personality, in a manner that had already been proposed by Ronald Fisher one century ago with his polygenic model of inheritance (Fisher, 1918). The large sets of genes for personality have a variety of functions in the brain and other parts of the body and partly overlap with other phenotypes. Yet, the exact biological mechanisms of this polygenic variation and the causal mechanisms among genes, personality, and related phenotypes largely remain to be understood.

It is expected that novel genotyping techniques as well as novel methodological approaches will further boost the field of molecular genetic research on personality. For instance, new methods could divert from a focus on detecting single variants using a too simplified focus on statistical significance and lead us to answer perhaps more interesting and relevant questions such as what the most important genetic predictors of personality are (by using variable selection methods) or how we can predict personality based on genetic marker data (by applying machine learning methods). Furthermore, it should be noted that some other promising technical and analytical innovations are also already underway. For example, next-generation sequencing, genome-wide gene expression and epigenetic studies, and omics studies have now become feasible, and big data analytical techniques can be employed to combine and make sense of such data (Auffray et al., 2016; Eisenstein, 2015). We expect that these techniques will find their application to personality.

Besides such genomic big data efforts, it will remain important to continue to evaluate the usefulness of smaller-sized single studies to further increase our insights into genetic versus environmental influences on personality. For instance, twin-family and adoption studies have shown that environmental factors, largely unshared among family members in adulthood, contribute to about an equal part of the variation in personality traits, yet we know little about which specific environmental factors account for that environmental variance and how they link to polygenic variation in our DNA. In addition, molecular genetic work on temperament and personality has rarely been integrated within single studies but would be important if we aim to obtain a lifespan understanding of how genetic variation affects personality development, while considering environmental influences as well. Carefully designed longitudinal studies that include both molecular genetic data as well as in-depth assessment of environmental factors across development would be needed to address these outstanding issues.

Hopefully, these lines of research (molecular genetic/omics approaches, versus environmental/longitudinal approaches) are not just further developed independently but brought together in multidisciplinary collaborations. A major challenge will lie in the interpretation of the outcomes of such multidisciplinary collaborative projects in the context of prevailing personality theories. Ultimately, we hope that the multitude of proposed future directions will yield novel knowledge on how genes are implicated in the development of individual differences in personality and how this relates to important health and well-being outcomes.

References

Almasy, L., & Blangero, J. (1998). Multipoint quantitative-trait linkage analysis in general pedigrees. *American Journal of Human Genetics, 62*(5), 1198–1211. https://doi.org/10.1086/301844

Amos, C. I. (1994). Robust variance-components approach for assessing genetic linkage in pedigrees. *American Journal of Human Genetics, 54*(3), 535–543.

Amrhein, V., Greenland, S., & McShane, B. (2019). Scientists rise up against statistical significance. *Nature, 567*(7748), 305–307. https://doi.org/10.1038/d41586-019-00857-9

Auffray, C., Balling, R., Barroso, I., Bencze, L., Benson, M., Bergeron, J., … Zanetti, G. (2016). Making sense of big data in health research: Towards an EU action plan. *Genome Medicine, 8*. https://doi.org/10.1186/s13073-016-0323-y

Ball, A. D., Stapley, J., Dawson, D. A., Birkhead, T. R., Burke, T., & Slate, J. (2010). A comparison of SNPs and microsatellites as linkage mapping markers: Lessons from the zebra finch (Taeniopygia guttata). *BMC Genomics, 11*. https://doi.org/10.1186/1471-2164-11-218

Belsky, J., Jonassaint, C., Pluess, M., Stanton, M., Brummett, B., & Williams, R. (2009). Vulnerability genes or plasticity genes? *Molecular Psychiatry, 14*(8), 746–754. https://doi.org/10.1038/mp.2009.44

Benjamin, J., Li, L., Patterson, C., Greenberg, B. D., Murphy, D. L., & Hamer, D. H. (1996). Population and familial association between the D4 dopamine receptor gene and measures of novelty seeking. *Nature Genetics, 12*(1), 81–84. https://doi.org/10.1038/ng0196-81

Bulik-Sullivan, B. K., Loh, P. R., Finucane, H. K., Ripke, S., Yang, J., Patterson, N., … Consortium SWG of the PG. (2015). LD score regression distinguishes confounding from polygenicity in genome-wide association studies. *Nature Genetics, 47*(3), 291–295. https://doi.org/10.1038/ng.3211

Carey, G. (1988). Inference about genetic correlations. *Behavior Genetics, 18*(3), 329–338.

Caspi, A., Hariri, A. R., Holmes, A., Uher, R., & Moffitt, T. E. (2010). Genetic sensitivity to the environment: The case of the serotonin transporter gene and its implications for studying complex diseases and traits. *American Journal of Psychiatry, 167*(5), 509–527. https://doi.org/10.1176/appi.ajp.2010.09101452

Cichon, S., Craddock, N., Daly, M., Faraone, S. V., Gejman, P. V., Kelsoe, J., … Psychiatric, G. C. S. (2009). A framework for interpreting genome-wide association studies of psychiatric disorders The Psychiatric GWAS Consortium Steering Committee. *Molecular Psychiatry, 14*(1), 10–17. https://doi.org/10.1038/mp.2008.126

Cloninger, C. R. (1987). A systematic method for clinical description and classification of personality variants – a proposal. *Archives of General Psychiatry, 44*(6), 573–588.

Cloninger, C. R., Svrakic, D. M., & Przybeck, T. R. (1993). A psychobiological model of temperament and character. *Archives of General Psychiatry, 50*(12), 975–990.

Cloninger, C. R., Van Eerdewegh, P., Goate, A., Edenberg, H. J., Blangero, J., Hesselbrock, V., … Begleiter, H. (1998). Anxiety proneness linked to epistatic loci in genome scan of human personality traits. *American Journal of Medical Genetics, 81*(4), 313–317. https://doi.org/10.1002/(sici)1096-8628(19980710)81:4<313::aid-ajmg7>3.0.co;2-u

de Moor, M. H. M., Costa, P. T., Terracciano, A., Krueger, R. F., de Geus, E. J. C., Toshiko, T., … Boomsma, D. I. (2012). Meta-analysis of genome-wide association studies for personality. *Molecular Psychiatry, 17*(3), 337–349. https://doi.org/10.1038/mp.2010.128

de Moor, M. H. M., van den Berg, S. M., Verweij, K. J. H., Krueger, R. F., Luciano, M., Vasquez, A. A., … Genetics Personality, C. (2015). Meta-analysis of genome-wide association studies for neuroticism, and the polygenic association with major depressive disorder. *JAMA Psychiatry, 72*(7), 642–650. https://doi.org/10.1001/jamapsychiatry.2015.0554

Ebstein, R. P., Novick, O., Umansky, R., Priel, B., Osher, Y., Blaine, D., … Belmaker, R. H. (1996). Dopamine D4 receptor (D4DR) exon III polymorphism associated with the human personality trait of novelty seeking. *Nature Genetics, 12*(1), 78–80. https://doi.org/10.1038/ng0196-78

Eisenstein, M. (2015). The power of petabytes. *Nature, 527*(7576), S2–S4.
Eysenck, H. J. (1967). *The biological basis of personality*. Springfield, MO: Thomas.
Eysenck, H. J., & Eysenck, M. W. (1985). *Personality and individual differences*. New York, NY: Plenum.
Ferreira, M. A. R. (2004). Linkage analysis: Principles and methods for the analysis of human quantitative traits. *Twin Research, 7*(5), 513–530. https://doi.org/10.1375/13690520423 35223
Fisher, R. A. (1918). The correlation between relatives on the supposition of Mendelian inheritance. *Transactions of the Royal Society of Edinburgh, 52*(2), 34.
Franic, S., Borsboom, D., Dolan, C. V., & Boomsma, D. I. (2014). The big five Personality traits: Psychological entities or statistical constructs? *Behavior Genetics, 44*(6), 591–604. https://doi.org/10.1007/s10519-013-9625-7
Fullerton, J., Cubin, M., Tiwari, H., Wang, C., Bomhra, A., Davidson, S., … Flint, J. (2003). Linkage analysis of extremely discordant and concordant sibling pairs identifies quantitative-trait loci that influence variation in the human personality trait neuroticism. *American Journal of Human Genetics, 72*(4), 879–890. https://doi.org/10.1086/374178
Gillespie, N. A., Johnstone, S. J., Boyce, P., Heath, A. C., & Martin, N. G. (2001). The genetic and environmental relationship between the interpersonal sensitivity measure (IPSM) and the personality dimensions of Eysenck and Cloninger. *Personality and Individual Differences, 31*(7), 1039–1051. https://doi.org/10.1016/s0191-8869(00)00200-2
Gillespie, N. A., Zhu, G., Evans, D. M., Medland, S. E., Wright, M. J., & Martin, N. G. (2008). A genome-wide scan for Eysenckian Personality dimensions in adolescent twin Sibships: Psychoticism, extraversion, neuroticism, and lie. *Journal of Personality, 76*(6), 1415–1445. https://doi.org/10.1111/j.1467-6494.2008.00527.x
Gusella, J. F., Wexler, N. S., Conneally, P. M., Naylor, S. L., Anderson, M. A., Tanzi, R. E., … Martin, J. B. (1983). A polymorphic DNA marker genetically linked to Huntingtons-disease. *Nature, 306*(5940), 234–238. https://doi.org/10.1038/306234a0
Haseman, J. K., & Elston, R. C. (1972). Investigation of linkage between a quantitative trait and a marker locus. *Behavior Genetics, 2*(1), 3–19. https://doi.org/10.1007/bf01066731
Heck, A., Pfister, H., Czamara, D., Muller-Myhsok, B., Putz, B., Lucae, S., … Ising, M. (2011). Evidence for associations between MDGA2 polymorphisms and harm avoidance – replication and extension of a genome-wide association finding. *Psychiatric Genetics, 21*(5), 257–260. https://doi.org/10.1097/YPG.0b013e3283457bfb
Hettema, J. M., van den Oord, E., An, S. S., Kendler, K. S., & Chen, X. N. (2009). Follow-up association study of novel neuroticism gene MAMDC1. *Psychiatric Genetics, 19*(4), 213–214. https://doi.org/10.1097/YPG.0b013e32832cec22
International Human Genome Sequencing Consortium. (2004). Finishing the euchromatic sequence of the human genome. *Nature, 431*(7011), 931–945. https://doi.org/10.1038/nature03001
Jang, K. L., Livesley, W. J., Ando, J., Yamagata, S., Suzuki, A., Angleitner, A., … Spinath, F. (2006). Behavioral genetics of the higher-order factors of the Big Five. *Personality and Individual Differences, 41*(2), 261–272. https://doi.org/10.1016/j.paid.2005.11.033
Johanssen, W. (1909). *Elemente der exakten Erblichkeitslehre*. Jena, Germany: Gustav Fischer.
Kim, S. E., Kim, H. N., Yun, Y. J., Heo, S. G., Cho, J., Kwon, M. J., … Kim, H. L. (2017). Meta-analysis of genome-wide SNP- and pathway-based associations for facets of neuroticism. *Journal of Human Genetics, 62*(10), 903–909. https://doi.org/10.1038/jhg.2017.61
Kuo, P. H., Neale, M. C., Riley, B. P., Patterson, D. G., Walsh, D., Prescott, C. A., & Kendler, K. S. (2007). A genome-wide linkage analysis for the personality trait neuroticism in the Irish affected sib-pair study of alcohol dependence. *American Journal of Medical Genetics Part B-Neuropsychiatric Genetics, 144B*(4), 463–468. https://doi.org/10.1002/ajmg.b.30478
Lander, E. S., Linton, L. M., Birren, B., Nusbaum, C., Zody, M. C., Baldwin, J., … International Human Genome Sequencing Consortium. (2001). Initial sequencing and analysis of the human genome. *Nature, 409*(6822), 860–921. https://doi.org/10.1038/35057062

Lesch, K. P., Bengel, D., Heils, A., Sabol, S. Z., Greenberg, B. D., Petri, S., … Murphy, D. L. (1996). Association of anxiety-related traits with a polymorphism in the serotonin transporter gene regulatory region. *Science, 274*(5292), 1527–1531. https://doi.org/10.1126/science.274.5292.1527

Lo, M. T., Hinds, D. A., Tung, J. Y., Franz, C., Fan, C. C., Wang, Y. P., … Chen, C. H. (2017). Genome-wide analyses for personality traits identify six genomic loci and show correlations with psychiatric disorders. *Nature Genetics, 49*(1), 152–156. https://doi.org/10.1038/ng.3736

Luciano, M., Hagenaars, S. P., Davies, G., Hill, W. D., Clarke, T. K., Shirali, M., … Deary, I. J. (2018). Association analysis in over 329,000 individuals identifies 116 independent variants influencing neuroticism. *Nature Genetics, 50*(1), 6–11. https://doi.org/10.1038/s41588-017-0013-8

Maher, B. (2008). Personal genomes: The case of the missing heritability. *Nature, 456*(7218), 18–21. https://doi.org/10.1038/456018a

Markon, K. E., Krueger, R. F., & Watson, D. (2005). Delineating the structure of normal and abnormal personality: An integrative hierarchical approach. *Journal of Personality and Social Psychology, 88*(1), 139–157. https://doi.org/10.1037/0022-3514.88.1.139

McCarthy, M. I., Abecasis, G. R., Cardon, L. R., Goldstein, D. B., Little, J., Ioannidis, J. P. A., & Hirschhorn, J. N. (2008). Genome-wide association studies for complex traits: Consensus, uncertainty and challenges. *Nature Reviews Genetics, 9*(5), 356–369. https://doi.org/10.1038/nrg2344

McCrae, R. R., & Costa, P. T. (1987). Validation of the 5-factor model of personality across instruments and observers. *Journal of Personality and Social Psychology, 52*(1), 81–90. https://doi.org/10.1037/0022-3514.52.1.81

Mendel, G. J. (1866). Versuche ueber pflanzenhybriden. *Verhandlungen des Naturforschunden Vereines in Bruenn, 4*, 3.

Merenda, P. F. (1987). Toward a 4-factor theory of temperament and or personality. *Journal of Personality Assessment, 51*(3), 367–374. https://doi.org/10.1207/s15327752jpa5103_4

Middeldorp, C. M., de Geus, E. J. C., Willemsen, G., Hottenga, J. J., Slagboom, P. E., & Boomsma, D. I. (2010). The serotonin transporter gene length polymorphism (5-HTTLPR) and life events: No evidence for an interaction effect on neuroticism and anxious depressive symptoms. *Twin Research and Human Genetics, 13*(6), 544–549. https://doi.org/10.1375/twin.13.6.544

Middeldorp, C. M., de Moor, M. H. M., McGrath, L. M., Gordon, S. D., Blackwood, D. H., Costa, P. T., … Boomsma, D. I. (2011). The genetic association between personality and major depression or bipolar disorder. A polygenic score analysis using genome-wide association data. *Translational Psychiatry, 1*. https://doi.org/10.1038/tp.2011.45

Munafo, M. R., Clark, T. G., Moore, L. R., Payne, E., Walton, R., & Flint, J. (2003). Genetic polymorphisms and personality in healthy adults: A systematic review and meta-analysis. *Molecular Psychiatry, 8*(5), 471–484. https://doi.org/10.1038/sj.mp.4001326

Munafo, M. R., Durrant, C., Lewis, G., & Flint, J. (2009). Gene x environment interactions at the serotonin transporter locus. *Biological Psychiatry, 65*(3), 211–219. https://doi.org/10.1016/j.biopsych.2008.06.009

Munafo, M. R., Freimer, N. B., Ng, W., Ophoff, R., Veijola, J., Miettunen, J., … Flint, J. (2009). 5-HTTLPR genotype and anxiety-related personality traits: A meta-analysis and new data. *American Journal of Medical Genetics Part B-Neuropsychiatric Genetics, 150B*(2), 271–281. https://doi.org/10.1002/ajmg.b.30808

Nagel, M., Jansen, P. R., Stringer, S., Watanabe, K., de Leeuw, C. A., Bryois, J., … Team, A. R. (2018). Meta-analysis of genome-wide association studies for neuroticism in 449, 484 individuals identifies novel genetic loci and pathways. *Nature Genetics, 50*(7), 920–927. https://doi.org/10.1038/s41588-018-0151-7

Nash, M. W., Huezo-Diaz, P., Sterne, A., Purcell, S., Hoda, F., Cherny, S. S., … Sham, P. C. (2004). Genome-wide linkage analysis of a composite index of neuroticism and mood-related scales in extreme selected sibships. *Human Molecular Genetics, 13*(19), 2173–2182. https://doi.org/10.1093/hmg/ddh239

Neale, B. M., Sullivan, P. F., & Kendler, K. S. (2005). A genome scan of neuroticism in nicotine dependent smokers. *American Journal of Medical Genetics Part B-Neuropsychiatric Genetics, 132B*(1), 65–69. https://doi.org/10.1002/ajmg.b.30095

Okbay, A., Baselmans, B. M., De Neve, J. E., Turley, P., Nivard, M. G., Fontana, M. A., … Study, L. C. (2016). Genetic variants associated with subjective well-being, depressive symptoms, and neuroticism identified through genome-wide analyses. *Nature Genetics, 48*(6), 624–633. https://doi.org/10.1038/ng.3552

Patnala, R., Clements, J., & Batra, J. (2013). Candidate gene association studies: A comprehensive guide to useful in silico tools. *BMC Genetics, 14*. https://doi.org/10.1186/1471-2156-14-39

Risch, N., Herrell, R., Lehner, T., Liang, K. Y., Eaves, L., Hoh, J., … Merikangas, K. R. (2009). Interaction between the serotonin transporter gene (5-HTTLPR), stressful life events, and risk of depression a meta-analysis. *Jama-Journal of the American Medical Association, 301*(23), 2462–2471.

Roberts, S. B., MacLean, C. J., Neale, M. C., Eaves, L. J., & Kendler, K. S. (1999). Replication of linkage studies of complex traits: An examination of variation in location estimates. *American Journal of Human Genetics, 65*(3), 876–884. https://doi.org/10.1086/302528

Schinka, J. A., Busch, R. M., & Robichaux-Keene, N. (2004). A meta-analysis of the association between the serotonin transporter gene polymorphism (5-HTTLPR) and trait anxiety. *Molecular Psychiatry, 9*(2), 197–202. https://doi.org/10.1038/sj.mp.4001405

Sen, S., Burmeister, M., & Ghosh, D. (2004). Meta-analysis of the association between a serotonin transporter promoter polymorphism (5-HTTLPR) and anxiety-related personality traits. *American Journal of Medical Genetics Part B-Neuropsychiatric Genetics, 127B*(1), 85–89. https://doi.org/10.1002/ajmg.b.20158

Shifman, S., Bhomra, A., Smiley, S., Wray, N. R., James, M. R., Martin, N. G., … Flint, J. (2008). A whole genome association study of neuroticism using DNA pooling. *Molecular Psychiatry, 13*(3), 302–312. https://doi.org/10.1038/sj.mp.4002048

Smith, D. J., Escott-Price, V., Davies, G., Bailey, M. E. S., Colodro-Conde, L., Ward, J., … O'Donovan, M. C. (2016). Genome-wide analysis of over 106,000 individuals identifies 9 neuroticism-associated loci. *Molecular Psychiatry, 21*(6), 749–757. https://doi.org/10.1038/mp.2016.49

Stranger, B. E., Stahl, E. A., & Raj, T. (2011). Progress and promise of genome-wide association studies for human complex trait Genetics. *Genetics, 187*(2), 367–383. https://doi.org/10.1534/genetics.110.120907

Terracciano, A., Sanna, S., Uda, M., Deiana, B., Usala, G., Busonero, F., … Costa, P. T. (2010). Genome-wide association scan for five major dimensions of personality. *Molecular Psychiatry, 15*(6), 647–656. https://doi.org/10.1038/mp.2008.113

van den Oord, E., Kuo, P. H., Hartmann, A. M., Webb, B. T., Moller, H. J., Hettema, J. M., … Rujescu, D. (2008). Genomewide association analysis followed by a replication study implicates a novel candidate gene for neuroticism. *Archives of General Psychiatry, 65*(9), 1062–1071. https://doi.org/10.1001/archpsyc.65.9.1062

van den Berg, S. M., de Moor, M. H. M., McGue, M., Pettersson, E., Terracciano, A., Verweij, K. J. H., … Boomsma, D. I. (2014). Harmonization of neuroticism and extraversion phenotypes across inventories and cohorts in the genetics of personality consortium: An application of item response theory. *Behavior Genetics, 44*(4), 295–313. https://doi.org/10.1007/s10519-014-9654-x

van den Berg, S. M., de Moor, M. H. M., Verweij, K. J. H., Krueger, R. F., Luciano, M., Vasquez, A. A., … Generation Scotland. (2016). Meta-analysis of genome-wide association studies for extraversion: Findings from the genetics of personality consortium. *Behavior Genetics, 46*(2), 170–182. https://doi.org/10.1007/s10519-015-9735-5

Verweij, K. J. H., Zietsch, B. P., Medland, S. E., Gordon, S. D., Benyamin, B., Nyholt, D. R., … Wray, N. R. (2010). A genome-wide association study of Cloninger's temperament scales: Implications for the evolutionary genetics of personality. *Biological Psychology, 85*(2), 306–317. https://doi.org/10.1016/j.biopsycho.2010.07.018

Verweij, K. J. H., Yang, J., Lahti, J., Veijola, J., Hintsanen, M., Pulkki-Raback, L., … Zietsch, B. P. (2012). Maintenance of genetic variation in human personality: Testing evolutionary models by estimating heritability due to common causal variants and investigating the effect of distant inbreeding. *Evolution, 66*(10), 3238–3251. https://doi.org/10.1111/j.1558-5646.2012.01679.x

Vilhjálmsson, B. J., Yang, J., Finucane, H. K., Gusev, A., Lindström, S., Ripke, S., … Schizophrenia Working Group of the Psychiatric Genomics Consortium, Discovery, Biology, and Risk of Inherited Variants in Breast Cancer (DRIVE) study. (2015). Modeling linkage disequilibrium increases accuracy of polygenic risk scores. *American Journal of Human Genetics, 97*(4), 576–592. https://doi.org/10.1016/j.ajhg.2015.09.001

Vinkhuyzen, A. A. E., Pedersen, N. L., Yang, J., Lee, S. H., Magnusson, P. K. E., Iacono, W. G., … Wray, N. R. (2012). Common SNPs explain some of the variation in the personality dimensions of neuroticism and extraversion. *Translational Psychiatry, 2*. https://doi.org/10.1038/tp.2012.27

Watson, J. D., & Crick, F. H. C. (1953). Molecular structure of nucleic acids – a structure for deoxyribose nucleic acid. *Nature, 171*(4356), 737–738. https://doi.org/10.1038/171737a0

Yang, J. A., Lee, S. H., Goddard, M. E., & Visscher, P. M. (2011). GCTA: A tool for genome-wide complex trait analysis. *American Journal of Human Genetics, 88*(1), 76–82. https://doi.org/10.1016/j.ajhg.2010.11.011

Zohar, A. H., Dina, C., Rosolio, N., Osher, Y., Gritsenko, I., Bachner-Melman, R., … Ebstein, R. P. (2003). Tridimensional personality questionnaire trait of harm avoidance (anxiety proneness) is linked to a locus on chromosome 8p21. *American Journal of Medical Genetics Part B-Neuropsychiatric Genetics, 117B*(1), 66–69. https://doi.org/10.1002/ajmg.b.10029