CHAPTER 10

SUMMARY

This thesis consists of eight studies that can be divided into four parts. The first part consists of two twin studies on Thought Problems (TP) where the longitudinal heritability is estimated, and rater bias of parental reports of TP in children and measurement invariance across age and sex of self-reports of TP in adolescents and adults is examined. In the second part of the thesis, the search for post-twinning *de novo* CNVs in monozygotic (MZ) twins is reported. The extent of CNV discordance between MZ twins is investigated in 50 MZ pairs selected for concordance and discordance on Attention Problems (AP), and in a group of ~1,100 MZ pairs, unselected for a particular phenotype. In the third part, genome-wide patterns reflecting Dutch ancestry are examined using single-nucleotide polymorphisms (SNPs) as typed on microarrays, and indels (short insertions or deletions) and larger deletions from Next Generation Sequence (NGS) data. In the final part, two studies are performed on ancestral non-random mating through migration and assortment and its effect on genome-wide variation through runs of homozygosity (ROHs).

Part I: Twin Studies: Rater Effects, Measurement Invariance, and Longitudinal Heritability of Thought Problems

The Thought Problems (TP) scale is an empirically derived set of items that measures symptoms common in several mental disorders: hallucinations, obsessive-compulsive symptoms, strange thoughts and behaviors, self-harm, and suicide attempts.

Chapter 2, shows that more than half of the variation of TP is explained by additive genetic factors in ~9,000 7-year-old twin pairs. The heritability estimated was based on parental ratings. The analyses modeled genetic and environmental influences on the commonly agreed upon part of the phenotype and on the unique views of the parents separately. The part of TP that parents agreed on explained ~67% of the variance, of which 76% was due to additive genetic influences. The unique part of the reported TP was also mostly heritable (maternal part: 61%, paternal part: 65%), indicating that the unique views of the parents very likely reflect real behavior of the child instead of rater bias. The unique views of the parents were, unlike the part they agreed on, also influenced by shared environmental factors (maternal part: 13%, paternal part: 13%). The remaining variance (24%) was explained by unique environmental influences and may also partly reflect measurement error.

In **chapter 3**, the strength and the structure of the relations between self-reported TP-items is investigated with an exploratory factor analysis (EFA), measurement invariance (MI) of the TP-scale is investigated across age and sex, and the extent of genetic and environmental influences is estimated longitudinally using the genetic relatedness of ~9,000 twin pairs and ~2,000 siblings. The EFA yielded a one-factor structure. The one-factor structure was then used in a

multigroup confirmatory factor analysis that led to the conclusion that the TP-scale is measurement invariant between adolescent (12-18 years old), young adult (19-27 years) and adult (28-59 years) males and females, in the sense that differences between these groups reflect real differences in the mean and/or variation of the construct measured by TP. About 37% of the variation in TP throughout adolescence, young adulthood, and adulthood is estimated to be due to additive genetic influences. TP was influenced by the same additive genetic component throughout the three age groups, with an additional genetic component arising during young adulthood, which keeps influencing the trait throughout adulthood as well. The remaining variance (63%) was explained by unique environmental influences and may also partly reflect measurement error.

Part II: Copy Number Variants: Post-Twinning Mutations and Discordance between Monozygotic Twins

In **chapter 4**, 50 MZ twin pairs were selected out of ~3200 MZ twin pairs based on their data on the Attention Problems (AP) scale: 17 concordant affected, 22 concordant unaffected, and 11 discordant pairs. The AP scale has been shown to be predictive for ADHD, and both AP and ADHD are highly heritable. CNVs were measured for this group of twins and in a subset (25 pairs) also for the parents. The presence of pre- and post-twinning CNVs was investigated, and an association analysis was conducted to test whether CNVs were associated with AP. Out of 26 *de novo* CNVs suggested by the Affymetrix 6.0 chip, three were replicated using qPCR: 1) a pre-twinning mutation (duplication) in a healthy twin pair, 2) a post-twinning mutation (deletion) in a concordant affected twin pair, and 3) a post-twinning mutation (duplication) in the affected twin of a discordant pair. The post-twinning mutations overlapped with genes that were associated in previous studies with the co-morbidity of psychiatric disorders that were also diagnosed in twins carrying the mutations. Besides more behavioral problems, the carriers of the post-twinning mutations also had a lower birth weight than their co-twin. In the entire sample of 50 twin pairs, association analyses of genome-wide CNV burden and AP showed that CNVs overlapping with genes were significantly larger in affected than in unaffected subjects. CNVs in non-genic regions did not show this association. For CNVs overlapping with genes, both deletions and duplications showed the same trend (i.e., larger in affected than unaffected individuals), but no significant differences, indicating that both contributed to the effect.

In **chapter 5** a genome-wide search was conducted for post-twinning CNVs in ~1,100 unselected MZ twin pairs, of which about half had their DNA extracted from buccal (mainly children), and the other half from blood (mainly adults). A total of 153 putative post-twinning *de novo* CNVs were found. The majority of these 153 CNVs resided in 15q11.2, of which the majority was

significantly overrepresented in blood-derived DNA, and observed in significantly older twins within the dataset of blood-derived DNA. Based on visual inspection of raw intensity signals done by multiple raters independently, eleven *de novo* CNVs were deemed suitable for a first series of qPCR follow-up experiments. All eleven resided in 15q11.2, of which two mutations within the same twin pair with buccal-derived DNA were validated by the qPCR experiments (~350kb and ~280 kb). The twins were thirteen years old at the time of sampling, and did not show large phenotypic differences based on parental and self-report questionnaires from ages 1 to 21.

Part III - Population Genetics: Genomic Structure of the Netherlands

In **chapter 6**, principal component analyses (PCAs) were conducted on ~500k genome-wide SNPs in 4,441 unrelated Dutch subjects and 1,014 unrelated subjects from 14 different world-wide populations from the 1000 Genomes dataset. Removing long-range linkage disequilibrium (LD) regions and LD-based SNP pruning (resulting in 130,248 SNPs) significantly changed the composition of the top ten PCs in the Dutch dataset, but barely influenced the top ten PCs in the 1000 Genomes dataset. Minimizing LD resulted in PCs with significantly higher correlations with Dutch geography, and also resulted in three ancestry-informative PCs: 1) the North–South PC, which differentiates the Southern provinces below the three major rivers from the Northern provinces, with the more urbanized West falling in between, 2) the East–West PC, which mainly differentiates the Northeastern part of the Netherlands from the rest, and 3) the Middle-Band PC, which separates the middle-band area of the Netherlands from the rest. The North-South PC showed several similarities with the European North-South cline: 1) a correlation of .66 with the 1000 Genomes PC that separates Northern from Southern European populations, 2) a significant correlation with genome-wide homozygosity (F; North = more homozygous), 3) a significant correlation with height (North = taller), 4) a signal of significant selection pressure with a SNP associated with blue/brown eye-color (North = more blue eyes). The extent of adaptive effects on the genetic differentiation within the Netherlands was investigated by comparing the distribution of alleles for ~500k SNPs between subpopulations identified by the three PCs. Besides the signal in HERC2 (i.e., eye color, which was the strongest signal), there were significant signals of diversifying selection pressures in 544 other SNPs, of which the majority resided in 184 genes, among which genes involved in brain function were significantly overrepresented.

In **chapter 7**, it was investigated whether genome-wide patterns reflecting ancestry were also detectable in structural variants (SVs) from a Next Generation Sequencing (NGS) dataset from a collaborative effort of four Dutch biobanks called *Genome of the Netherlands* (GoNL). PCA was conducted on 490

unrelated individuals from across the Netherlands on common indels (<20 bp) and common larger deletions (20 – 10,000 bp) separately (where common means: MAF > 1%). Geographical location and the three ancestry-informative PCs from chapter 6 (projected onto GoNL individuals) were used to determine which indel PCs and larger deletion PCs were likely to reflect ancestry. Indels showed three ancestry-informative PCs, and larger deletions showed five. The indel PC explaining most variation showed a correlation of .86 with the North-South SNP PC and showed the same geographic distribution. The other two indel PCs showed a geographic distribution that was similar to the East-West SNP PC from chapter 6, but correlated only ~.36 with the East-West SNP PC, and showed a greater genetic distance between the Eastern and Western subpopulations. The five larger deletion PCs showed significant but lower correlations with the SNP PCs from chapter 6 (with the significant correlations ranging from .16 to .30) and showed no clear geographic distributions when plotted on the Dutch map (although there were significant correlations with geography ranging from .18 to .26). Ancestry-informative SNP PCs and indel PCs show similar degrees of differentiation across SNPs, indels, and larger deletions. Larger deletion PCs however seem to capture genome-wide patterns that mostly reflect larger deletions. Larger deletion PCs were also the only ancestry-informative PCs to not show a significant spouse correlation. It is not yet clear whether the discrepancy between ancestry-signals from larger deletions and from the smaller SNPs and indels are due to larger deletions capturing older variation, or due to ancestry signals being weaker due to a limited amount of measured larger deletions.

Part IV - Runs of Homozygosity: How Ancestral Behaviors Influence Current Genetic Variation

In **chapter 8**, religious affiliation shows a significant association with autozygosity, which was measured by F_{roh}: the proportion of the genome that consists of runs of homozygosity (ROHs). It also became apparent that the geographic distribution of the North-South PC that was discovered in chapter 6 matches that of the major religious denominations in the Netherlands, which has had a stable geographic distribution for about four centuries. The distribution of religions in the current dataset showed that this distribution was maintained, but with an overall increase of non-religious individuals, which have become more numerous in more recent times. The spouse correlations for religious affiliation are quite high (.73), in line with the high spouse correlation of the North-South PC and historical documentation of strong religious assortment among the Dutch. Post-hoc association analyses revealed that the significant religion-F_{roh} association was due to a difference between the religious and non-religious group, with the non-religious group showing significantly fewer/shorter ROHs, which is likely

explained by the relatively recent absence of denominational restrictions on mate selection for non-religious people, which increased variation in the gene pool of possible mates. The non-religious group also showed significantly more cases of major depressive disorder (MDD), which caused a nominally significant association between autozygosity and MDD. This spurious association disappeared after correcting for religious affiliation. Ancestry-informative PCs were not able to correct for this, since they also did not capture the autozygosity differences between religious and non-religious groups.

In **chapter 9**, a significant association between educational attainment and F_{roh} is observed. Like in chapter 8, additional evidence shows that this association is not due to causal ROHs decreasing the chances for achieving a higher educational attainment, but due to ancestral mating behavior. Parental educational attainment was much more significantly associated with offspring F_{roh}, with higher educated parents having offspring with a smaller genomic proportion of ROHs. Parents with higher educational attainment showed significantly higher distances between their own birthplace and the birthplace of their offspring and a higher distance to the birthplace of their spouse. The distance between paternal and maternal birthplaces was also significantly associated with F_{roh}, and when included in the F_{roh}-education regression, the association between F_{roh} and parental educational attainment disappeared, while the association between F_{roh} and parental birthplaces was still significant. This indicates that the association between F_{roh} and (parental) educational attainment was due to higher educated parents showing higher migration rates, increasing the chances of choosing another higher and more mobile educated spouse with a different ancestral background (as spouse correlations were also very significant). This effect was also visible through significantly decreasing correlations between ancestry-informative PCs and geography in offspring of parents with a higher education.

CHAPTER 11

GENERAL DISCUSSION

Behavior genetics is the discipline that studies the role of genetics in human and other animal behavior. In the first half of this thesis, classical behavior genetics topics are explored, such as the role of phenotypic measurement and the twin design in estimating genetic and environmental influences on human traits, and molecular genetic data are incorporated to explore the extent of genetic similarity in monozygotic twins. In the second half of the thesis, molecular genetic, geographic, and phenotypic data are incorporated to examine genetic variation on a population level within a relatively small but densely populated country, and the role of behavior therein.

Behavior ← Genetics: Twin Studies

The twin study design, developed around a century ago,[1; 2] allows for deriving heritability estimates by comparing the phenotypic resemblance between monozygotic (MZ) twins with that of dizygotic (DZ) twins.[3] Family, twin, and adoption studies demonstrated in the early 20th century that our genome plays a strong and significant role in the development of psychological and psychiatric traits.[4] In the mid-20th century however, the prevailing view was that environmental influences were largely responsible for the development of these traits, until the 1960s and 1970s, when a more balanced view on the nature-nurture debate gained ground again in the scientific community, in part through twin and adoption studies.[5] Many have utilized this design by gathering a large number of phenotypic measurements of twin pairs, often using questionnaire data, mostly self-reports when it comes to adults, and parental and/or teacher reports on children. It is estimated that over 1.5 million twins and their families participate in twin studies worldwide.[6]

Part I - Twin Studies: Rater Effects, Measurement Invariance, and Longitudinal Heritability of Thought Problems

In the first part of the thesis, the impact of genetic and environmental influences are estimated for the empirically derived combination of questions referred to as "Thought Problems" (TP), which is one of the more difficult problem behavior scales to analyze in a population-based sample, unless large sample sizes are available, due to positive item answers having a very low prevalence. This scale has often been used in the context of how predictive it is for several psychiatric traits, such as OCD, schizophrenia, bipolar disorder, substance abuse, social phobia, anxiety and mood disorders, and is likely to measure a single underlying construct. The results from chapter 3 show that the TP scale measures one factor, and shares additive genetic influences across age. As detailed in the discussion of chapter 3, the

symptoms measured by the scale (hallucinations, OCD-symptoms, strange thoughts and behaviors, self-harm and suicide attempts) suggest that TP may be measuring a liability for a so-called "schizo-obsessive" disorder. Schizophrenia and OCD have been shown to occur together more often than expected by chance, and share common functional circuits and dysfunctions of neurotransmitter systems.[7]

The heritability of TP in young children estimated in chapter 2 (63%) is considerably larger than the heritability of the multiple older age groups in chapter 3 (37%). Based on these results, we cannot yet conclude that this means that TP is more heritable in children than in adolescents and adults. The estimates in chapter 2 were obtained with the so-called rater bias model, which decomposes the variance in scores of parental reports into a part of the phenotype that both parents agree on (referred to as the common part), and parts that reflect parents' unique views on the trait (the rater specific parts), and gives estimates on genetic and environmental influences on all these components separately. In chapter 3 the heritability estimates were based on self-reports from twins and their siblings. The sudden drop in heritability from children to adolescents has been observed previously for Attention Problems (AP) as well,[8] and is likely explained by the change from parental reports to self-reports as soon as children reached the age of 12.

This drop in heritability may be the result of rater effects similar to what we modeled in chapter two and was recently addressed by Kan et al (2014).[9] Kan et al (2014)[9] showed analytically that if rater specific factors are genetically influenced, heritability estimates depend on whether both twins are rated by the same individual (which is the case in chapter 2, where the children are both rated by the father and/or by the mother), or each by a different individual (like in chapter 3, where both twins are rated by themselves, i.e., two different individuals). When they are each rated by a different individual (i.e., themselves), only the heritability of the component of the trait that they both agree on (the common part) is estimated, while the rater specific component is subscribed to environmental influences. When they are both rated by the same person, their mother for example, the effects that only the mother observes (rater specific) genetic factors are added to the heritability estimates. It seems that we assessed the following four components of the TP trait in these two studies: 1) the common part of parental reports of 7-year-olds, 2) the rater specific part of maternal reports in 7-year-olds, 3) the rater specific part of paternal reports in 7-year-olds, and 4) the common part of self-reports in young adolescents, young adults, and adults (12-59 years old). The first three components are likely to overlap with the fourth, but we would need additional data to formally test this by including parental reports on 12-59 year old twins, and self-reports for the 7-year old twins. These additional data are difficult to obtain, especially in large numbers, so for now we should be careful in interpreting different heritability estimates derived from different raters and different studies. We

cannot say much with certainty yet about the longitudinal heritability changes that may or may not occur between children and young adolescents. If we follow the reasoning of Kan et al (2014),[9] the heritability estimates of the TP in adolescents and adults (currently estimated at 37%) are likely to increase if twins would rate themselves as well as their co-twins. What makes different raters report different components of the trait? Whatever the cause, these different components do seem to be "real", as they seem to be influenced by genes. Item level analyses on data from multiple raters may give us a better idea on what the common and rater specific components of these traits represent phenotypically. These different perspectives on the same underlying trait could give us a more complete picture and important information concerning the measurement of psychiatric (endo)phenotypes. Using more informants on better defined and more accurately modeled traits may provide more insight in the different heritable parts of a trait and may thus be helpful in the quest for finding causal genetic variants.

Part II - Copy Number Variants: Post-Twinning Mutations and Discordance between Monozygotic Twins

In chapters 4 and 5 the assumption of the twin study design that MZ twins are 100% genetically identical was investigated. Genetic differences that lead to phenotypic differences between MZ twins would confound heritability estimates from twin studies (they would lead to an underestimation of the genetic contribution), but could potentially be used to identify causal genetic variants, as was demonstrated for example in the identification of the causal variant underlying Van der Woude Syndrome.[10] *De novo* copy number variants (CNVs) have been identified as possible major risk factors for several psychiatric disorders.[11-13] The majority of current molecular genetic studies focus on single nucleotide polymorphisms (SNPs), while structural variants (SVs), such as CNVs, may have a two to four times higher mutation rate and cover a larger part of the genome,[14; 15] but are unfortunately also harder to reliably measure by microarray technology. Microarrays are better at measuring whether a nucleotide is an A or a G (i.e., a SNP) than at detecting how many times a certain DNA segment occurs.[16] This makes the chance for a false positive finding much higher than finding an actual genetic difference between MZ twins. Studies using microarrays for this goal without attempting to validate their *de novo* CNVs with qPCR usually report many more CNVs differences between MZ twins than studies that do (see reference[17] for example, which reports 21 post-twinning mutations in 2 MZ pairs, in comparison to the many studies in Table 1 from chapter 5, where the majority of putative mutations did not get validated).

Thus, searching for post-twinning de novo mutations in CNVs was an endeavor with relatively low chances of succeeding. The chance of success was

likely increased by heavily selecting discordant MZ pairs from a large sample of ~3200 MZ pairs as detailed in chapter 4. Some interesting results were found when we screened 50 MZ pairs selected for AP (11 discordant), and the raw intensity signals of the 1.3 Mb post-twinning deletion on chromosome 4 (Figure 1b of chapter 4) were especially convincing, and confirmed. Another promising feature of the chapter 4 results is that the somatic mutations overlapped with genes where CNVs were previously implicated with the co-morbidity of psychiatric diseases that were also observed in the carriers of the mutations. Among the 50 MZ pairs from chapter 4, we also validated a pre-twinning *de novo* CNV (i.e., present in both twins, but not the parents) in 15q11.2, a region that was grossly overrepresented among the putative post-twinning mutations in the ~1,100 MZ pairs from chapter 5 (90 out of 152). This is either an unusually unstable genomic region, or our approach is for some reason more sensitive for detecting CNVs in this region. Eleven of the putative post-twinning 15q11.2 mutations for qPCR replication, among which two (from the same twin pair) were validated; a higher validation rate than the qPCR replication experiments in the 50 selected MZ pairs of chapter 4, where we validated 2 out of 18 putative post-twinning mutations. The validation rate may still not seem high, but the rest of the 79 putative mutations in 15q11.2 are significantly more often obtained from blood samples, of which the majority are adults, as opposed to the buccal samples, that are mainly from children. In addition, blood samples with a putative 15q11.2 mutation are also from significantly older twins than blood samples in which we did not detect a putative discordance. This is in line with recent findings of *de novo* mutation rates increasing with age.[18;19] Given this association with age, it was not expected that the one twin pair in which the discordance was validated with qPCR was a 13 year old twin pair with a buccal sample, which did not show any striking phenotypic discordance. We are currently preparing for another round of qPCR replications among the rest of the putative somatic mutations, which should give us a better idea on somatic mutation rates for relatively large (>100 kb) CNVs, and on whether large *de novo* CNV mutations in the 15q11.2 region, which have been associated with a wide range of psychiatric and cognitive disorders, can be phenotypically tolerated.

Behavior → Genetics: Ancestral Influences on Genetic Variation

The molecular structure of DNA, and thereby the molecular mechanisms underlying the coding of genetic information and inheritance, was discovered in 1953.[20-22] This discovery and the major technological advances that followed were largely responsible for commencing the so-called Genomic Era, which is still ongoing, and paved the way for mapping genetic variation through genome-wide genotyping. These advances led to a gold rush for causal genetic variants, resulting in large international collaborations analyzing vast amounts of genomes through genome-wide association studies (GWASs), where we are reaching the maximum of practically feasible sample sizes.[23] Some psychiatric traits, such as depressive or anxiety disorders, turn out to be among the most difficult complex traits with respect to finding consistent genetic association signals with the current available sample sizes.[24; 25] In the understandable hurry to disentangle the genetic etiology of complex traits and disease, understanding the larger patterns of genetic variation (and where they come from) within populations contributing to these meta- and mega-analyses may not get the urgency it deserves.

Part III - Population Genetics: Genomic Structure of the Netherlands

Because of the small geographic area, the Netherlands is among the regions sampled from when a genetically homogeneous sample is desired.[26-28] However, even in such small areas considerable genetic heterogeneity may exist. We used a principal component analysis (PCA) on genome-wide single nucleotide polymorphisms (SNPs) to summarize the largest patterns of genetic variation in the Netherlands. Principal components (PCs) reflecting ancestry differences have been shown to effectively correct for population stratification in genetic association studies, i.e., allele frequency differences due to systematic ancestry differences, which can confound signals in GWASs if they correlate with the studied phenotype.[29] Extracting reliable signals for ancestry differences from a relatively homogeneous dataset required a different and more stringent approach than in a dataset consisting of multiple more differentiated populations (e.g., the 1000 Genomes dataset). This allowed for three Dutch ancestry-informative PCs capturing relatively small ancestry differences to be extracted from microarray SNP data, which showed clear and significant correlations with geography (a North-South, an East-West, and a Middle-Band distribution), and are not independent of complex traits. SNP data extracted from the *Genome of the Netherlands* (GoNL) dataset, a Next Generation Sequence (NGS) dataset representing participants from all regions from the Netherlands from multiple Biobanks, show the same geographic distributions

when projecting the PCs derived from the Affymetrix 6 SNPs. When conducting a PCA on common indels from the GoNL dataset, the North-South and East-West cline are also clearly visible. The absence of the "Middle-Band cline" among the indel PCs might be explained by a lack of power, since that only emerged in 4,441 unrelated subjects when LD was minimized beyond the recommended levels, and when outliers due to a non-Dutch ethnic/ancestral background and non-genetic artifacts were removed. The GoNL dataset was considerably smaller (490 unrelated individuals), and non-genetic artifacts related to quality differences were still present in PCs derived from indels, which is not entirely unexpected given that they are harder to reliably measure than SNPs. Results from the GoNL dataset do confirm however that common indels are largely in LD with common SNPs included in microarrays, which means that their variation is also largely captured by common SNPs, which supports the reliability of indels imputed from microarray SNPs for GWASs.

Height is significantly correlated with the PC explaining most variation, which captures the Dutch North-South cline. Height is a classical example of a complex trait and has long served as a model for the investigation of the genetic etiology underlying complex traits[30-33]: it is influenced by many genetic variants, relatively easy to measure reliably on a large scale, and therefore turned out as one of the more successful complex traits when it comes detecting causal genetic variants.[34] We show that improving measures for ancestry by decreasing LD results in PCs that are more effective in correcting for inflated statistics in GWASs on height that are caused by systematic North-South ancestry differences, which represent more than just causal SNPs. It is very likely that the causal variants underlying the height differences between North and South differ between these regions because of selection pressures. This is the case for European North-South height differences as well.[35] The European North-South cline correlates highly with the Dutch North-South cline and shows several other similarities, such as a significant correlation with genome-wide homozygosity due to the serial founder effect that was initiated with the ancient successive out-of-Africa migrations, and selection pressures on many of the same genes. This does not necessarily mean that these events (north-ward migration and diversifying selection) took place within the borders of the Netherlands; it could also be that Southern Europeans have migrated more to the South of the Netherlands, and/or Northern Europeans more to the Northern parts.

By comparing the distribution of alleles between subpopulations identified by the ancestry-informative PCs, we were able to detect a relatively large number of SNPs in genes under diversifying selection pressures. These variants likely had relatively large effects on phenotypes that were important for survival and/or reproductive success in Dutch ancestors. The variant with the strongest signal is

the key determinant of human blue/brown eye color, which also shows significant selection pressures in Europe,[36; 37] where (like in the Netherlands) blue eyes are more prevalent in Northern regions.[38] Phenotypes under diversifying selection likely included brain-related traits, as several important and well-known brain genes showed significant signals, such as SLC6A4 (a.k.a. SERT; encodes the serotonin transporter), BDNF (encodes the brain-derived neurotrophic factor), NRXN3 (encodes neurexin-3-alpha), GRIN2A (encodes a subunit for the NMDA receptor), GRM7 (encodes a metabotropic glutamate receptor), and AUTS2 (autism susceptibility candidate 2). In addition, genes involved in neurotransmission of nervous tissue were significantly overrepresented among selection pressure signals overall. SLC6A4, one of the most studied genes in candidate gene studies in the context of psychiatric traits (especially regarding depression), showed recent selection pressures in other populations as well.[39-41] It remains to be elucidated why we are able to pick up these signals, while GWASs on behavioral and psychiatric traits rarely show significant associations in these classical candidate genes, especially when using sample sizes close to ours (<5,000 subjects). A possible explanation could be that nature "measures" these traits across many more generations than scientists do, or perhaps scientists and nature have a different opinion about what they call a "disease" or an important trait. One thing seems clear however: the ancestry differences for these genes are larger than the causal effects they have on current measurable behavioral and psychiatric traits.

Part IV - Runs of Homozygosity: How Ancestral Behaviors Influence Current Genetic Variation

Genetic variants in the Netherlands also show non-causal relationships with behavioral traits. The geographic distributions captured by the ancestry-informative PCs are visible because of relatively low levels of migration rates (and thus low levels of gene flow) in the relatively recent Dutch history. Non-random mating through assortment and migration however has created non-random genetic differences between certain phenotypic groups. The Dutch North-South PC shows the same geographic distribution of Protestants and Catholics in the last couple of centuries. When populations are separated geographically and socially for longer periods of time, they will eventually diverge from each other genetically as well (largely due to genetic drift). Protestants and Catholics have been strongly segregated in the Netherlands for centuries, which is also visible in the high levels of assortment during these times and in our dataset (spouse correlation = .73). We are likely picking up the genetic consequences of that with the North-South PC, which is also the PC that shows the strongest assortment (spouse correlation = .56, which is >3 times larger than the spouse correlation of the other two ancestry-informative PCs). Non-religious individuals however have rapidly increased in numbers during

the last 50 years. They are less restricted to mates with similar ancestries, as they are more likely to migrate and are free of denominational restrictions in their partner choice. This makes non-religious individuals less related to their mates than religious individuals, leading to offspring with significantly smaller genomic proportions of runs of homozygosity (ROHs, i.e., consecutive homozygous SNPs). We see a similar effect for education: higher educated individuals are more likely to migrate and pick a higher educated partner who is more likely to have come from a different geographic region. This makes higher educated spouse pairs genetically less similar to each other than lower educated spouse pairs, leading to offspring carrying significantly fewer/shorter ROHs. This non-random mating behavior caused the ancestry-informative PCs (which represent the largest patterns of genome-wide variation when excluding patterns due to LD) to be more mixed in the offspring, which caused these PCs to show lower correlations with geography in non-religious and higher educated individuals. These systematic differences affect many non-causal variants, which can lead to spurious associations with traits related to religion and education, and this can lead to wrong conclusions, especially when the spurious association happens to coincide with existing hypotheses. The ancestry-informative PCs did not capture systematic homozygosity differences of this subtle nature sufficiently to account for them in association analyses. These effects would probably be more absent in the GoNL dataset, where spouse pairs were not selected randomly with respect to shared ancestry (spouse pairs were chosen to be born in the same province, which was detectable through inflated spouse correlations for ancestry-informative PCs). This likely decreased the chance for sequencing higher educated or non-religious trios, since their parents are more likely to share different ancestral backgrounds.

The expected consequences of these confounding effects are not much different than those following from the classical population stratification issues, like with height, where the assumption that GWASs are conducted in genetically homogeneous samples is violated. This can especially be a problem when these systematic genetic differences within the sample are related to the trait under investigation. Religion and educational attainment are associated with many other psychiatric, behavioral, and cognitive traits of interest. We detect the confounding effects using a genome-wide measure of homozygosity and PCs summarizing genome-wide variation (both computed using ~500k directly measured SNPs), but have yet to explore the impact this might have on single-SNP associations in GWASs on these traits. Much larger sample sizes are needed to investigate this further, and it is important that the presence of these effects is also investigated in other populations contributing to meta-analyses. Educational attainment needed >125k subjects from 57 cohorts to find three borderline significant associations with p-values ranging from 2.1×10^{-9} to 4.2×10^{-9}.[42] Admittedly, educational

attainment is a difficult heterogeneous phenotype to analyze, for example because of cross-cultural and cross-generation differences in availability and access to education. It is one of the few practical proxies for IQ however that is able to reach sample sizes required to detect signals from the many alleles with weak effects among the millions of likely non-contributing variants. It is worth investigating whether these confounding effects vary between populations with different social, historical, and demographic backgrounds, and whether statistical power can be improved in meta-analyses by accounting for this non-random variation. A recent study on the relationship between MDD and F_{roh} that included nine datasets from five countries found MDD-F_{roh} associations with opposite directions of effects between datasets that were consistent across countries.[43] The association in the Dutch dataset from chapter 8, of which a considerable part was included in that study, disappeared when accounting for religious affiliation, which was significantly associated with MDD in our dataset. When dealing with behavioral, cognitive, or psychiatric traits that may correlate with social, historical, and/or demographic factors, alternative explanations for genetic associations should be considered before concluding the finding of risk increasing genetic variants. In order to do this, we must have a good understanding of what drives genetic variation on a population level.

Behavior ↔ Genetics: Main Conclusions and Future Perspective on the Field

Part I: Twin studies aiming to estimate genetic and environmental influences on complex behavioral and psychiatric traits should be aware of the impact of rater effects in their estimates. Heritability estimates may show considerable variation depending on who provides the report and on how many individuals provide the report. Including different raters when collecting phenotypic measurements may provide different perspectives on certain traits that may all turn out to contain useful information.

Part II: Searching for genetic differences between MZ twins is a promising endeavor that may lead to the discovery of novel causal genetic variants. Microarray data is not yet optimal for this goal, as it is only suited for detecting relatively large CNVs and contains a considerable amount of false positive signals. Sequencing phenotypically discordant twins may be a more suited approach, since that makes it possible to scan the entire genome without ascertainment biases that microarrays may have in their probe selection, and are better suited for finding smaller genetic differences as well.

Part III: I would recommend consortia contributing to large GWAS meta- and mega-analyses to explore the main patterns of variation in their population, and also explore how social, historical, and demographic factors shaped their population structure. Especially in more homogeneous populations, one should make sure that PCAs conducted for this goal are carried out with care, by making sure that patterns of variation such as LD patterns and non-genetic artifacts are accounted for, since those can be larger than the relatively small ancestry differences within their population. These ancestry differences may be associated with phenotypic measures of interest. Even though these ancestry differences are relatively small, they may be greater than the very small effects the many individual SNPs have on complex traits.

Part IV: The presence of more recent ancestral influences that may have led to non-random mating should also be explored, because they may create systematic differences in genome-wide homozygosity due to systematic differences in parental relatedness. Heritable behavioral traits contributing to these differences may be associated with additional traits of interest. This is especially important for genetic studies on the effects of inbreeding, but may also result in less confounded GWAS analyses.

References

1. Galton, F. (1876). The history of twins, as a criterion of the relative powers of nature and nurture. *The Journal of the Anthropological Institute of Great Britain and Ireland* 5, 391-406.
2. Rende, R.D., Plomin, R., and Vandenberg, S.G. (1990). Who discovered the twin method? *Behavior Genetics* 20, 277-285.
3. Boomsma, D., Busjahn, A., and Peltonen, L. (2002). Classical twin studies and beyond. *Nature Reviews Genetics* 3, 872-882.
4. Slater, E. (1936). The Inheritance of Manic-depressive Insanity (Section of Psychiatry). *Proceedings of the Royal Society of Medicine* 29, 981-990.
5. Plomin, R., Owen, M.J., and McGuffin, P. (1994). The genetic basis of complex human behaviors. *Science* 264, 1733-1739.
6. Hur, Y.-M., and Craig, J.M. (2013). Twin registries worldwide: an important resource for scientific research. *Twin Research and Human Genetics* 16, 1-12.
7. Bottas, A., Cooke, R.G., and Richter, M.A. (2005). Comorbidity and pathophysiology of obsessive–compulsive disorder in schizophrenia: Is there evidence for a schizo-obsessive subtype of schizophrenia? *Journal of Psychiatry and Neuroscience* 30, 187.
8. Kan, K.J., Dolan, C.V., Nivard, M.G., Middeldorp, C.M., van Beijsterveldt, C.E., Willemsen, G., and Boomsma, D.I. (2013). Genetic and environmental stability in attention problems across the lifespan: evidence from the Netherlands twin register. *Journal of the American Academy of Child and Adolescent Psychiatry* 52, 12-25.
9. Kan, K.-J., van Beijsterveldt, C.E.M., Bartels, M., and Boomsma, D.I. (2014). Assessing genetic influences on behavior: Informant and context dependency as illustrated by the analysis of Attention Problems. *Submitted*.
10. Kondo, S., Schutte, B.C., Richardson, R.J., Bjork, B.C., Knight, A.S., Watanabe, Y., Howard, E., de Lima, R.L.F., Daack-Hirsch, S., et al. (2002). Mutations in IRF6 cause Van der Woude and popliteal pterygium syndromes. *Nature genetics* 32, 285-289.
11. Cook Jr, E.H., and Scherer, S.W. (2008). Copy-number variations associated with neuropsychiatric conditions. *Nature* 455, 919-923.
12. Sebat, J., Lakshmi, B., Malhotra, D., Troge, J., Lese-Martin, C., Walsh, T., Yamrom, B., Yoon, S., Krasnitz, A., et al. (2007). Strong association of de novo copy number mutations with autism. *Science* 316, 445-449.
13. Xu, B., Roos, J.L., Levy, S., Van Rensburg, E., Gogos, J.A., and Karayiorgou, M. (2008). Strong association of de novo copy number mutations with sporadic schizophrenia. *Nature Genetics* 40, 880-885.
14. Lupski, J.R. (2007). Genomic rearrangements and sporadic disease. *Nature Genetics* 39, S43-S47.
15. van Ommen, G.-J.B. (2005). Frequency of new copy number variation in humans. *Nature Genetics* 37, 333-334.
16. Baker, M. (2010). Genomics: the search for association. *Nature* 467, 1135-1138.
17. Maiti, S., Kumar, K.H.B.G., Castellani, C.A., O'Reilly, R., and Singh, S.M. (2011). Ontogenetic de novo copy number variations (CNVs) as a source of genetic individuality: studies on two families with MZD twins for schizophrenia. *PLoS One* 6, e17125.
18. Forsberg, L.A., Rasi, C., Razzaghian, H.R., Pakalapati, G., Waite, L., Thilbeault, K.S., Ronowicz, A., Wineinger, N.E., Tiwari, H.K., et al. (2012). Age-related somatic structural changes in the nuclear genome of human blood cells. *Am J Hum Genet* 90, 217-228.
19. Kong, A., Frigge, M.L., Masson, G., Besenbacher, S., Sulem, P., Magnusson, G., Gudjonsson, S.A., Sigurdsson, A., Jonasdottir, A., et al. (2012). Rate of de novo mutations and the importance of father/'s age to disease risk. *Nature* 488, 471-475.
20. Franklin, R.E., and Gosling, R.G. (1953). Molecular configuration in sodium thymonucleate. *Nature* 171, 740-741.
21. Watson, J.D., and Crick, F.H. (1953). Molecular structure of nucleic acids. *Nature* 171, 737-738.
22. Wilkins, M.H.F., Stokes, A.R., and Wilson, H.R. (1953). Molecular structure of nucleic acids: molecular structure of deoxypentose nucleic acids. *Nature* 171, 738-740.
23. Visscher, P.M., Brown, M.A., McCarthy, M.I., and Yang, J. (2012). Five years of GWAS discovery.

The American Journal of Human Genetics 90, 7-24.
24. Burmeister, M., McInnis, M.G., and Zöllner, S. (2008). Psychiatric genetics: progress amid controversy. *Nature Reviews Genetics* 9, 527-540.
25. Sullivan, P.F., Daly, M.J., and O'Donovan, M. (2012). Genetic architectures of psychiatric disorders: the emerging picture and its implications. *Nature Reviews Genetics* 13, 537-551.
26. Rietschel, M., Mattheisen, M., Degenhardt, F., Kahn, R.S., Linszen, D.H., van Os, J., Wiersma, D., Bruggeman, R., Cahn, W., et al. (2011). Association between genetic variation in a region on chromosome 11 and schizophrenia in large samples from Europe. *Molecular Psychiatry* 17, 906-917.
27. Zeegers, M.P., Khan, H.S., Schouten, L.J., van Dijk, B.A., Goldbohm, R.A., Schalken, J., Shajahan, S., Pearlman, A., Oddoux, C., et al. (2010). Genetic marker polymorphisms on chromosome 8q24 and prostate cancer in the Dutch population: DG8S737 may not be the causative variant. *European Journal of Human Genetics* 19, 118-120.
28. Buizer-Voskamp, J.E., Muntjewerff, J.-W., Strengman, E., Sabatti, C., Stefansson, H., Vorstman, J.A., and Ophoff, R.A. (2011). Genome-wide analysis shows increased frequency of copy number variation deletions in Dutch schizophrenia patients. *Biological Psychiatry* 70, 655-662.
29. Price, A.L., Patterson, N.J., Plenge, R.M., Weinblatt, M.E., Shadick, N.A., and Reich, D. (2006). Principal components analysis corrects for stratification in genome-wide association studies. *Nature Genetics* 38, 904-909.
30. Fisher, R.A. (1919). XV.—The Correlation between Relatives on the Supposition of Mendelian Inheritance. *Transactions of the Royal Society of Edinburgh* 52, 399-433.
31. Galton, F. (1886). Hereditary stature. *Nature* 33, 317.
32. Hemani, G., Yang, J., Vinkhuyzen, A., Powell, J.E., Willemsen, G., Hottenga, J.-J., Abdellaoui, A., Mangino, M., Valdes, A.M., et al. (2013). Inference of the Genetic Architecture Underlying BMI and Height with the Use of 20,240 Sibling Pairs. *The American Journal of Human Genetics* 93, 865-875.
33. Yang, J., Benyamin, B., McEvoy, B.P., Gordon, S., Henders, A.K., Nyholt, D.R., Madden, P.A., Heath, A.C., Martin, N.G., et al. (2010). Common SNPs explain a large proportion of the heritability for human height. *Nature Genetics* 42, 565-569.
34. Allen, H.L., Estrada, K., Lettre, G., Berndt, S.I., Weedon, M.N., Rivadeneira, F., Willer, C.J., Jackson, A.U., Vedantam, S., et al. (2010). Hundreds of variants clustered in genomic loci and biological pathways affect human height. *Nature* 467, 832-838.
35. Turchin, M.C., Chiang, C.W., Palmer, C.D., Sankararaman, S., Reich, D., and Hirschhorn, J.N. (2012). Evidence of widespread selection on standing variation in Europe at height-associated SNPs. *Nature Genetics* 44, 1015-1019.
36. Chen, H., Patterson, N., and Reich, D. (2010). Population differentiation as a test for selective sweeps. *Genome Research* 20, 393-402.
37. McEvoy, B.P., Montgomery, G.W., McRae, A.F., Ripatti, S., Perola, M., Spector, T.D., Cherkas, L., Ahmadi, K.R., Boomsma, D., et al. (2009). Geographical structure and differential natural selection among North European populations. *Genome Research* 19, 804-814.
38. Frost, P. (2006). European hair and eye color: A case of frequency-dependent sexual selection? *Evolution and Human Behavior* 27, 85-103.
39. Voight, B.F., Kudaravalli, S., Wen, X., and Pritchard, J.K. (2006). A map of recent positive selection in the human genome. *PLoS Biology* 4, e72.
40. Crespi, B., Summers, K., and Dorus, S. (2007). Adaptive evolution of genes underlying schizophrenia. *Proceedings of the Royal Society B: Biological Sciences* 274, 2801-2810.
41. Gelernter, J. (2014). SLC6A4 polymorphism, population genetics, and psychiatric traits. *Human Genetics*, 1-3.
42. Rietveld, C.A., Medland, S.E., Derringer, J., Yang, J., Esko, T., Martin, N.W., Westra, H.-J., Shakhbazov, K., Abdellaoui, A., et al. (2013). GWAS of 126,559 Individuals Identifies Genetic Variants Associated with Educational Attainment. *Science* 340, 1467-1471.
43. Power, R.A., Keller, M.C., Ripke, S., Abdellaoui, A., Wray, N.R., Sullivan, P.F., and Breen, G. (2014). A recessive genetic model and runs of homozygosity in major depressive disorder. *American Journal of Medical Genetics Part B: Neuropsychiatric Genetics*.