## Discovering the genetic architecture of the mind

Karlsson Linnér, R.

2019

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

# General summary

Twin and family studies have convincingly established that most, if not all, human traits are heritable to some degree. Yet, few genetic variants have been discovered that can be attributed for the resemblance in psychology and behavior that is observed in families. Until recently, the available technology restricted the sample size of molecular genetic studies, which limited researchers to be able to detect only highly penetrant variants. At the same time, repeated empirical observations suggested that the genetic influences on genetically complex traits, such as psychology and behavior, are orders of magnitude smaller compared to what was previously thought and distributed across a very large set of genetic variants. Now, modern genotyping technologies are drastically changing the investigation of the molecular genetic basis of complex traits by rapidly decreasing the cost of measuring individual genotypes. This advancement has led to a surge of data generation in epidemiological cohorts and large biobank initiatives and we are finally at a point where samples sizes are large enough to discover the very small genetic effects that influence highly polygenic traits.

Because of that development, there is now a cascade of scientific efforts, across fields and disciplines, with the objective to discover the genetic architecture of complex traits and disorders, in some aspects comparable to the space race of the 20th century. Many researchers aspire to put a flag on previously uncharted genetic territory. These efforts are considered to be a potential game changer in the investigation of mental traits, not only for the detection, prevention, and treatment of various mental disorders, but also for how empirical research can be conducted in fields like psychology and the social sciences. In that spirit, this thesis investigates the molecular genetic basis of various mental traits, such as behaviors, moods, and preferences, with the intent to contribute to the first steps towards discovering the genetic architecture of the human mind.

This thesis consists of four empirical studies that study a few specific measures of psychology and behavior in relation to a particular type of molecular genetic variation. The objective of the first two empirical chapters is to discover novel associations with single-nucleotide polymorphisms (SNPs), the most occurring form of genetic variation, by performing genome-wide association studies in hundreds of thousands and up to a million individuals. The objective of the third empirical chapter is to discover novel associations with CpG methylation, an epigenetic marker, by performing an epigenome-wide association study. The objective of the fourth empirical chapter is to critically review and meta-analyze the literature on a reported gene-environment interaction.

In **Chapter 2**, we study individual differences in the overall willingness to take risks, which we refer to as general risk tolerance, in combination with self-reported adventurousness and four risky behaviors: automobile speeding, alcohol consumption, smoking initiation, and number of lifetime sexual partners. By studying hundreds of thousands to almost a million individuals we discover hundreds of novel genetic loci associated with these traits. Yet, the loci account for only a small part of the missing heritability. Thus, there are many genetic variants left for future studies to discover as larger samples become available. Nonetheless, the findings reveal several interesting aspects of the traits' genetic architectures, such as the possible involvement of structural variation, in the form of inversions. We find that most of the associations are located outside of genes, and thus, are likely to be involved in the regulation of gene expression. Also, the study estimates great genetic overlap across general risk tolerance, risky behaviors, and many other traits from various research domains. For general risk tolerance, bioinformatic analyses finds strong enrichment of the central nervous system, and weaker enrichment of the category immune/hematopoietic. The study uncovers evidence in favor of biological mechanisms that were previously not thought to be related to risk taking, in the form of

glutamatergic and GABAergic neurotransmission, and casts doubt on the pathways and candidate genes that have previously been hypothesized to influence risk taking.

In **Chapter 3**, we investigate three mood phenotypes: subjective well-being, depressive symptoms, and neuroticism. The three traits are strongly correlated phenotypically and there is a shared genetic basis, which motivates a joint study of the three. Importantly, larger genetic studies of depression have been warranted for a long time. By performing genome-wide association studies in roughly 150,000 to 300,000 individuals we discover a handful of associated loci, which are among the first reported for these traits. Yet, their joint effects are so small that the study barely scratches the surface of their genetic architectures. Nonetheless, the findings are a first step in the discovery of the genetic variants responsible for the missing heritability of depression. We indeed find great genetic overlap across the three traits, as well as with anxiety disorders. Associations with neuroticism tag two inversion polymorphisms, which provides evidence in favor of the premise that structural variants are involved in determining individual differences in mental traits. Bioinformatic analyses finds strong enrichment of the central nervous system for all three traits, and weaker enrichment of the category adrenal/pancreas for subjective well-being and depressive symptoms.

In **Chapter 4**, we investigate whether educational attainment, which is a cognitive trait influenced by the social environment but also a major life experience, is associated with CpG methylation across the genome. Methylation is an epigenetic mechanism that is considered to play a major role in aging and a range of medical conditions. We perform an epigenome-wide association study in almost 11,000 individuals, which is by far the largest study on the association between a social factor and methylation. Disappointingly, we identify only a small number of epigenetic associations, and most do not appear to be robustly associated with educational attainment but instead driven by confounding bias attributable to own and/or maternal smoking. Overall, the overwhelming effect of smoking exposure on methylation appears to be a major source of concern for epigenetic studies of the social environment, because many such factors correlate with smoking to a varying degree. Nonetheless, because of the large sample size our estimates provide an approximate upper bound to the strength of association that can be expected for a biologically distal factor. Importantly, the findings suggest that the association with a social environmental influence, which is experienced over many years early in life, appears to be many times weaker than that of risk factors with a direct biological impact, such as smoking, alcohol consumption, and BMI.

In **Chapter 5**, we critically review and meta-analyze the literature that has hypothesized that antisocial behavior is associated with a gene-environment interaction between adverse life events and the gene region *5-HTTLPR*. The number of published studies is not great, but the meta-analysis includes about 8,000 individuals. The study identifies a significant interaction effect. However, a range of methodological shortcomings severely restrict the evidential value in support of that interaction and the possibility to draw a firm conclusion. The major sources of concern are differences in study design and statistical procedures, inadequate statistical modeling of all relevant covariate×gene and covariate×environment interaction terms, and the high risk of bias from population stratification because of admixed-ancestry samples and lack of genetic principal components. In addition, a more recent genome-wide association study lends little support for a direct effect of *5-HTTLPR* on antisocial behavior, which casts further doubt on the hypothesis that *5-HTTLPR* would be involved in the moderation of the effect of adverse life events.