Introduction

The study of how we are formed as social animals, the reasons for why we think and behave the ways we do, and ultimately why we end up with the life course outcomes we do, is arguably the core of the social sciences broadly speaking. Why are some people more likely to be highly educated, end up in high status jobs, be healthier and live longer, be more politically engaged, and, in general, happier than others? Understanding the origins of such social, economic and political inequalities in life course outcomes between individuals and segments of society is of fundamental importance for a wide range of questions.

The metaphor of the blank slate, stating that we are born as empty canvases and socialized by family and society into complex social beings, has in large part guided mainstream social scientific research for decades. However, in addition to social and economic determinants, recent interdisciplinary research shows decisively that almost all individual-level outcomes are to some degree also influenced by genetics (Polderman et al. 2015). This finding applies to outcomes like educational attainment, social inequalities and political participation, just as it does to more intuitively genetically influenced traits such as health, cognitive abilities and personality. Against this backdrop a consensus is steadily growing that the assumptions imbued in the classic blank slate metaphor need to be replaced by a more nuanced account of the social *and* genetic origins of life outcomes (Pinker 2003).

Furthermore, omitting the genetic part of inter-generational transmission – that is, failing to take into account that we are not only raised by our parents, but we also inherit a combination of their DNA – neglects an integral part of the explanation of life outcomes because genetic differences between individuals not only *add* to social and environmental influences but also co-vary and *interact* with them in complex ways (Spinath & Bleidorn 2017). Studying either factor in isolation therefore necessarily paints an incomplete picture. Consequently, considering genetic influences by no means negates social influences, but rather provides an additional layer of explanation that can substantially improve our understanding of how they work. As such, it can also aid in developing more effective policies that deal with the social roots and consequences of inequality.

The knowledge of how, more precisely, genetic factors interact with social and environmental conditions in influencing life course outcomes, is so far very limited (Conley & Rauscher 2013). Nevertheless, these so-called gene-by-environment effects (G*E) are widely believed to be pervasive for social and behavioral traits. They arise when the type or magnitude of the effect of a genetic factor depends on the environmental conditions in which it is expressed. We can suspect, for example, that a predisposition toward scholarly achievement may only matter under the right environmental circumstances, or that certain genetic risk factors only matter if they are also *combined* with social risk factors. The adage that genetics loads the gun, but environment pulls the trigger, is an apt summary.

Despite a strong suspicion that G*E effects are common and important, to date the research on G*E phenomena has been hampered by a number of practical and methodological problems. These problems have resulted in a lot of previous research on the subject being under-powered, often lacking credible identification strategies, and therefore subject to problematic biases and considerable uncertainty. The purpose of this project is to take the burgeoning field of social science

genomics to the next level, incorporating a number of methodological innovations and unique data described below, to credibly and robustly investigate G*E effects. This ambition puts us at the cutting edge of this exciting interdisciplinary research field.

Project description

Theory and method

Up until recently, the most common approach taken to study G*E effects on complex social traits has been to consider single genes in so called candidate gene studies. In a candidate gene study, a researcher specifies ex ante hypotheses about a small set of genes or genetic markers (typically in the 1–10 range) and then runs association analyses between each of these pre-determined genetic markers and the outcome in focus. Ideally, these hypotheses are derived from the known biological function of the gene or genetic marker. In practice, the hypotheses are often based on previously reported associations with the same outcome or a related outcome, or the choice of genetic markers is a result of their availability in the data set the researchers are using.

An informative and also prominent example of this approach is an often-cited study by Caspi et al. published in *Science* in 2003. Based on a sample of less than 1000 individuals, the authors of this study presented evidence showing that the impact of stressful life events on depression and suicidality was significantly worse for individuals with a particular version of the serotonin transporter gene (Caspi et al. 2003). These results are both interesting and provocative and, consequently, the Caspi et al. study evoked a lot of debate, both within and outside of academia.

Caspi et al. (2003) also serve as a good example of the two major methodological challenges researchers studying G*E interactions must wrestle with (Benjamin et al. 2012). First, whenever G is a single gene or genetic marker, its effect size (both "main effect" and interactions) is likely to be very small, implying that most existing studies, which are based on at most a few thousand individuals, are dramatically underpowered. A second challenge is providing evidence for a causal gene-environment interaction since most existing research is correlational in the sense that no strategy for identifying exogenous variation in environmental conditions is used. As a result of this the Caspi et al. study, as well as the great majority of other similar G*E studies on complex social traits, has subsequently failed to replicate in large-scale efforts (Conley & Rauscher 2103; Spinath & Bleidorn 2017; Culverhouse et al. 2018). Thus, in order to solve the problems that have affected previous G*E research, new approaches – empirically, methodologically and theoretically – are necessary.

To get a grasp of our proposed approach to studying G*E effects a few more words on the problems with candidate gene studies are necessary. The candidate gene approach is now considered problematic in most applications, as it has become increasingly understood that complex behavioral phenotypes are not influenced by just one, or even a handful, of genetic variants. Rather, such traits are influenced by a very large number of genes, likely in the hundreds or thousands, each with a very small effect. These effects, taken all together, sum up to the total genetic heritability of the trait. The phenomenon of heritability being composed of many genes of small effect has come to be referred to as the *fourth law of behavior genetics*, or the polygenicity of

behavioral traits (Chabris et al. 2015). A consequence of this is that any study of a complex human trait attempting to measure the effect of a single gene, since any given gene is going to have a very small effect, is going to require enormous sample sizes to be well-powered. In almost any practical G*E application, where both genotype data and an array of other individual and environmental variables are necessary, such sample sizes do not exist. Furthermore, the tiny effects of single genes are rarely meaningful in themselves for advancing a substantive understanding of how genetics shape life outcomes. A serious investigation of G*E effects on complex social outcomes therefore needs a different approach.

What would such an approach look like? First and foremost, some way of identifying the multiplicity of genes implicated in a polygenic trait is necessary. Enter the *genome-wide association study* (GWAS). The dramatic decreases in costs for genotyping and the consequently increasing availability of very large samples of genotyped individuals (approaching the millions) have recently made the methodological innovation of GWAS possible.

To understand the GWAS approach a brief primer on some basic concepts will be helpful. Human DNA is composed of a sequence of approximately 3 billion pairs of nucleotide molecules. At the overwhelming majority of these 3 billion locations in the genome, there is no variation in the nucleotides across individuals. The segments of DNA in which individuals do differ are called polymorphisms. What we here sometimes refer to as genetic markers or genetic variants refer to the simplest and most common kind of genetic polymorphism called a single-nucleotide polymorphism (SNP). SNPs are locations in the DNA sequence in which individuals may differ from each other in terms of a single nucleotide. At the vast majority of SNP locations, there are only two possible nucleotides that occur. The nucleotide of a SNP that is more common in a population is called the major allele, and the nucleotide that is less common is called the minor allele. At conception, each individual inherits half of their DNA from the mother and half from the father. For a given SNP, one allele is transmitted from each parent. The gene, and hence the protein it produces, is affected by the genetic material received from both parents, but it does not matter which material came from which parent. Therefore, for a particular SNP, there are three possibilities: An individual has zero minor alleles, one minor allele, or two minor alleles. This number (0-2) is called the individual's genotype for this particular SNP. Finally, genotyping is the process of determining which genetic variants an individual possesses at a very large number (several millions) of SNP locations.

In essence, a GWAS consists of an atheoretical mapping of how each of several million SNPs correlates with an outcome. That is, a GWA study entails running millions of analyses in which the outcome in focus (e.g., some complex human social trait) is regressed separately on each measured SNP. Obviously this approach would result in massive multiple-testing problems if the common significance threshold of $p<0.05$ would be used. Therefore a much more stringent genome-wide significance threshold of $p<0.00000005$ is employed in a GWA study. In a well-powered study, this exercise results in a set of genome-wide significant “hits” that can then be further investigated for downstream effects and possible mechanisms. Recent studies on for example educational attainment (Okbay et al. 2016a), subjective wellbeing and happiness (Okbay et al. 2016b), reproductive behavior (Barban et al. 2016) and cognitive ability (Trampush et al. 2017) have found large numbers of genome-wide significant hits. The same studies also suggest that sample sizes of 100,000 individuals

or more are necessary in order to detect the very small effect sizes expected for the relationship between single SNPs and complex social traits.

The GWAS approach can also help us overcome the first methodological challenge mentioned above – the fact that most, if not all, previous G*E studies on complex social traits are in all likelihood dramatically underpowered. Rather than chasing often elusive effects of particular genes, a more robust approach is to use the results of GWA studies and construct so-called *polygenic scores* (PGS) for each individual. A PGS is a composite index of all genetic variants *across* the genome that aggregates their additive explanatory power. As a simplified example, consider if there are just two SNPs in the genome, for which a given individual can have either zero, one or two minor alleles. Assume, further, that previous genome-wide association studies show that each additional minor allele for the first SNP is statistically associated with a one week increase in lifetime educational duration and a half week increase for the second SNP. In this case, a PGS would simply consist of adding the genotype (that is, the number of minor alleles) for the first SNP multiplied by 1, and the genotype for the second SNP multiplied by 0.5. The single resulting number becomes a simple measure of an individual's genetic predisposition towards the number of weeks of educational attainment. Results from a GWA study, extending the example with two SNPs, can be used to construct a PGS based on the correlation between each of millions of genotyped SNPs and an outcome of interest.

Because PGS's are far more predictive of complex behavioral outcomes than single genetic variants, they can be used in many analyses that would be statistically underpowered if conducted using single variants. For example, the most recent PGS for educational attainment based on the results of an upcoming GWA study involving several of the members in this project predicts 10-11% of the variation in the level of education *out of sample*, despite the fact that each genetic variant only accounts for a miniscule part of the differences in education between individuals (Lee et al. forthcoming). This genetic effect summarizes all possible pathways (like cognitive skills, personality traits etc.).

Using PGS's as a measure of individual genetic predisposition, instead of single genetic markers, solves the problem polygenicity induces. Sample sizes required for well-powered studies of G*E effects using scores drop dramatically to the thousands or even hundreds. Further, the size of the total genetic effect becomes orders of magnitude larger, making the results substantively meaningful.

The second challenge concerns the possibility of credibly identifying *causal* G*E interactions: that is, we need a robust *identification strategy*. The key here is to find instances of exogenous environmental variation that is not correlated with genetic factors (otherwise, any observed gene-environment interaction might simply reflect an underlying gene-gene interaction) or other relevant environmental factors (possibly rendering the interaction with the environmental stimulus of interest spurious and meaningless). Here, too, most of the existing G*E literature falls short (Conley & Rauscher 2013).

To get around this problem, several experimental and quasi-experimental approaches can be drawn upon. First, we intend to utilize different policy reforms that have generated as-good-as-random variation between individuals in, among other things, educational attainment. A prime example is the introduction of comprehensive nine-year mandatory schooling in Sweden in the 1950s and 1960s.

The implementation of this reform was structured in such a way that the amount of mandatory schooling among children differed depending on municipality of residence and birth year. This makes it close to ideal as a source of environmental variation, since it is unlikely to be correlated with other environmental factors that vary over time. Previous research has shown that this comprehensive school reform affected individuals' life chances in terms of educational attainment and income (Meghir & Palme 2005), health (Lager & Torssander 2012), criminal behavior (Hjalmarsson et al. 2015), cognitive abilities (Lundborg et al. 2014), and political engagement (Lindgren et al. 2017).

There are also other types of policy processes that induce exogenous environmental variation that can be used, such as discontinuities in treatment over a continuous variable (often called a regression discontinuity design, RDD). As an example, people born right before and right after election day are not systematically different in any way other than that those born right before can vote the first time when they are 18, while those born right after will only get their first chance to vote when they are a couple of years older (depending on the length of the electoral cycle). This induces credibly exogenous variation in the onset of electoral participation that can be used to study different processes such as voting habit formation (Meredith, 2009) or trickle-up socialization from children to parents (Dahlgaard 2018). Similarly, nominal-income-threshold criteria in welfare benefit policies may induce effects on labor supply and other life outcomes.

A second approach is to use natural experiments not related to policy changes, but to randomness in natural processes. These can occur at both the contextual and the individual level. An example of as-good-as-random contextual environmental variation is weather calamities and natural disasters. These have been used to study, among other things, the impact of the Swedish 2005 storm Gudrun on party preferences (Eriksson 2016). At the individual level, the genders of one's children are random. Moreover, the gender of one's second child is known to predict the chances of having a third child and therefore becomes a credible instrument for family size that, in turn, has been shown to have downstream effects on a host of social, economic and political behavior (Black et al. 2005).

Examples of studies to be implemented within the project

Combining the individual-level polygenic scores with credible techniques for identifying the causal effects of environmental variation is in itself a major methodological advance that makes it possible to perform robust gene-environment analyses. This project will thus be able to address age-old questions with an updated and significantly more powerful toolbox.

Over the course of the project we intend to investigate whether the effects on behavioral outcomes of changes in environmental conditions induced by policy reforms and natural experiments (such as the ones mentioned above) are reduced or amplified by individual differences in genetic predispositions. The variety of outcomes we propose to study can be divided into three broad categories. The first category is education and labor market outcomes. These include for example years of education, cognitive skills, income, and employment status. Second, we will study social and political behaviors. These are, among others, personality traits, political ideology and electoral participation, pro-social and altruistic behaviors like donations and charity work, and having a criminal record. Third, we will also focus on health-related behaviors and outcomes. This category includes both behavioral variables such as tobacco and alcohol use, and health-related outcomes such as prevalence of cardiovascular disease or psychiatric illness, height, weight and life span.

A non-exhaustive list of examples of the types of questions that can be addressed using our proposed combined approach of polygenic scores and plausibly exogenous variation in environmental stimuli is:

- Do the effects of comprehensive school reforms on income, cognitive skills, health, anti-social behavior and political participation (Meghir & Palme 2005; Lager & Torssander 2012; Hjalmarsson et al. 2015; Lundborg et al. 2014; Lindgren et al. 2017) become more pronounced among those with a low or a high genetic predisposition towards education?
- Are education, labor market, social, political and health behaviors positively or negatively affected by having more siblings (Black et al. 2005), and do such effects vary depending on genetic endowments?
- Is policy preference formation of having a daughter or sister (Healy & Malhotra 2013) moderated by genetic predispositions towards certain personality traits?
- Can access to more expansive welfare benefits decrease the incidence of psychiatric illness among individuals with higher genetic risk?
- Does early voting eligibility have subsequent effects on political behavior (Meredith 2018; Dahlgaard 2018), and are such effects moderated by genetic endowments?

Data

We will use a wide variety of data sources to carry out our studies. Most importantly, we already have approval from the ethical review board in Uppsala (Dnr 2017/083) for linking a large sample of genotyped Swedish twins (N≈43000) to register data on a large set of life outcome measures such as educational attainment, labor market participation, voter turnout, political candidacy, health measures, criminal behavior and cognitive and non-cognitive abilities based on tests taken at conscription. This core sample will further be augmented with all (non-genotyped) parents, siblings and children to the genotyped twins (total N≈207000). This type of register data is very rare in international comparison, and gives us unique opportunities to contribute to the field.

Moreover, we have access to a number of sources that include both genotype data and information on relevant social, political, economic and health-related outcomes, for example the UK Biobank (N≈500000), Add Health (N≈10000), the Wisconsin Longitudinal Study (N≈10000) and the Health and Retirement Study (N≈20000). For all the individuals in these samples we can construct PGSs measuring genetic predispositions for multiple traits based on the results in available GWA studies. We will also be able to construct PGSs based on GWA studies and genetic data sources that are released during the project period.

Organization and research team

The field of social science genomics is a highly interdisciplinary and international endeavor. Reflecting this, the proposed project involves eight researchers across two different disciplines from five universities in three different countries. The eight project members have extensive experience in the social science genomics field: Sven Oskarsson (PI) (Uppsala University); David Cesarini (New York University); Christopher Dawes (New York University); Aysu Okbay (Vrije Universitet Amsterdam); Rafael Ahlskog (Uppsala University); Patrick Turley (Massachusetts General Hospital and Broad Institute); Daniel Benjamin (University of Southern California); and Magnus Johannesson

(Stockholm School of Economics). The project will last for five years and the team will to varying degrees devote their time to the project: Ahlskog 70%; Oskarsson and Okbay 50%; Cesarini and Dawes 40%; Benjamin and Turley 30%; and Johannesson 20%. The research grant is intended to fund part of the salaries for Ahlskog (50%), Okbay (20%) and Oskarsson (10%) and one US based research assistant (100%). Apart from salaries the research grant is intended to cover data purchases, team meetings and conference participation.

The applicants have been involved in several recent landmark studies on the genetic etiology of different behavioral traits published in leading journals such as *Nature* and *Science* (Rietveld et al. 2013; Okbay et al. 2016a; Okbay et al. 2016b; Turley et al. 2018; Lee et al. Forthcoming). The applicants have also published extensively on both social genomics and more traditional social science topics in leading journals in their respective fields such as *American Economic Review*, *Quarterly Journal of Economics*, *American Political Science Review*, *American Journal of Political Science*, *Psychological Science* and *Behavior Genetics*.

International collaboration

As already stressed, the field of social science genomics is by necessity a highly interdisciplinary and international endeavor. An important reason for this is that all social science genomics research has to confront the major hurdle that any true association between a complex human trait and a genetic marker is likely to be extremely small. As a result, very large samples of genotyped respondents with measures of the behavioral outcomes are needed to credibly identify genetic associations.

The medical genetics community was able to attain sufficiently large sample sizes by forming "consortia" to meta-analyze GWAS results from multiple cohorts. To enable large samples in studies of behavioral traits, the researcher active within the social genomics field followed the same research strategy by developing a research infrastructure, the Social Science Genetic Association Consortium (SSGAC). Co-applicants Daniel Benjamin and David Cesarini are co-founders of the SSGAC, and all applicants are active members of the consortium. The SSGAC was launched in 2011 and is now a fully functional research consortium. The consortium website (www.thessgac.org) contains a catalog of ongoing initiatives, updates on recent publications, and posted results (to facilitate replication and follow-up work). In addition to coordinating major GWAS efforts, the SSGAC also coordinates the collection of harmonized measures of important social science outcomes for GWAS analyses. The SSGAC has an Advisory Board comprising prominent researchers across a variety of disciplines.

As part of our collaboration with the SSGAC, we have access to data provided by several major biobanks including large samples of genotyped individuals with data on different social and economic outcomes, such as the UK Biobank. Several of the studies described in the project plan will be conducted in collaboration with an international team of researchers from a variety of disciplines connected to the SSGAC, including (i) economists Jonathan Beauchamp (University of Toronto) and Kevin Thom (New York University) (ii) statistical geneticists Peter Visscher (University of Queensland), Jian Yang (University of Queensland) and Matthew Robinson (University of Lausanne), (iii) psychologist James Lee (University of Minnesota), and (iv) molecular biologist Tõnu Esko (University of Tartu).

Dissemination plan

As discussed above the project team members have a proven track record of publishing in top-tier international scientific journals. Apart from presenting the results at conferences and workshops across different disciplines, including political science, economics and psychology, the primary academic dissemination will also be by publishing the results of the studies in peer-reviewed journals. Moreover, results will be published as open access in order to make them publically available, and by the principles of Open Science (individual level data cannot be made available since it would compromise the integrity of the research subjects, but procedures for data analysis can be).

Furthermore, we believe that the increasing use of PGS's is an important step towards integrating social-science genomics with more traditional social science research. During the grant period, many new genetic data sources will become available, enabling the construction of PGS's with meaningful predictive power for a wider range of behavioral traits. We intend to write an article outlining best practices for constructing such scores and showcasing the predictive power of PGS's for a range of social, economic and political traits. As part of the project we will also construct and make publicly available PGS's for genotyped respondents in several major social-science surveys. We believe that such a repository of PGS's has the potential to greatly simplify the inclusion of genetic factors into traditional empirical social science research, without compromising the integrity of the respondents (since no actual genetic data is shared).

We will also strive for public outreach using media contacts, press releases and social media to disseminate results of our project to the general audience. Several of the project members have extensive earlier experience of presenting social genomics research in radio, newspapers, blogs and via public seminars and lectures.

In summary, we firmly believe that social science genomics is ripe with opportunities for major advances in scientific understanding of how and why individual life course outcomes vary across the economic, political and social domains. The main contribution of this project lies in being able to perform robust G*E analyses on a wide range of outcomes, using – among other sources – unique Swedish register data. This will provide clues about how people's life outcomes are formed in a complex interplay between nature and nurture – questions that are relevant across all of the social sciences.

References

Barban, N, R Jansen, … MC Mills. (2016). "Genome-wide analysis identifies 12 loci influencing human reproductive behavior." *Nature Genetics,* 48:1462-1472.

Black, SE, PJ Devereux & KG Salvanes. (2005). "The More the Merrier? The Effect of Family Size and Birth Order on Children's Education." *Quarterly Journal of Economics*, 120(1):669-700.

Benjamin, DJ, D Cesarini, … P Lichtenstein. (2012). "The Promises and Pitfalls of Genoeconomics." *Annual Review of Economics*, 4:627-662.

Caspi, A, K Sugden, TE Moffitt, … R Poulton. (2003). "Influence of Life Stress on Depression: Moderation by a Polymorphism in the 5-HTT Gene." *Science*, 301(5631):386-389.

Chabris, CF, JJ Lee, … DI Laibson. (2015). "The Fourth Law of Behavior Genetics." *Current Directions in Psychological Science*, 24(4):304-312.

Conley, D & E Rauscher. (2013). "Genetic Interactions with Prenatal Social Environment: Effects on Academic and Behavioral Outcomes." *Journal of Health and social Behavior*, 54(1):109-127.

Culverhouse, RC, NL Saccone, AC Horton, … LJ Bierut. (2018). "Collaborative Meta_Analysis finds no Evidence of a Strong Interaction between Stress and 5-HTTLPR Genotype Contributing to the Development of Depression." *Molecular Psychiatry*, 22:133-142.

Dahlgaard, JO. (2018). "Trickle-Up Political Socialization: The Causal Effect on Turnout of Parenting a Newly Enfranchised Voter." *American Political Science Review*. Published online February 22.

Eriksson, L. (2016). "Winds of Change: Voter Blame and Storm Gudrun in the 2006 Swedish Parliamentary Election." *Electoral Studies*, 41:129-142.

Healy, A & N Malhotra. (2013). "Childhood Socialization and Political Attitudes: Evidence from a Natural Experiment." *Journal of Politics*, 75(4):1023-1037.

Hjalmarsson, R, H Holmlund & MJ Lindquist. (2015). "The Effect of Education on Criminal Convictions and Incarceration: Causal Evidence from Micro-data." *The Economic Journal*, 125(587):1290-1326.

Lager, AC & J Torssander. (2012). "Causal effect of education on mortality in a quasi-experiment on 1.2 million Swedes." *Proceedings of the National Academy of Sciences USA*, 109(22):8461-8466.

Lee, JJ, R Wedow., A Okbay, ... D Cesarini. (2018). "Gene discovery and polygenic prediction from a 1.1-million-person GWAS of educational attainment." Forthcoming in *Nature Genetics*.

Lindgren, K-O, CT Dawes & S Oskarsson. (2017). "Can Political Inequalities be Educated Away? Evidence from a Large-Scale Reform." *American Journal of Political Science* 61(1):222-236.

Lundborg, P, A Nilsson & D-O Rooth. (2014). "Parental Education and Offspring Outcomes: Evidence from the Swedish Compulsory School Reform." *American Economic Journal: Applied Economics*, 6(1):253-278.

Meghir, C & M Palme (2005). "Educational Reform, Ability and Family Background." *American Economic Review*, 95(1):414-424.

Meredith, M. (2009). "Persistence in Political Particpation." *Quarterly Journal of Political Science*, 4(3):187-209.

Okbay, A, JP Beauchamp, ... DJ Benjamin. (2016a). "Genome-wide association study identifies 74 loci associated with educational attainment." *Nature,* 533:539–542.

Okbay, A, BM Baselmans, … D Cesarini. (2016b). "Genetic variants associated with subjective well-being, depressive symptoms, and neuroticism identified through genome-wide analyses." *Nature Genetics,* 48:624-633.

Pinker, S. (2003) *The Blank Slate: The Modern Denial of Human Nature*. Penguin Press.

Polderman, TJC, B Benyamin, ... D Posthuma. (2015). "Meta-analysis of the heritability of human traits based on fifty years of twin studies." *Nature Genetics,* 47:702-709.

Rietveld, N, S Medland, … P Koellinger. (2013). "GWAS of 126,559 Individuals Identifies Genetic Variants Associated with Educational Attainment." *Science*, 340(6139):1467-1471.

Spinath, FM & W Bleidorn. (2017). "The New Look of Behavioral Genetics in Social Inequality: Gene-Environment Interplay and Life Chances." *Journal of Personality*, 85(1):5-9.

Trampush, JW, MLZ Yang, … T Lencz. (2017). "GWAS meta-analysis reveals novel loci and genetic correlates for general cognitive function: a report from the COGENT consortium." *Molecular Psychiatry,* 22:336-345.

Turley, P, RK Walters, O Maghzian, A Okbay, JJ Lee, ... DJ Benjamin. (2018). "Multi-trait analysis of genome-wide association summary statistics using MTAG." *Nature Genetics*. Published online January 1.