

This electronic thesis or dissertation has been downloaded from the King's Research Portal at <https://kclpure.kcl.ac.uk/portal/>



Leveraging family data for complex trait genomics

Cheesman, Rosa

Awarding institution:
King's College London

The copyright of this thesis rests with the author and no quotation from it or information derived from it may be published without proper acknowledgement.

END USER LICENCE AGREEMENT



Unless another licence is stated on the immediately following page this work is licensed

under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International

licence. <https://creativecommons.org/licenses/by-nc-nd/4.0/>

You are free to copy, distribute and transmit the work

Under the following conditions:

- Attribution: You must attribute the work in the manner specified by the author (but not in any way that suggests that they endorse you or your use of the work).
- Non Commercial: You may not use this work for commercial purposes.
- No Derivative Works - You may not alter, transform, or build upon this work.

Any of these conditions can be waived if you receive permission from the author. Your fair dealings and other rights are in no way affected by the above.

Take down policy

If you believe that this document breaches copyright please contact librarypure@kcl.ac.uk providing details, and we will remove access to the work immediately and investigate your claim.

Leveraging family data for complex trait genomics

Rosa Cheesman

PhD Thesis, March 2020

Social, Genetic, and Developmental Psychiatry Centre
King's College London

Submitted for the degree of
Doctor of Philosophy in Behavioural Genetics

Abstract

Fundamental questions about the roles of genes, environments, and their interplay in the development of complex behaviours have traditionally been the domain of twin and family study designs. More recently, the rapidly growing availability of large genomic datasets, comprised of unrelated individuals, has generated exciting novel insights. However, there are major stumbling blocks. It has been difficult to capture the full heritability of complex traits, as estimated from family data, with measured DNA. Moreover, the genetic influence identified using DNA may be confounded with environmental effects. A deeper integration of genomic research with family-based quantitative genetics helps to address these issues and push knowledge further.

This thesis includes four empirical chapters. Each increases the accuracy of findings in complex trait genomics by using cutting-edge methods in combination with insights and data from twin and family studies. The first chapter demonstrates the utility of longitudinal twin data for defining a measure of emotional problems. This measure is more heritable and powerful for genomic research than individual time-specific assessments. The second chapter presents evidence that genomic data from families in the UK Biobank can be leveraged to capture additional genetic effects on neuroticism and educational attainment. The third chapter uses an adoption design to measure how polygenic scores for educational attainment pick up environmental effects provided by relatives. The fourth chapter uses genomic data on parent-offspring trios to show that environmental effects of parents on childhood emotional symptoms can be indexed using the parental genome. Limitations of these approaches, potential directions for future research, and overarching conclusions are presented in the final chapter.

Overall, this thesis demonstrates that family-based quantitative genetic research is not only the foundation and justification for complex trait genomics. It is of enduring importance, especially when integrated with new genomic tools. Family-based genetic research charts the direction for complex trait genomics.

Table of Contents

Abstract	2
Table of Contents	3
Table of Figures	6
Table of Tables	7
Acknowledgements	8
Statement of authorship	9
Publications resulting from chapters of this thesis	10
Other publications	10
Chapter 1. Background	11
Introduction	11
A brief history of complex trait genetics	11
Challenges in genomic research on complex traits	15
Missing heritability	15
Family environmental bias	18
Aims and structure of the thesis	23
Improving measurement (Chapter 2)	23
Capturing rarer genetic effects (Chapter 3)	23
Removing environmental bias (Chapter 4)	23
Addressing long-standing questions about the environment (Chapter 5)	24
References	25
Chapter 2. Extracting stability increases the SNP heritability of emotional problems in young people	27
Abstract	28
Introduction	28
Materials and methods	29
Sample	30
Genotyping	30
Measures	30
Statistical analyses	30
Code availability	32
Results	32
Latent trait analyses	32
Heritability analyses	32
Polygenic scoring	33
Genome-wide association	34
Discussion	34
References	35
Chapter 3. Familial Influences on Neuroticism and Education in the UK Biobank	37
Abstract	38
Introduction	38
Methods	40

Sample	40
Genotyping	40
Measures	40
Analyses.....	40
Results.....	42
Identification of family members.....	42
Kin-based SNP heritability method accounting for environmental similarity: GREML-KIN	42
GREML-LDMS-I to investigate the effects of less common variants	43
Discussion	44
References	46
Chapter 4. Comparison of adopted and non-adopted individuals reveals gene-environment interplay for education in the UK Biobank.....	48
Abstract.....	49
Introduction	49
Method.....	50
Sample, genotype quality control and phenotype definition.....	50
Statistical analyses.....	50
Results.....	52
Sample analysed.....	52
Phenotypic results.....	53
SNP heritability estimates	53
Polygenic prediction results	53
Supplemental analyses	54
Discussion	55
References	57
Chapter 5. How important are parents in the development of child anxiety and depression? A genomic analysis of parent-offspring trios in the Norwegian Mother, Father and Child Cohort Study.....	59
Abstract.....	60
Introduction	61
Methods	64
Sample	64
Measures	64
Genotype quality control	65
Statistical analyses.....	65
Software	68
Results.....	69
RDR model.....	69
Pedigree model	69
Discussion	71
References	74
Chapter 6. Discussion	78
Overview of findings	78
Chapter-specific findings	79

Conclusions about cross-cutting themes	80
Missing heritability of complex traits.....	80
Gene-environment interplay and causality.....	82
General limitations	85
Comparability of results across chapters	85
Generalisability of findings	85
Unknown quantities in genomics.....	87
Future directions	89
Estimating various environmental effects and disentangling biases.....	89
Estimating how individuals shape their experiences.....	90
Investigating gene-environment interactions.....	90
Answering longitudinal multivariate questions with family genomic data	91
Conclusions	91
References	92
Appendices.....	95
Appendix 1 – Supplementary Information for Chapter 2.....	96
Appendix 2 – Supplementary Information for Chapter 3.....	120
Appendix 3 – Supplementary Information for Chapter 4.....	125
Appendix 4 – Supplementary Information for Chapter 5.....	135

Table of Figures

Chapter 1. Background

Figure 1. Methods for heritability estimation take advantage of different relatedness relationships..	15
Figure 2. Genetic variation can have direct and indirect effects on an individual	19
Figure 3. Differences between concepts in gene-environment interplay research	21
Figure 4. Overview of factors influencing the twin-SNP heritability gap	22

Chapter 2. Extracting stability increases the SNP heritability of emotional problems in young people

Figure 1. Workflow diagram of current analyses	30
Figure 2. A simplified diagram of our CFA model	31
Figure 3. Twin and SNP heritabilities of the CFA- and IRT-derived scores for stable emotional problems, and of the 12 individual measures	33

Chapter 3. Familial Influences on Neuroticism and Education in the UK Biobank

Figure 1. Kinship plotted against IBS0 for all first-degree relatives in the UK Biobank. Blue = siblings; yellow = parent– offspring pairs	43
Figure 2. Variance component estimates for neuroticism and education	44

Chapter 4. Comparison of adopted and non-adopted individuals reveals gene-environment interplay for education in the UK Biobank

Figure 1. Estimates of the variance explained by common SNPs for years of education, and of the variance explained by polygenic scores for education polygenic scores, in adoptees compared to individuals who were reared with their relatives	54
Figure 2. Regression of years of education on polygenic score for education, comparing 6311 adoptees to a sample of 6,500 non-adopted individuals	54
Figure 3. Mean education years (standardised) per decile of polygenic score for education years (also standardised), for adopted and non-adopted individuals	55

Chapter 5. How important are parents in the development of child anxiety and depression? A genomic analysis of parent-offspring trios in the Norwegian Mother Father and Child Cohort Study

Figure 1. Variance component estimates from the RDR model	70
Figure 2. Pedigree-based variance component estimates	70

Table of Tables

Chapter 2. Extracting stability increases the SNP heritability of emotional problems in young people

Table 1. Proportion of phenotypic variance predicted in stable emotional problems, and in 12 anxiety and depression measures, by polygenic scores for Major Depressive Disorder and for General Anxiety Disorder	33
--	----

Chapter 3. Familial Influences on Neuroticism and Education in the UK Biobank

Table 1. Sample sizes for different family relationships in the UK Biobank.....	43
Table 2. Results of GREML-LDMS-I variance components analyses for neuroticism and education using six minor allele frequency and LD bins	44

Chapter 4. Comparison of adopted and non-adopted individuals reveals gene-environment interplay for education in the UK Biobank

Table 1. Comparative analysis of phenotypes in adoptees versus non-adopted individuals.....	53
---	----

Acknowledgements

First and foremost, thank you to Professors Thalia Eley and Robert Plomin, who have been wonderful supervisors. I have loved being part of the EDIT lab and the TEDS team. Your enthusiasm and good humour helped me to feel at home in both families. Thalia, you have anchored and guided my work, whilst encouraging me to pursue ideas. Your insights have been invaluable. Robert, you have been very intellectually generous with me throughout my PhD. I have benefited enormously from our weekly discussions.

I am grateful to have been based in the Social, Genetic, & Developmental Psychiatry Centre for 4 years, and am thankful to the students and staff who make it a great place to be, both intellectually and socially. Special thanks go to those who have given a lot of their time to work with me: Tom McAdams, Gerome Breen, Joni Coleman, Jean-Baptiste Pingault, and Eivind Ystrom. I would like to acknowledge the brilliant TEDS team, including Kaili Rimfeld, Nic Shakeshaft, Margherita Malanchini and the rest of the 'CFG', who have been my role models, and of course Rachel Ogden and Andy McMillan, who keep everything running smoothly and never made me feel like the token TEDS twin.

My PhD journey was very much shared with Chris Rayner. Big thanks to you for being there to support me, share mince pies with me, and make me laugh constantly through highs and lows. Thanks also to Yasmin Ahmadzadeh for being a ray of sunshine and for telling me to do a PhD in the first place. Andrea Allegrini and Tom McGregor, I'll miss sharing the desk island with you (along with the coffee and holes in the floor).

Additionally, this would not have been possible without the mentorship of my Human Sciences tutors Professor Maggie Snowling and Dr Rosalind Harding, who nurtured my interests in gene-environment interplay and encouraged me to apply to the SGDP.

Huge thanks to Riccardo Marioni and Nic Timpson for being great examiners. I'll be reflecting on our discussion for a long time. Your insights will greatly strengthen my work going forward.

Finally, thank you to Mum, Dad and Margie. Your love and support saw me through the final stages of writing, in the midst of the coronavirus pandemic. This is dedicated to you.

Statement of authorship

All work presented in the thesis is my own, except where acknowledged in the text. Data collection for all study samples used herein was completed by research teams prior to my involvement. All investigations were conceived and carried out by me, as first author, in collaboration with colleagues included in the author lists at the start of each chapter. Chapter-specific author contribution summaries are shown below.



Rosa Cheesman

Chapter 2. R.C and T.C.E conceived and designed the investigation; R.C carried out statistical analyses, wrote and revised the manuscript; all co-authors critically reviewed the manuscript.

Chapter 3. R.C conceived and designed the investigation, carried out the genotype data quality control and analysis, wrote and revised the manuscript; R.C discussed results with T.C.E, G.B, J.R.I.C, K.L.P and C.R; all co-authors provided comments on manuscript drafts.

Chapter 4. R.C and G.B conceived and designed the study. R.C analysed the data and wrote and revised the manuscript. A.H and J.R.I.C conducted genotype data quality control. All co-authors provided critical revisions and approved the final version of the manuscript.

Chapter 5. R.C conceived and designed the investigation, carried out genotype data quality control and analysis, and wrote and revised the manuscript; all co-authors discussed results, and provided comments on manuscript drafts.

Publications resulting from chapters of this thesis

Published:

Cheesman, R., Major Depressive Disorder Working Group of the Psychiatric Genomics Consortium, Purves, K.L., Pingault, J-B., Breen, G., Rijdsdijk, F., Plomin, R. and Eley, T.C. 2018. Extracting stability increases the SNP heritability of emotional problems in young people. *Translational psychiatry* 8(1), p. 223.

Cheesman, R., Coleman, J., Rayner, C., Purves, K.L., Morneau-Vaillancourt, G., Glanville, K., Choi, S.W., Breen, G. and Eley, T.C. 2019. Familial influences on neuroticism and education in the UK biobank. *Behavior Genetics*. doi:10.1007/s10519-019-09984-5

Accepted/in press:

Cheesman, R., Hunjan, A., Coleman, J., Ahmadzadeh, Y., Plomin, R., McAdams, T., Eley, T.C. Breen, G. 2019. Comparison of adopted and non-adopted individuals reveals gene-environment interplay for education in the UK Biobank. In press *Psychological Science*.

In preparation:

Cheesman, R., Moen Eilertsen, E., Ahmadzadeh, Y.I., Gjerd, L.C., Hannigan, L.J., Havdahl, A., Young, A.I., Eley, T.C., Ystrom E., McAdams, T.A. How important are parents in the development of child anxiety and depression? A genomic analysis of parent-offspring trios in the Norwegian Mother Father and Child Cohort Study.

Other publications

The following is a subset of relevant papers that have been published or accepted for publication over the course of the PhD research:

Cheesman, R., Selzam, S., Ronald, A., Dale, P. S., McAdams, T. A., Eley, T.C., Plomin, R. J. 2017. Childhood behaviour problems show the greatest gap between DNA-based and twin heritability. *Translational Psychiatry*. 7(12), p.1284.

Allegrini, A.G. *, Cheesman, R.*, Rimfeld, K., et al. 2020. The p factor: Genetic analyses support a general dimension of psychopathology in childhood and adolescence. *JCPP* (*=joint first authors). <https://doi.org/10.1111/jcpp.13113>

Murphy, J.*, Cheesman, R.*, Gregory, A., Lau, J., Ehlers, A., Catmur, C., Bird, G. and Eley, T. Estimating the stability of heartbeat counting in middle childhood: a twin study. 2019. *Biological psychology* (*=joint first authors), 148, p.107764

Cheesman, R., Rayner, C., Eley, T. C. 2019. The genetic basis of child and adolescent anxiety. *Pediatric Anxiety Disorders*. Ed. Compton. Elsevier.

Purves, K.L., Coleman, J.R.I., Meier, S.M., Rayner, C., Davis, K.A.S., Cheesman, R., ..., Hettema, J.M., Hotopf, M., Hougaard, D., Hübel, C., Kan, C., McIntosh, A.M., Mors, O., Bo Mortensen, P., Nordentoft, M., Werge, T., Nicodemus, K.K., Mattheisen, M., Breen, G. and Eley, T.C. 2019. A major role for common genetic variation in anxiety disorders. *Molecular Psychiatry*. doi:10.1038/s41380-019-0559-1

Rayner, C., Coleman, J.R.I., Purves, K.L., Cheesman, R., Huebel, C., Gaspar, H., Glanville, K., Krebs, G., Morneau-Vaillancourt, G., Breen, G., Eley, T.C. 2019. Genetic influences on treatment-seeking for common mental health problems in the UK Biobank. *Behaviour Research and Therapy*, p.103413.

Morneau-Vaillancourt, G., ... Cheesman, R., ... Eley, T.C. 2020. The genetic and environmental hierarchical structure of anxiety and depression in the UK Biobank. *Depression and Anxiety*.

Chapter 1. Background

Introduction

Over the past two decades, major advances have been made in identifying the molecular genetic basis of complex human traits. This progress has been achieved by applying genomic tools to large samples of unrelated individuals. However, several issues stand in the way of further developments using this population approach. In light of the current barriers, it is timely to return to insights and data structures that have served the field of quantitative genetics since the early 20th century. Well before genetic variation could be directly measured, resemblances between twins and families were enabling discoveries about nature, nurture and their interplay. This chapter sets forth the historical context of complex trait genetics, from traditional quantitative genetic, to newer genomic methods, and the divide between the two. Then, current challenges in genomics are introduced, and the advantages of integrating twin and family approaches into the genomic era are described. The thesis calls for a deeper synthesis of the ‘two worlds’ of genetics, to generate more accurate and novel findings.

A brief history of complex trait genetics

Quantitative genetics

To what extent do genetics (‘nature’) or environment (including parental ‘nurture’) make us who we are? In pursuing answers to this enduring question, the observation of resemblance between children and their parents has been of central concern. Francis Galton, who popularised the phrase ‘nature and nurture’ in its contemporary sense, invented regression specifically to gain an understanding of familial resemblance. It was not until Ronald Fisher’s landmark contribution in 1918 that genetic theory could explain the origins of familial resemblance and trait variation in the population (Fisher 1918). Indeed, the roots of modern quantitative genetics – the study of genetic and environmental contributions to phenotypic variance and covariance in a population – can be traced to his paper, which reconciled an ongoing conflict between Mendelians and biometricians (Fisher 1918). It had seemed that the complex quantitative traits studied by biometricians – characteristics with underlying continuous dimensions such as human height – could not result from the patterns of inheritance that applied to discrete Mendelian traits. Fisher showed that a complex quantitative trait could be explained by Mendelian inheritance if many genes affect the trait (i.e., polygenicity) (Plomin et al. 2009). Fisher's theory implied that genetic and environmental sources of variation could be partitioned by quantifying phenotypic correlations between known relatives, without knowledge of specific genes affecting the trait (Vinkhuyzen et al. 2013). The central concept of ‘heritability’ or ‘genetic variance’, meaning the proportion of population variance attributable to genetic factors, was thus in effect defined by Fisher.

Since the 1930s, thousands of heritability estimates for complex human traits have been obtained by contrasting trait resemblance between family members of varying degrees of genetic relatedness, such as monozygotic and dizygotic twins, extended pedigrees, and adoptees. In particular, the natural experiment provided by the existence of two types of twins has offered an elegantly simple method for investigating the relative roles of nature and nurture. Because of this, twin studies have had a profound impact on our understanding of what makes us who we are. The logic behind this design is

that resemblance in monozygotic pairs results from them sharing all their genetic (A) and shared environmental (C) factors:

$$(1) r_{MZ} = A + C.$$

Note: r =correlation

In contrast, resemblance in dizygotic pairs is due to them sharing half their genetic variance in addition to the shared environment:

$$(2) r_{DZ} = 1/2A + C.$$

It follows that heritability can be estimated as twice the difference between monozygotic and dizygotic correlations:

$$(3) A = 2(r_{MZ} - r_{DZ})$$

Shared environment, representing environmental influences that result in family members resembling one another, is calculated as the monozygotic twin correlation minus the heritability of the trait:

$$(4) C = r_{MZ} - A$$

The remaining variance not explained by A or C loads on the non-shared environmental component (E), representing influences that make family members different from one another. This might include illness or accidents, but also randomness and measurement error.

Notably, the above equations pertain to an additive genetic model, where the cumulative effect of all variants is assumed to be approximated by their sum. Twin model-fitting can be used to estimate both additive and non-additive components of variance. However, it is not possible to simultaneously estimate three components – additive, non-additive and shared environmental variance – from only two pieces of relevant information (the MZ and DZ covariances). Accordingly, most applications employ the ACE model rather than a non-additive model. Modelling typically relies on structural equation modelling approaches, rather than multiplication and subtraction.

Heritability estimates vary widely but for most complex traits, twin-based estimates are moderate to high (~0.4-0.8) (Polderman et al. 2015). The ubiquitous heritability has been called the ‘first law of behaviour genetics’ (Turkheimer 2000). The field of quantitative genetics has by no means been limited to what might now seem a ‘rudimentary’ task of estimating heritability (Haworth and Plomin 2010). Extensions of the classical univariate twin model have allowed hypotheses to be addressed about genetic and environmental influences and their interplay across the life-course, across generations, traits and contexts. One well-replicated example is the finding from longitudinal twin studies that genetic factors drive stability in symptoms of depression and anxiety across ages, whereas environmental factors explain change (Nivard et al. 2015; Hannigan et al. 2017).

There are important nuances about heritability to note. First, traits have no ‘true’ heritability values. Heritability is not a fixed statistic but depends on genetic and non-genetic causal factors acting in the sample of the study. Genetic and environmental variance components may vary systematically across

different social contexts and policy environments (Jencks 1980). Therefore, researchers tend to emphasise the patterns that emerge from many different studies, rather than specific point estimates derived from individual samples. Second, heritability estimates from twin, pedigree, and adoption studies may be affected by the strong assumptions necessary for employing these approaches. For example, violations of the equal environment assumption of the classical twin design, that monozygotic and dizygotic twins have equally similar family environments, may lead to overestimation of heritability.

The assumptions of models for decomposing variance into genetic and environmental components, and the consequences of these assumptions, are discussed in detail in the empirical chapters of this thesis, particularly Chapter 2.

Genomics

With the advent of increasingly cheap genotyping technologies and characterisation of the haplotype structure of the genome (International HapMap Consortium 2003), the application of indirect statistical methods to families was no longer the exclusive domain for studying nature and nurture. In the genomic era, researchers could directly measure genome-wide DNA variation and its effects in samples of thousands of individuals. Attention in genetics turned towards scanning the genome for measured variations between people that accounted for the heritability estimated in family-based quantitative genetic designs.

Over the past two decades, the genome-wide association study (GWAS) has offered a powerful and hypothesis-free way to test for specific genetic variants that are associated with quantitative variation (if the phenotype is a dimensional trait) or are more common amongst cases than controls (in clinical samples). The main kind of genetic variation surveyed through GWAS is the single nucleotide polymorphism (SNP). These single base pair changes are the most frequent form of human genetic variation. GWAS usually assess SNPs that are common in the population, in that they have a minor allele frequency (MAF) above 1%. A notable feature of GWAS is the typically vast scale of samples (tens, hundreds of thousands) required to achieve power to detect individual SNP effects. This is largely due to the highly conservative significance threshold (of $5e-8$, to correct for testing of 1 million independent SNP associations), together with small individual SNP effect sizes (which follows from Fisher's polygenic or 'infinitesimal' model).

Would GWAS supplant or re-invigorate quantitative genetic research? On one hand, the essence of the GWAS paradigm was informed by quantitative genetics. The focus on many common genetic variants, and the implementation of GWAS as a linear regression of a trait on SNP dosage (0, 1, or 2 of an allele), match Fisher's derivation of genetic variation (Visscher and Goddard 2019).

However, for valid reasons, genomics research largely left behind at least three aspects of classical quantitative genetics: methods, samples and insights. First, regarding **methods**, if we could find the specific genetic variants responsible for heritability of complex traits, there would perhaps be no need for traditional family-based methods, because genetic influence could now be assessed directly rather than implied. In many ways, genomic methods (GWAS and post-GWAS) push insights further than

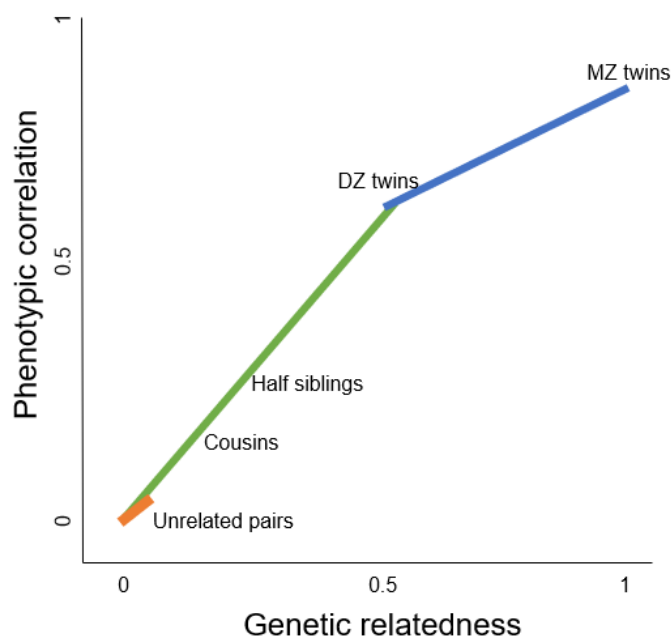
quantitative genetic methods can. For example, they can investigate specific genes involved in traits, how many there are, their individual and interactive effects. Second, genomic approaches demanded the development of population-based **samples** excluding all relatives, whether known (pedigree data) or unknown (cryptic relatedness). This has the effect of controlling for bias from environmental influences shared by more genetically similar individuals, and to ensure independence of participants' genotypes, as assumed in GWAS. Third, twin studies revealed relevant **insights** that are seldom used in genomics. For example, since the power of a GWAS partially depends on the sheer heritability of the trait, GWAS phenotypes could be optimised according to twin-based evidence on the most heritable measure, age, and rater. However, the importance of sample size, and the cost of having to collect new phenotype data in addition to genotyping, has led to the neglect of such insights in favour of somewhat crude phenotypes that are available in many existing datasets, such as years of education, which is a standard demographic variable (Lee et al. 2018).

The population-level GWAS approach has been successful for complex traits, with an astonishing pace of discovery of tens of thousands of robust SNP-trait associations (Visscher et al. 2017; Claussnitzer et al. 2020). Insights about traits from GWAS have been wide-ranging, indicating what genes are involved (meta-analysis / pathway), how they work beyond the sequence (epigenetics, transcriptomics, proteomics), and what can we do with them (translational genomics). Of the many complex psychiatric phenotypes to which GWAS approach has been applied, one poster child is schizophrenia, a highly heritable disorder involving cognitive and perceptual impairments. Over 12 years, community consortium efforts allowed expansion from 320 cases and 325 controls (Mah et al. 2006) to 40,675 cases and 64,643 controls in the latest meta-analysis, in which 179 independent SNP associations were identified (Pardiñas et al. 2018). Associated genes have yielded important insights into pathophysiological mechanisms in schizophrenia. In 2016, a schizophrenia-associated genetic variant in the complement C4 gene was found to increase expression of protein C4A in brains of mice, where it affects synaptic pruning. The resulting synaptic dysfunction could be the consequence of this excessive pruning leading to the development of schizophrenia (Sekar et al. 2016). Given that common diseases are not simply aggregations of related Mendelian conditions but are highly polygenic (Plomin et al. 2009) it follows that not only the significantly associated 'hits' contain useful information. Polygenic scores, which index the genetic liability that individuals carry for a specific trait by aggregating genetic risk across thousands of genetic variants, not only the significant SNPs, were first applied to show that schizophrenia genetic risk is associated with bipolar disorder (International Schizophrenia Consortium et al. 2009). Since then, these polygenic scores have been used to address pressing questions, for example about early manifestations of genetic risk for schizophrenia (Allegrini et al. 2020), even when schizophrenia is not assessed in the target population. None of this is possible using traditional family-based quantitative genetic designs.

Importantly for this thesis, the genomic era also brought new methods to estimate heritability, which no longer relied on comparisons of correlations between types of family members. The foremost such method – Genomic-RElatedness-based restricted Maximum-Likelihood (GREML), which is implemented in the Genome-wide Complex Trait Analysis toolkit (GCTA) (Yang et al. 2011) – takes

advantage of the random genetic differences that exist between unrelated individuals in a population. The GREML approach involves estimating the extent that genome-wide genetic relatedness among unrelated individuals can predict their trait similarity. Because genome-wide genetic relatedness is calculated using directly genotyped SNPs, the GREML method effectively adds up the variance explained by SNPs captured on the array (and those correlated with measured SNPs), known as **SNP-based heritability**. This restriction to common SNPs is significant because it also applies to genome-wide association studies. GREML SNP heritability therefore reveals the upper limit on how much variance in a trait can be captured using measured common genetic variation – i.e. the ceiling for success of a well-powered GWAS of a given phenotype in a given population. GREML estimates of genetic influence are expected to be lower than twin study-based heritability, mainly because GREML only detects the effects of common genetic variants, whereas the twin design picks up effects of any DNA variation. Figure 1 shows how GREML (yellow) capitalises on small variations in genomic similarity among unrelated individuals, whereas the classical twin design (blue) uses information on twins who share 100% or 50% of their genes. The third method (green) is discussed in Chapter 3.

Figure 1. Methods for heritability estimation take advantage of different relatedness relationships. Blue line= classical twin design; green = GREML-KIN method; orange = single component GREML. Phenotypic correlations reflect an imaginary trait similar to educational attainment with a high MZ correlation of ~0.80



Challenges in genomic research on complex traits

There are many important and interesting challenges at the cutting edge of complex trait genomics (e.g. understanding the biological pathways of genome-wide significant SNP associations). This section does not attempt to cover all of these, but introduces **two key dilemmas** – missing heritability and family environmental bias. These are tackled with empirical data in the main body of the thesis.

Missing heritability

Despite the overarching success of the GWAS paradigm, it has proven difficult to capture the full magnitude of twin and pedigree-based genetic influence on complex traits with directly measured

genetic variation – known as the ‘missing heritability’ problem (Manolio et al. 2009). Vast efforts are required to obtain individual genetic associations, reflecting the fact that complex traits are shaped by many common genetic variants with small effects, which require very large GWAS samples to find. Not only do individual hits explain tiny fractions of phenotypic variance (missing GWAS heritability) – the aggregate variance explained by common SNPs, as estimated using GREML, explains only a small proportion of familial clustering and is often less than half the size of twin heritability (missing SNP heritability – the sense in which the term ‘missing heritability’ will be used hereafter). The first application of GREML, to understand variation in height, revealed substantial missing heritability. Although SNP heritability exceeded the variance explained by genome-wide significant SNPs (suggesting room for GWAS success), it was much lower than pedigree estimates (Yang et al. 2010).

In a study aiming to gauge the breadth of the missing heritability for anthropometric, cognitive, and mental health traits in childhood and adolescence, we obtained directly comparable GREML and twin heritability estimates for the same measures in the same sample – the UK Twins Early Development Study (Cheesman et al. 2017). Particularly for behaviour problems, we found surprisingly low SNP heritability estimates (average 6%; non-significant) in comparison to twin heritabilities (average 52%). This is in line with SNP heritability estimates for related adult mental health traits such as neuroticism (also 6%) (Vinkhuyzen et al. 2012). For anthropomorphic and cognitive measures in TEDS, SNP-based estimates were low but closer to expectation – half those of twin-based estimates. If SNP heritabilities are truly this low, then GWAS, *particularly* of childhood emotional and behaviour problems, will struggle to find variants involved in their aetiology and to generate predictive polygenic scores. It may be fruitless to conduct and interpret such studies if the overall signal from common variation in the genome is negligible.

Possible explanations for missing heritability

There are multiple known factors that influence why SNP heritabilities are so low compared with twin estimates. The following section introduces the factors in turn, while considering their utility as targets to maximise SNP heritability for GWAS. Notably, the severity of the missing heritability problem, and the importance of the individual factors in explaining the heritability gap, differs across traits.

Effects of untagged and rarer variants

A large part of the problem is that SNP heritability estimates capture a subset of the true genetic variation present across a population. This is because genotyping platforms are unable to measure rare variants or de novo mutations. In contrast, the twin design estimates a latent effect of all DNA sequence differences. The continuing divide between efforts targeting rare variants with large effects and genome-wide searches for SNPs influencing complex traits seems to echo the historical debate between Mendelian and biometric views of human genetics (Claussnitzer et al. 2020). Future identification of significant DNA-based heritability will rely on further synthesis of these perspectives. Whole genome sequencing offers an excellent opportunity for recovering complex trait heritability with measured DNA (Wainschtein et al. 2019), but until it becomes less expensive, other approaches can help to pick up effects of alleles across the frequency spectrum. This is the focus of Chapter 3.

Phenotypic measurement issues

Phenotypic measurement is important because it determines the amount of reliable variance (relative to error) to be partitioned into genetic and environmental sources. The vast SNP-twin heritability gap for mental health but not for objectively assessable traits like height, seen in TEDS and other datasets, may be because psychological traits are especially hard to measure reliably. Improved measurement will not narrow the gap between twin and SNP heritability where they are estimated in the same sample. However, anything that increases SNP heritability by creating better measures or by improving the *use* of measures will help build towards gene-finding and prediction. This is the rationale for Chapter 2.

Other factors

First, low **sample size** partially explains the heritability gap because much more statistical power is required for SNP than for twin heritability, in order to detect small effects of common DNA variants. Standard errors are always larger for SNP heritability than for twin estimates because GREML is based on comparisons of small pairwise genetic differences amongst unrelated individuals compared to the genetic difference of 100 vs. 50% for monozygotic and dizygotic twins. Put simply, it is more difficult to precisely estimate differences between two correlations if they are small correlations. As GWAS samples continue to grow, this is becoming less and less of a problem.

Second, violation of twin study assumptions about the environment may lead to **overestimation of heritability**, such that less heritability may in fact be missing. Extended twin family designs can relax some of the assumptions of the classical twin design. For some traits but not others, this provides evidence that twin heritability is inflated (Coventry and Keller 2005).

Third, an important source of trait variation not captured by GREML is **non-additive genetic effects**. SNP heritability estimates only detect *additive* effects of causal variants tagged by common SNPs on current DNA arrays. Although it is assumed that the combined effect of alleles at a locus can be modelled as the sum of their individual effects, alleles may not operate in this linear fashion. One allele could mask the effect of another (dominance), and there may also be interactions at the gene level (epistasis).

Fourth, **gene-environment interactions**, whereby genetic effects depend on environmental context, complicate matters further. The idea is intuitive and likely to occur, but results are mixed and not well-replicated. Gene-environment interactions studies are practically difficult when reliant upon measurement of both genes and environments. Also, results are difficult to interpret because different types of interactions affect different heritability methods in different ways. For example, in the twin design, a gene-by-shared environment interaction may inflate heritability estimates (genes make MZ twins more similar than DZ twins; shared environmental factors make both sets of twins more similar, so the interaction may make MZ twins even more similar). Conversely, a gene-by-unique-environment interaction may deflate heritability estimates relative to both shared and nonshared environmental variance (genes moderate the effect of a factor that makes individuals more different from one another; MZ twins share more genes so may be made even more different from one another by that

factor). For SNP heritability estimates, gene-by-shared environment interactions are unlikely to contribute because the unrelated participants are unlikely to share an environment. It is possible that both gene-by-environment interactions and non-additivity make genomic research more difficult and lead to overestimation of twin heritability, making both phenomena plausible candidates for explaining some of the ‘missing heritability’.

Summary

Heritability is at the heart of genomic research and is the benchmark against which genetic association studies and genetic predictors of complex human traits are measured. The accuracy of our benchmark is therefore of vital importance. Missing heritability poses a threat to the statistical power and utility of modern genomic designs. Luckily, there are several tractable options to capture more genomic signal.

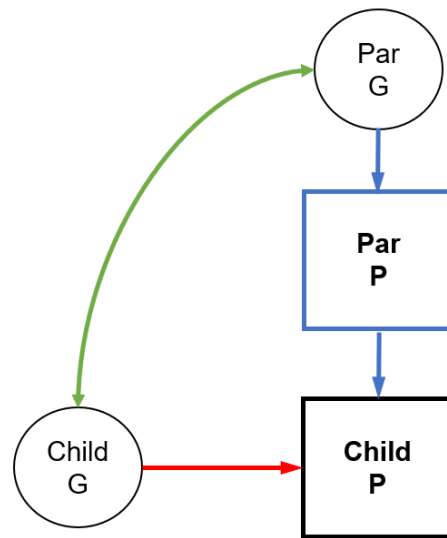
Family environmental bias

The previous section dealt with a central challenge in complex trait genomics: missing heritability. A second key issue, which has come to the fore more recently, is the potential biasing of estimates of genetic influence by environmentally-mediated effects of relatives’ genomes.

Associations between individuals’ DNA and their traits may not simply reflect genetic ‘propensity’ originating in individuals’ own genotypes, but may also pick up environmentally mediated effects coming from their genetic relatives, particularly parents. This occurs because, in typical families, children receive both their genome and their rearing environment from their parents. Consider the example of educational attainment using Figure 2. First, genes inherited by the child have **direct** effects on their own traits, such as learning and education (red arrow). Second, this genetic variation is correlated with parental genetic variation (*green double headed arrow in Fig. 2 represents the genetic correlation*), which may affect the child **indirectly**, for example by contributing to parental education, and to influential aspects of the home environment of the child such as reading books together (blue arrow). Importantly, genetic relatedness is entangled with environmental sharing, so direct and indirect genetic contributions to phenotypic similarity are difficult to separate. Without appropriate methods to take parent genetic effects into account, estimates of child genetic effects will be a combination of the correlated indirect parental effect and the within-individual effect.

It is essential to note that the within-individual direct genetic effect is only direct in the sense that it does not stem from another individual’s genotype. It is also not necessarily ‘purely’ genetic. Especially for complex behaviours, genotypes lead to phenotypes via intermediate pathways, and are expressed in the context of environments. For example, a locus that shows a replicable association with lung cancer risk (CHRNA5-A3-B4) operates largely via quantity of tobacco exposure (Gage et al. 2016).

Figure 2. Genetic variation can have direct and indirect effects on an individual. Genes do not only influence an individual's phenotype because they are inherited and operate directly (red arrow). This variation is present in the parents (green double headed arrow), and may affect how parents shape the rearing environment (indirect genetic effect; blue path). Note: square=observed measure, circle=latent factor, G= genetics, P=phenotype, Par=parents



In 2018, two influential studies provided striking demonstrations of indirect genetic effects on educational attainment (Kong et al. 2018; Bates et al. 2018). Using genomic data on parent-child trios, they showed that education-linked genetic variation *that parents did not pass on to their child*, was still predictive of child attainment. Although this result concerns a polygenic score, it constitutes an environmental effect that the authors call **genetic nurture**. Alleles that were not transmitted to offspring can only take effect *indirectly* via heritable parental behaviour that children are exposed to, not via direct inheritance of DNA. This approach enables the parent genetic effect (the blue path in Figure 2) to be isolated over and above the effect of transmitted genes (green and red paths in Figure 2). Importantly, Kong et al. showed that polygenic score associations estimated without including parental genotype data may pick up effects of unobserved parental genomes, and so cannot be taken to represent individual genetic propensity. In their study, unmodelled parental genetic effects captured 30% of the predictive power of an educational attainment polygenic score in the child generation (Kong et al. 2018). It has since been shown that this issue of bias from parental genetic effects also applies to GREML SNP heritability estimates for some traits (Young et al. 2018). Although GREML uses unrelated individuals to avoid shared family environmental confounding, it does not use genetic information on parents and so is unable to distinguish between the direct effects of inherited genetic variants (heritability) and the effects of genetic variants in parents acting through the environment.

However, the idea of indirect genetic effects is not completely new. Quantitative genetics researchers have always grappled with the difficulty of separating genetic variance from the variance due to the environment shared by members of a family. It is well-known that genetic effects on behaviour do not 'stop at the skin', and include effects on individuals' environments. A review of biases affecting genomic methods (Young et al. 2019), has been criticised for overlooking decades of pre-genomic era human genetic studies (Anon 2020) – although they did acknowledge the long history of animal breeding studies of indirect genetic effects (e.g. Lush et al. 1934). The fact that parenting behaviour is

genetically influenced, and therefore linked to effects of genes transmitted to children, is well-established in human quantitative genetics. ‘Genetic nurture’ overlaps with a key concept in quantitative genetics. Indeed, **passive gene-environment correlation** refers to how children passively inherit genetic material from their parents, who also create the family environment (Plomin et al. 1977). This phenomenon has been well theorised and modelled in quantitative genetics. For example, the adoption design breaks the genetic link between parents and children. Associations between parent behaviours and later child traits are then more easily interpreted as environmental. The children-of-twins design allows similar insights by comparing avuncular correlations. For example, the association between parental divorce and offspring emotional problems is not due to a causal exposure effect, but instead represents a marker of shared genetic predisposition (D’Onofrio et al. 2007).

Note that passive gene-environment correlation is distinguished from **active and evocative gene-environment correlations**, which reflect how individuals’ own genotypes lead to their phenotypes – they select and evoke environments based on their genetically influenced traits. In contrast to the passive type, these forms of gene-environment correlation do not depend on parents’ genotypes so would not be classed as biasing factors if included in estimates of ‘direct’ genetic influence.

The concepts of genetic nurture and passive gene-environment correlation are overlapping to the extent that some argue that they are interchangeable (Selzam et al. 2019). However, the following text delineates some differences, which can be traced to the scientific context (research questions and available data) in which they were developed. Highlighting conceptual discrepancies helps to clarify what is new about modern genomic research on gene-environment interplay.

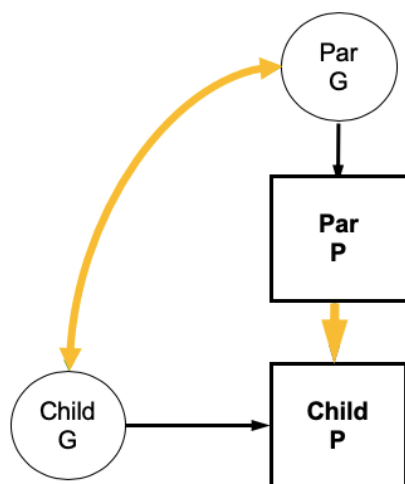
Figure 3 shows that passive gene-environment correlation and genetic nurture are both captured within the diagram from Figure 2. However, the two concepts typically refer to different parts of the diagram (highlighted in yellow). Passive gene-environment correlation (left, Figure 3), the classic concept in quantitative genetics, refers to A: **shared genes** between parents and children (yellow double-headed arrow) inducing B: **a correlation between genetic and environmental effects**, which may cause intergenerational phenotypic associations (yellow arrow from parent to child phenotype) to be partially C: **genetically confounded**.

In contrast, in the genomic era, it is now possible to isolate genomic effects that are A: **not shared** with parents and thus to calculate genetic nurture as the independent effect of parent genetics on a child phenotype (yellow path flowing from parent genome, via parent behaviour, to the child phenotype; right, Figure 3). Genetic nurture, *if shown to operate over and above child genetic influences*, must be B: an **environmentally mediated effect**, rather than the effect of correlation between genetic and environmental effects. Therefore, genetic nurture refers to a novelty from the DNA era – indexing effects of parents using the parental genome. Additionally, instead of focusing on confounding of putative environmental effects, genetic nurture research focuses on C: **environmental confounding** of genetic effects.

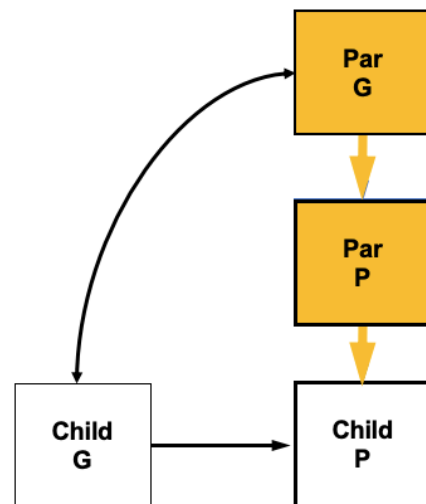
‘Genetic nurture’ is hereafter used to label any indirect genetic effect that can be shown to operate through the environment, over and above effects of inherited offspring genetics. ‘Passive gene-environment correlation’ will refer to how genetic propensities that were passed down to offspring are also associated with environment-providing parent behaviours. This term sometimes refers to situations where measured parent behaviours are not necessarily affecting the child phenotype but are at least in part just epiphenomena reflecting shared genes. Here, it refers to cases where parent genes *do* influence the child phenotype indirectly via parent behaviours, and also act directly through the child. In other words, passive gene-environment correlation denotes a correlation between direct and nurturing genetic effects. Chapter 5 and Appendix 4 describe and compare concepts and methods in detail.

Figure 3. Differences between concepts in gene-environment interplay research. Yellow highlighting emphasises how passive gene-environment correlation and genetic nurture refer to different sections of the diagram of aetiological links between parents and children. Note that the genetic parts of the genetic nurture diagram are in squares to reflect that genotypes are measured (unlike in quantitative genetics, focused on latent influences [circles]). G= genetics, P=phenotype, Par=parents

Passive gene-environment correlation



Genetic nurture



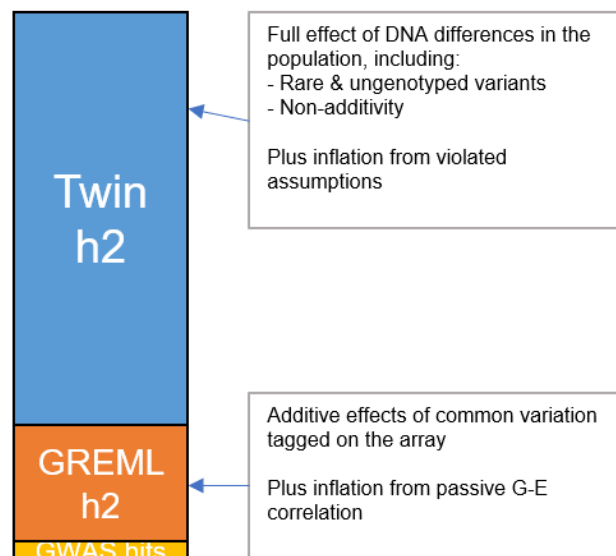
New genomic methods allow direct and indirect genetic effects to be estimated separately (Bates et al. 2018; Kong et al. 2018; Young et al. 2018). However, existing methods usually require intergenerational genomic data, which are scarce. Also, these methods have not yet been applied to investigate environmentally mediated effects on childhood traits for which parents are particularly likely to be influential, such as childhood mental health. No studies have theoretically or empirically compared findings from these new methods to traditional quantitative designs or sought to harmonise language between disciplines. This thesis seeks to close these gaps.

Summary

Confounding between genetic relatedness and environmental sharing greatly complicate interpretations of genomic studies. Family genomic data, such as genotyped trios, are needed to disentangle direct genetic effects from correlated environmentally mediated effects of the parent genome.

The two key issues – missing heritability and family environmental bias due to passive gene-environment correlation – are related. Interestingly, the latter biases SNP heritability estimates in the opposite direction to most of the factors considered in the missing heritability section. Passive gene-environment correlation has been shown to inflate heritability in genomic designs, yet in the twin design it would load onto the shared environment component instead. Figure 4 summarises the key factors contributing to SNP and twin heritability estimates (including ‘true’ sources of variation and biases). Factors that affect both heritability methods (e.g. phenotypic measurement) are excluded from the diagram.

Figure 4. Overview of factors influencing the twin-SNP heritability gap. Note that this is for an imaginary phenotype. As will be discussed in other chapters, the magnitude of the gap varies by phenotype



Aims and structure of the thesis

This thesis aims to address several cutting-edge issues in complex trait genomics by taking advantage of methods, samples, and insights from quantitative genetics. The two key issues that are confronted are estimating heritability accurately i.e. capturing the genetic component of traits in full and without bias (all chapters) and understanding the complex roles of the environment (Chapters 4 and 5). In exploring opportunities to resolve contemporary challenges in complex trait genomics, the thesis focuses on several socially important (and socially contingent), commonly assessed quantitative traits: emotional problems including anxiety and depression, neuroticism (a personality trait closely linked to emotional problems), and educational attainment.

Improving measurement (Chapter 2)

Beyond increasing sample sizes, SNP heritability can be increased by improving the assessment of phenotypes. The relevance of family data is that decades of twin studies tell us where to look to construct a more heritable phenotype.

This first study demonstrates the utility of longitudinal twin data for defining a measure of emotional problems that is more heritable and useful for genomics research.

Capturing rarer genetic effects (Chapter 3)

Recent whole-genome sequence heritability work suggests that rare variants are to blame for the missing heritability, at least for height and BMI. Other than whole-genome sequencing, another way to detect rare genetic effects is to focus analysis on families rather than on population-level variation among unrelated individuals. The latter approach, of which GREML is the prime example, cannot detect variants that are not in linkage disequilibrium with (correlated with) common genotyped SNPs. However, if a similar approach is applied to family data (e.g. siblings, parent-offspring pairs; green line in Figure 1), the higher genetic relatedness among relatives increases the correlation between genotyped SNPs and causal variants, because they are more likely to be inherited together (Hill et al. 2018). The increase in linked variants from using this method helps to capture effects of additional genetic variation such as rare variants, copy number variants, and structural variants.

This second study presents evidence that genomic data on families in the UK Biobank can be leveraged to capture additional genetic effects on neuroticism and educational attainment.

Removing family environmental bias (Chapter 4)

There are huge methodological benefits to knowing the processes that have produced variation in offspring generation. **Principled separation of genetic and environmental influences is only facilitated by family data.** A prime example from quantitative genetics is the natural experiment of adoption. This rules out passive gene-environment correlation, because children are not genetically related to their parents.

This third study combines the adoption design with genomic tools (polygenic scores and GREML) to measure the extent that estimates of genetic influence on educational attainment tag the environment provided by relatives in the UK Biobank. The fourth study also presents a way to remove family

environmental bias (using trio genomic data to control for shared genes), but the focus of the chapter is shifted from heritability estimation to estimating environmental effects.

Addressing long-standing questions about the environment (Chapter 5)

The preceding sections focused on biases in heritability estimates, including environmental bias, and how family data can be used to eliminate this. The other side of the coin is that the same genomic methods for removing family environmental bias can shed light on the importance of parent behaviour in novel ways. As mentioned above, due to gene-environment interdependence, causal environmental effects of parents are difficult to isolate. Genetically sensitive parenting research has long been able to combat this, but has limitations, such as reliance upon highly specific family structures (e.g. extended twin families). Recent developments in family-based genomics can overcome several drawbacks of traditional designs, and new methods allow environmental effects of parents to be indexed using the parent genome.

This final study applies a trio-based genomic method to unpick environmental from genetic transmission of anxiety and depression, using genotyped trios in a multi-generational Norwegian cohort.

Collectively, these studies demonstrate the value of family-based genomic research. Combining cutting-edge genomics methods with insights and data from family studies gives rise to more accurate and novel findings about genetic and environmental contributions to complex phenotypes.

References

- Allegriani, A.G., Cheesman, R., Rimfeld, K., et al. 2020. The p factor: genetic analyses support a general dimension of psychopathology in childhood and adolescence. *Journal of Child Psychology and Psychiatry, and Allied Disciplines* 61(1), pp. 30–39.
- Anon 2020. Deconstructing the history of genotype-phenotype associations in humans [Online]. Available at: <https://science.sciencemag.org/content/deconstructing-history-genotype-phenotype-associations-humans> [Accessed: 27 February 2020].
- Bates, T.C., Maher, B.S., Medland, S.E., McAloney, K., Wright, M.J., Hansell, N.K., Kendler, K.S., Martin, N.G. and Gillespie, N.A. 2018. The Nature of Nurture: Using a Virtual-Parent Design to Test Parenting Effects on Children's Educational Attainment in Genotyped Families. *Twin Research and Human Genetics* 21(2), pp. 73–83.
- Cheesman, R., Selzam, S., Ronald, A., et al. 2017. Childhood behaviour problems show the greatest gap between DNA-based and twin heritability. *Translational Psychiatry*.
- Claussnitzer, M., Cho, J.H., Collins, R., et al. 2020. A brief history of human disease genetics. *Nature* 577(7789), pp. 179–189.
- Coventry, W.L. and Keller, M.C. 2005. Estimating the extent of parameter bias in the classical twin design: a comparison of parameter estimates from extended twin-family and classical twin designs. *Twin Research and Human Genetics* 8(3), pp. 214–223.
- Fisher, R.A. 1918. The correlation between relatives on the supposition of Mendelian inheritance. *Trans R Soc Edinb* 52, pp. 399–433.
- Gage, S.H., Davey Smith, G., Ware, J.J., Flint, J. and Munafò, M.R. 2016. G = E: What GWAS Can Tell Us about the Environment. *PLoS Genetics* 12(2), p. e1005765.
- Hannigan, L.J., Walaker, N., Waszczuk, M.A., McAdams, T.A. and Eley, T.C. 2017. Aetiological influences on stability and change in emotional and behavioural problems across development: a systematic review. *Psychopathology review* 4(1), pp. 52–108.
- Haworth, C.M.A. and Plomin, R. 2010. Quantitative genetics in the era of molecular genetics: learning abilities and disabilities as an example. *Journal of the American Academy of Child and Adolescent Psychiatry* 49(8), pp. 783–793.
- Hill, W.D., Arslan, R.C., Xia, C., et al. 2018. Genomic analysis of family data reveals additional genetic effects on intelligence and personality. *Molecular Psychiatry* 23(12), pp. 2347–2362.
- International HapMap Consortium 2003. The international hapmap project. *Nature* 426(6968), pp. 789–796.
- International Schizophrenia Consortium, Purcell, S.M., Wray, N.R., et al. 2009. Common polygenic variation contributes to risk of schizophrenia and bipolar disorder. *Nature* 460(7256), pp. 748–752.
- Jencks, C. 1980. Heredity, environment, and public policy reconsidered. *American Sociological Review* 45(5), p. 723.
- Kong, A., Thorleifsson, G., Frigge, M.L., et al. 2018. The nature of nurture: Effects of parental genotypes. *Science* 359(6374), pp. 424–428.
- Lee, J.J., Wedow, R., Okbay, A., et al. 2018. Gene discovery and polygenic prediction from a genome-wide association study of educational attainment in 1.1 million individuals. *Nature Genetics* 50(8), pp. 1112–1121.
- Lush, J.L., Hetzer, H.O. and Culbertson, C.C. 1934. Factors affecting birth weights of swine. *Genetics* 19(4), pp. 329–343.
- Mah, S., Nelson, M.R., Delisi, L.E., et al. 2006. Identification of the semaphorin receptor PLXNA2 as a candidate for susceptibility to schizophrenia. *Molecular Psychiatry* 11(5), pp. 471–478.
- Manolio, T.A., Collins, F.S., Cox, N.J., et al. 2009. Finding the missing heritability of complex diseases. *Nature* 461(7265), pp. 747–753.

- Nivard, M.G., Dolan, C.V., Kendler, K.S., et al. 2015. Stability in symptoms of anxiety and depression as a function of genotype and environment: a longitudinal twin study from ages 3 to 63 years. *Psychological Medicine* 45(5), pp. 1039–1049.
- Pardiñas, A.F., Holmans, P., Pocklington, A.J., et al. 2018. Common schizophrenia alleles are enriched in mutation-intolerant genes and in regions under strong background selection. *Nature Genetics* 50(3), pp. 381–389.
- Plomin, R., DeFries, J.C. and Loehlin, J.C. 1977. Genotype-environment interaction and correlation in the analysis of human behavior. *Psychological Bulletin* 84(2), pp. 309–322.
- Plomin, R., Haworth, C.M.A. and Davis, O.S.P. 2009. Common disorders are quantitative traits. *Nature Reviews. Genetics* 10(12), pp. 872–878.
- Polderman, T.J.C., Benyamin, B., de Leeuw, C.A., et al. 2015. Meta-analysis of the heritability of human traits based on fifty years of twin studies. *Nature Genetics* 47(7), pp. 702–709.
- Sekar, A., Bialas, A.R., de Rivera, H., et al. 2016. Schizophrenia risk from complex variation of complement component 4. *Nature* 530(7589), pp. 177–183.
- Selzam, S., Ritchie, S.J., Pingault, J.-B., Reynolds, C.A., O'Reilly, P.F. and Plomin, R. 2019. Comparing Within- and Between-Family Polygenic Score Prediction. *American Journal of Human Genetics* 105(2), pp. 351–363.
- Turkheimer, E. 2000. Three laws of behavior genetics and what they mean. *Current Directions in Psychological Science* 9(5), pp. 160–164.
- Vinkhuyzen, A.A.E., Pedersen, N.L., Yang, J., et al. 2012. Common SNPs explain some of the variation in the personality dimensions of neuroticism and extraversion. *Translational psychiatry* 2, p. e102.
- Vinkhuyzen, A.A.E., Wray, N.R., Yang, J., Goddard, M.E. and Visscher, P.M. 2013. Estimation and partition of heritability in human populations using whole-genome analysis methods. *Annual Review of Genetics* 47, pp. 75–95.
- Visscher, P.M. and Goddard, M.E. 2019. From R.A. Fisher's 1918 paper to GWAS a century later. *Genetics* 211(4), pp. 1125–1130.
- Visscher, P.M., Wray, N.R., Zhang, Q., et al. 2017. 10 years of GWAS discovery: biology, function, and translation. *American Journal of Human Genetics* 101(1), pp. 5–22.
- Wainschtein, P., Jain, D.P., Yengo, L., et al. 2019. Recovery of trait heritability from whole genome sequence data. *BioRxiv*.
- Yang, J., Benyamin, B., McEvoy, B.P., et al. 2010. Common SNPs explain a large proportion of the heritability for human height. *Nature Genetics* 42(7), pp. 565–569.
- Yang, J., Lee, S.H., Goddard, M.E. and Visscher, P.M. 2011. GCTA: a tool for genome-wide complex trait analysis. *American Journal of Human Genetics* 88(1), pp. 76–82.
- Young, A.I., Benonisdottir, S., Przeworski, M. and Kong, A. 2019. Deconstructing the sources of genotype-phenotype associations in humans. *Science* 365(6460), pp. 1396–1400.
- Young, A.I., Frigge, M.L., Gudbjartsson, D.F., et al. 2018. Relatedness disequilibrium regression estimates heritability without environmental bias. *Nature Genetics* 50(9), pp. 1304–1310.

Chapter 2. Extracting stability increases the SNP heritability of emotional problems in young people

This chapter is an exact copy of a peer-reviewed publication:

Cheesman, R., Major Depressive Disorder Working Group of the Psychiatric Genomics Consortium, Purves, K.L., Pingault, J-B., Breen, G., Rijdsdijk, F., Plomin, R. and Eley, T.C. 2018. Extracting stability increases the SNP heritability of emotional problems in young people. *Translational psychiatry* 8(1), p. 223.

Supplementary Information is included in Appendix 1.

ARTICLE

Open Access

Extracting stability increases the SNP heritability of emotional problems in young people

Rosa Cheesman¹, Major Depressive Disorder Working Group of the Psychiatric Genomics Consortium, Kirstin L. Purves¹, Jean-Baptiste Pingault^{1,2}, Gerome Breen^{1,3}, Fruhling Rijdsdijk¹, Robert Plomin¹ and Thalia C. Eley^{1,3}

Abstract

Twin studies have shown that emotional problems (anxiety and depression) in childhood and adolescence are moderately heritable (~20–50%). In contrast, DNA-based ‘SNP heritability’ estimates are generally <15% and non-significant. One notable feature of emotional problems is that they can be somewhat transient, but the moderate stability seen across time and across raters is predominantly influenced by stable genetic influences. This suggests that by capturing what is in common across time and across raters, we might be more likely to tap into any underlying genetic vulnerability. We therefore hypothesised that a phenotype capturing the pervasive stability of emotional problems would show higher heritability. We fitted single-factor latent trait models using 12 emotional problems measures across ages 7, 12 and 16, rated by parents, teachers and children themselves in the Twins Early Development Study sample. Twin and SNP heritability estimates for stable emotional problems ($N = 6110$ pairs and 6110 unrelated individuals, respectively) were compared to those for individual measures. Twin heritability increased from 45% on average for individual measures to 76% ($se = 0.023$) by focusing on stable trait variance. SNP heritability rose from 5% on average (n.s.) to 14% ($se = 0.049$; $p = 0.002$). Heritability was also higher for stable within-rater composites. Polygenic scores for both adult anxiety and depression significantly explained variance in stable emotional problems (0.4%; $p = 0.0001$). The variance explained was more than in most individual measures. Stable emotional problems also showed significant genetic correlation with adult depression and anxiety (average = 52%). These results demonstrate the value of examining stable emotional problems in gene-finding and prediction studies.

Introduction

Anxiety and depression, known collectively as emotional problems, are highly prevalent in childhood and adolescence (median prevalence 2–24% for anxiety and 0.2–17% for depression)¹. The average age of onset for anxiety is strikingly young (between 11 and 14), and 50% of anxiety disorders begin before age 11². Similarly, 25% of

depression cases are diagnosed before age 19². Emotional problems often take a chronic course throughout life³, and predict numerous difficulties in adulthood: anxiety⁴, bipolar disorder⁵, disruptive disorders and schizophrenia². Given that emotional problems are common and disabling, it is essential to unravel their origins in childhood to identify risk factors and develop preventative methods and treatments.

Twin studies have shown that emotional problems in childhood and adolescence are moderately heritable (~20–50%)^{6–10}. In contrast, DNA-based ‘SNP heritability’ estimates for childhood emotional problems are generally low and non-significant (in an underpowered and

Correspondence: Rosa Cheesman (rosa.cheesman@kcl.ac.uk)

¹Social, Genetic and Developmental Psychiatry Centre, Institute of Psychiatry, Psychology and Neuroscience, King's College London, London, UK

²Division of Psychology and Language Sciences, University College London, London, UK

Full list of author information is available at the end of the article.

Members of Major Depressive Disorder Working Group of the Psychiatric Genomics Consortium are listed in Supplementary Information

© The Author(s) 2018



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

inconsistent literature—see Supplementary Table 1)¹¹. Low SNP heritability estimates suggest that genome-wide association studies will struggle to shed light on the genetic architecture of early emotional problems, because both SNP heritability and genome-wide association studies are limited to the additive effects of measured common variants (single-nucleotide polymorphisms (SNPs)). Association studies seeking to identify genetic variants influencing the heritability of childhood anxiety and depression have so far been underpowered and unsuccessful^{12–14}. Also, polygenic scores (which aggregate the effects of thousands of genetic variants into a single index of risk for each individual) for a range of traits explain <1% of the variance in emotional problems in childhood and adolescence^{15–17}.

A potential way to improve power to identify significant SNP heritability is to tap into the genetic core of emotional problems by assessing problems that are stable across time. Longitudinal twin studies have shown that emotional problems are only moderately stable across childhood and adolescence, but that stability is predominantly influenced by stable genetic influences^{18–23}. Moreover, longitudinally assessed stable anxiety traits are more heritable than anxiety at a single time-point^{24–26}. In one study, the heritability of adolescents' stable trait anxiety sensitivity across ages 14, 15 and 17 was 61%, whereas age-specific heritability was zero and non-significant apart from at one age²⁴.

Aggregation across reporters and across measures also yields a more heritable core phenotype. Twin research suggests that variation in the part of children's behaviour that raters see in common captures more of the genetic action than rater-specific variation²⁷. For example, two studies found higher twin heritability for rater-common than rater-specific parts of variance in childhood anxiety (~46% vs ~17%)^{28,29}. Further, the substantial covariation between anxiety and depression in childhood and adolescence is largely genetically influenced^{10,30}.

Analysing phenotypic measurements can also increase heritability through improved reliability. Latent trait modelling allows us to account for measurement error, and more robustly summarise measurements of the same latent characteristics across time, measures and raters on one scale. In genetic research, reducing measurement error variance by definition reduces environmental variance. This increases the proportion of phenotypic variance explained by genetic variance—heritability (both twin and SNP). Previous studies have used latent trait modelling including factor analysis^{31,32} and Item Response Theory³³ and obtained more reliable psychopathology measures, from which more accurate heritability estimates can be derived, for example, a latent 'general childhood psychopathology' factor had a high SNP heritability of 38%, reflecting pervasiveness across

domains and raters, as well as lower error³². However, no research thus far has applied latent modelling approaches to longitudinal multi-rater emotional problems data, and estimated both twin and SNP heritability.

Our primary hypothesis was that a phenotype capturing the stability of emotional problems across ages, across measures, and across raters would yield higher heritability than individual anxiety and depression measures. To test our hypothesis, we estimated twin and SNP heritabilities of stable emotional problems phenotypes derived from latent modelling. The use of measures assessing both anxiety and depression across multiple raters increases the likelihood of capturing stability in the emotional problems trait, free of rater- and measure-specific views and bias. By 'stability', we therefore refer to trait variance in childhood emotional problems that is shared across time, situations (raters), and measures.

Our secondary hypothesis was that the heritability of a stable factor based on all questionnaire items would yield higher heritability than a scale-level approach. Items have unique properties, so heritability can be assessed more accurately by optimally weighting items and reducing the proportion of variance accounted for by item-specific measurement error³³. To test this, we compared the twin and SNP heritabilities of factor scores derived from two complementary methods: scale-level Confirmatory Factor Analysis (CFA) and Item Response Theory (IRT). In supplementary analyses, we assessed whether heritability was also higher for a crude composite constructed without latent modelling. We compared the independent contributions of stability across age and across raters by analysing cross-age composites for each rater and cross-rater composites for each age. We also investigated the extent that the heritability of stable emotional problems is inflated by individuals with persistent, severe symptoms.

In sum, the present study estimates twin and SNP heritability of a stable emotional problems phenotype constructed from 12 measures from three ages and three raters. These are compared to the twin and SNP heritabilities of the individual anxiety and depression measures. We also extend previous work by assessing the prediction of our stable emotional problems phenotype by polygenic scores for adult anxiety (UK Biobank)³⁴ and major depression (PGC)³⁵. Additionally, we report results from genome-wide genetic correlation analyses. Our research therefore sheds light on both the genetic architecture of early emotional problems, and links with adult emotional problems.

Materials and methods

Sample

The sample is from the Twins Early Development Study (TEDS), a multivariate, longitudinal study of >10,000 twin pairs representative of England and Wales, recruited

1994–1996³⁶ Analyses were conducted on a sub-sample of unrelated individuals with available emotional problem data and genome-wide genotyping, plus their co-twins (6110 pairs). Informed consent was obtained from all subjects.

Genotyping

Full details are in the Supplementary Information. Genotypes were obtained using the AffymetrixGeneChip 6.0 ($N = 3665$) and HumanOmniExpressExome-8v1.2 arrays ($N = 4649$). Typical quality control procedures were followed (e.g., samples were removed based on call rate < 0.99 , SNPs were removed if minor allele frequency was $< 0.5\%$). Genotypes from the two platforms were separately imputed and then merged.

Measures

All anxiety and depression variables from ages 7, 12 and 16, and from self-, parent- and teacher ratings, were included. Across these ages, anxiety and depression were measured with 12 scales. See Supplementary Table 2 and Fig. 1 for descriptive statistics and histograms of the scales.

The emotional problems subscale of the Strengths and Difficulties Questionnaire (SDQ)³⁷ at ages 7, 12 and 16 was completed by parents^{7,12}, teachers^{7,12} and children themselves^{12,16}.

The Moods and Feelings Questionnaire (MFQ)³⁸ (self- and parent-rated) assessed depressive symptoms at ages 12 and 16. Note that for all MFQ measures except self-

rated at age 16, 11 rather than 13 items were collected, due to similarity to SDQ emotional problems items.

The Anxiety-Related Behaviours Questionnaire (ARBQ)⁶ was administered at age 16 to assess parent ratings of anxiety symptoms, behaviour, emotions and cognition.

The Childhood Anxiety Sensitivity Index (CASI)³⁹ was used at age 16 to measure self-rated anxiety sensitivity (i.e. fear of the experience of anxiety, and the belief that anxiety has negative consequences).

Overall scale scores

Total scores were derived for each scale by taking a mean of the items, requiring at least half the items to be present (e.g. at least 9 for the 18-item CASI). All items had three response categories, taking integer values 0–2. Hence each scale has a range of values from 0 to $(2 \times \text{number of items})$.

Statistical analyses

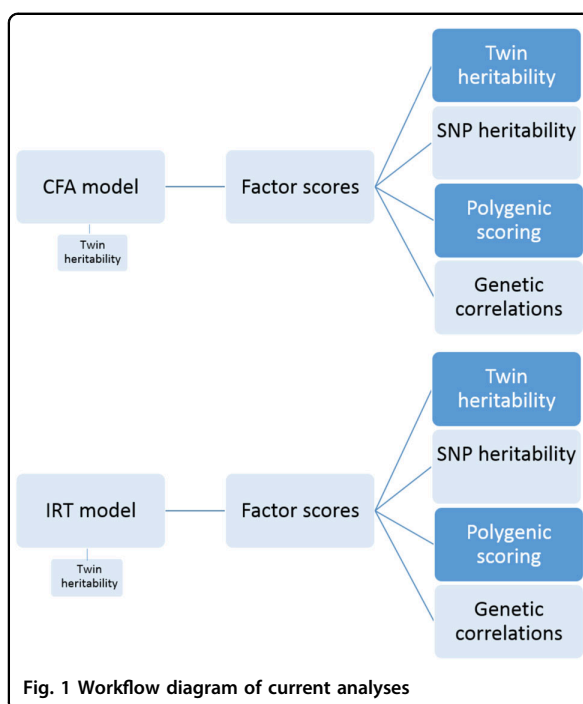
Overview

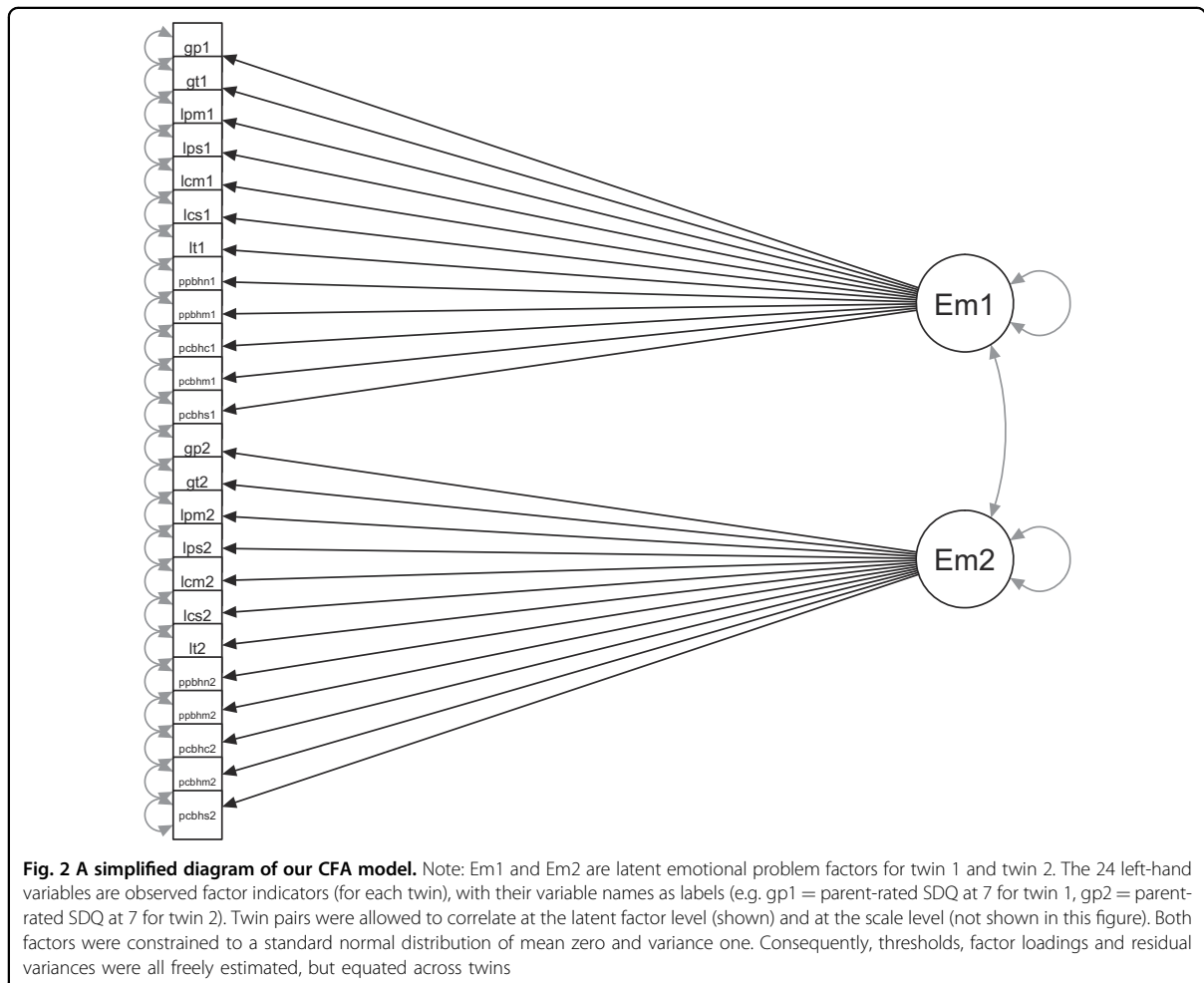
Figure 1 depicts our workflow. We used two latent trait approaches to model our 12 longitudinal multi-rater measures: CFA at the scale level, and IRT at the item-level. In the first stage, we fitted latent trait models to reflect a stable emotional problems factor, accounting for the structure of the twin data (allowing twin heritability to be estimated). In stage two, we extracted individual factor scores for both twins in pair from this latent model ($N = 6110$). Stage three involved four analyses using the stable emotional problems factor scores: (i) estimating twin heritability; (ii) estimating SNP heritability; (iii) predicting stable emotional problems with polygenic scores for adult emotional problems; and (iv) genome-wide association analysis of stable emotional problems. In the fourth stage, we used summary statistics from GWA to explore genetic correlations between stable emotional problems and adult psychopathology. Each stage is explained further below.

To test whether heritability results held with a simpler, non-latent modelling approach, and without combining raters, we conducted sensitivity analyses using simple mean composites across all 12 variables, across age for each variable, and across variables at each age (see Supplementary Figure 4). We also tested whether heritability results held when individuals with persistent severe problems were excluded (see Supplementary Information).

Confirmatory Factor Analysis (CFA)

Figure 2 presents our twelve-indicator one-factor CFA model for twin pairs. The model contains 12 observed continuous variables for each twin of a pair. We included all 12 measures across ages and raters. Two correlated latent emotional problems factors are each measured by the 12 observed factor indicators.





The model was estimated in Mplus, taking into account within-twin-pair correlations at the factor level and at the scale level ($N = 6110$ pairs). Individuals' factor scores were extracted for estimation of twin and SNP heritability.

The CFA model with twin data also allowed us to obtain twin heritability estimates directly—i.e. by simultaneously estimating the latent trait model and directly decomposing the variance of two latent factors (one for each twin) into genetic and environmental components. In Supplementary Table 7 we present the estimates from simultaneous latent trait twin analyses, plus twin heritability estimates for extracted factor scores. The latter approach can introduce more error than simultaneous modelling³³.

Item Response Theory (IRT)

IRT models describe the relationship between individuals' responses to specific questionnaire items and their level of the 'latent variable' being measured. In the standard two-parameter IRT model, the parameters (item

difficulty and discrimination) are equivalent to thresholds and factor loadings in CFA. However, IRT involves categorical observed variables and logistic regressions (rather than continuous and linear). We used all 112 items from the 12 longitudinal cross-rater measures of anxiety and depression in an IRT model. A diagram would look the same as the CFA model in Fig. 1, but with 112 rather than 12 factor indicators for each twin of a pair. Sample size was the same as for CFA.

Heritability analyses of factor scores

Factor scores obtained from the above models for 6110 twin pairs were used to estimate twin and SNP heritabilities. For comparison, we estimated twin and SNP heritabilities for the 12 individual measures.

In the twin design, differences in within-pair correlations for MZ and DZ twins are used to estimate genetic, shared environmental and non-shared environmental effects on traits. Greater MZ than DZ similarity indicates

genetic influence. Within-pair similarity that is not due to genetic factors is attributed to shared environmental influences. Non-shared environment accounts for individual-specific factors that influence differences among siblings from the same family, plus measurement error. Twin model fitting analysis using full-information maximum likelihood was carried out with structural equation modelling software OpenMx in R⁴⁰.

SNP heritability analyses were conducted using one twin of each pair. Analyses were restricted to individuals with available genotyping and factor scores. During estimation of SNP heritability, individuals with missing covariates and excessive genetic relatedness were removed (one from each pair of individuals with pairwise identity-by-descent (IBD) of >0.025 (third degree relatives)), reducing the final samples to 6002 and 6001, respectively.

SNP heritability was estimated using genomic relatedness matrix restricted maximum likelihood (GREML), implemented in the Genome-wide Complex Trait Analysis (GCTA) program⁴¹. For this method, we calculate genetic similarity for each pair of unrelated individuals across all genotyped SNPs. To decompose trait variance, genetic similarity is used to predict phenotypic similarity. GCTA only detects additive genetic effects tagged by common SNPs (here, allele frequencies >5%) in our DNA arrays; the residual component includes any other source of variance, including non-additive genetic effects, rare variants, environment, gene–environment interaction and error. We used sex and the first 10 principal components as covariates. See Supplementary Information for details.

Polygenic score analyses

Polygenic scores aggregate the effects of thousands of SNPs from genome-wide association studies, including variants that do not achieve genome-wide significance, to provide individual-specific ‘genetic propensity’ estimates. An individual’s polygenic score is the sum of their allele count weighted by the effect size for each SNP, as derived from GWAS. We used the high-resolution approach in PRSice 2⁴² to obtain the most predictive polygenic score (with the best *p*-value threshold for inclusion of SNPs) for each phenotype. Our analyses included 10,000 permutations to obtain more stringent empirical *p*-values.

In this study, we generated polygenic scores using summary statistics from PGC Depression (130,664 MDD cases and 330,470 controls; without 23&Me—results based on summary statistics for a subset of 10,000 variants in a sample including 23&Me are reported in Supplementary Table 8)³⁵ and UK Biobank case–control anxiety (25,453 probable Generalised Anxiety Disorder cases and 58,113 controls)³⁴ genome-wide association studies. We predicted phenotypic variance in the CFA- and IRT-derived stable emotional problems scores with the polygenic scores for depression and anxiety. To compare the

prediction of the stable scores with that of individual measures, the 12 individual measures were also regressed on the two polygenic scores. We used sex and the first 10 principal components as covariates.

Genome-wide association analysis

Genome-wide association analysis was performed on the CFA-derived stable emotional problems phenotype (*N* = 6110 individuals) using PLINK v1.90b3.31⁴³.

Subsequent analyses were conducted using the summary statistics: calculation of genetic correlations with UK Biobank anxiety and PGC depression using LD Score Regression 44 (color and linked) and with a range of phenotypes from publicly available external GWAS using LD Hub⁴⁵.

Code availability

Computer code used in our analyses is available from the authors upon request.

Results

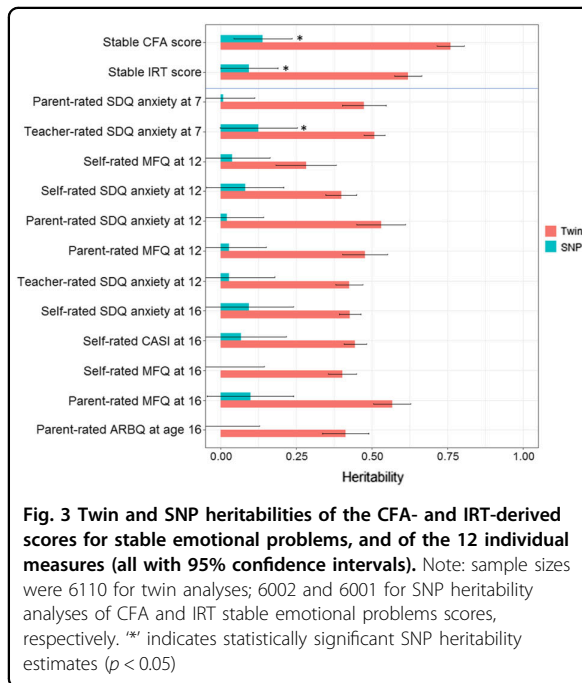
Latent trait analyses

A CFA single-factor model using 12 longitudinal multi-rater emotional problems measures was fitted to the data. All scales loaded significantly on the stable emotional problems factor (*p* < 0.0005). As expected given the low stability of anxiety and depression across childhood (see Supplementary Figure 2 for phenotypic correlations), model fit was relatively poor. However, our aim was to extract a stable developmental factor and estimate its heritability, and not to offer the best explanation of covariance. See Supplementary Tables 4–6 for CFA results.

Heritability analyses

Results support our hypothesis that extracting stable variance increases the heritability of childhood emotional problems. Figure 3 shows that twin heritability increased from 45% on average (range: 28% (se = 0.05; 95% CI = 0.18–0.38) – 57% (se = 0.02; 95% CI = 0.51–0.63)) to 76% (se = 0.02; 95% CI = 0.72–0.81). SNP heritability rose from 5% on average (range: 0% (se = 0.07; 95% CI = –0.15 to 0.15) – 13% (se = 0.07; 95% CI = –0.001 to 0.25)) to 14% (se = 0.05; 95% CI = 0.04–0.24; *p* = 0.002). For most individual anxiety and depression measures, point estimates are low and the intervals based on standard errors cross zero. None of these individual measures had a significant SNP heritability (*p* < 0.05), except teacher-rated SDQ at age 7.

Scale- and item-level approaches increased SNP heritability equally. Promisingly, IRT factor scores are more normally distributed (see Supplementary Figure 3), but CFA is easier to understand and execute. We thus used CFA but not IRT factor scores in subsequent analyses. Twin heritability estimates for simultaneous latent trait-



twin modelling and extracted factor scores were equivalent: point estimates were almost identical and SE intervals were overlapping. See Supplementary Table 7 for full twin and SNP heritability results. Supplementary Figure 4 contains results from sensitivity analyses confirming that aggregation across age and raters increases heritability even when not using latent modelling to account for unreliability. These results also suggest that heritability is still higher when aggregating across age for each rater, and across raters for each age. We show in Supplementary Information that there is no statistically significant difference in heritability upon removal of individuals with persistently severe emotional problems, although point estimates were lower, so replication of this analysis in an independent sample is needed.

Polygenic scoring

Polygenic scores for anxiety and depression significantly explain variance in stable emotional problems (Table 1). More variance is explained than in most of the individual measures (~0.4% vs. ~0.2% on average, although comparison of R^2 statistics is difficult, given the differing target sample sizes and the lack of error estimates).

Table 1 Proportion of phenotypic variance predicted in stable emotional problems, and in 12 anxiety and depression measures, by polygenic scores for Major Depressive Disorder and for General Anxiety Disorder

	R^2		p -value		Empirical p -value	
	PGC MDD	UKBB ANX	PGC MDD	UKBB ANX	PGC MDD	UKBB ANX
Stable CFA score	0.0048	0.0041	<0.0001	<0.0001	0.0001*	0.0001*
Stable IRT score	0.0042	0.0035	<0.0001	<0.0001	0.0001*	0.0001*
Parent-rated SDQ at 7	0.0007	0.0012	0.0481	0.0115	0.1305	0.038*
Teacher-rated SDQ at 7	0.0031	0.003	0.0002	0.0003	0.0013*	0.0010*
Self-rated MFQ at 12	0.0041	0.0008	<0.0001	0.0544	0.0002*	0.1623
Self-rated SDQ at 12	0.001	0.0022	0.0362	0.0014	0.1041	0.0046*
Parent-rated SDQ at 12	0.0017	0.0015	0.0054	0.0087	0.0186*	0.0313*
Parent-rated MFQ at 12	0.0015	0.0025	0.0086	0.0021	0.0277*	0.0030*
Teacher-rated SDQ at 12	0.0008	0.0012	0.0791	0.0019	0.2087	0.034*
Self-rated SDQ at 16	0.0015	0.0031	0.0106	0.0002	0.0325*	0.0005*
Self-rated CASI at 16	0.003	0.0024	0.0006	0.0013	0.0022*	0.0054*
Self-rated MFQ at 16	0.0018	0.0013	0.0075	0.0241	0.0217*	0.0771*
Parent-rated MFQ at 16	0.003	0.0016	0.0007	0.0135	0.0023*	0.0135*
Parent-rated ARBQ at 16	0.003	0.0017	0.0004	0.0089	0.0017*	0.0300*

Note: R^2 indicates proportion of variance predicted by each of the two polygenic scores in the 2 stable and 12 individual phenotypes; *PGC MDD* polygenic scores for Depression (from the PGC without 23&Me data); *UKBB ANX* polygenic scores for Anxiety (from UK Biobank); Empirical p -values were obtained from 10,000 permutations to correct for multiple testing of p -value thresholds, and '*' indicates target TEDS phenotypes that were significantly predicted by polygenic scores at the more stringent empirical p -value thresholds; Supplementary Figures 5, 6 show that prediction is reliable across multiple thresholds; Supplementary Table 8 shows that similar results were obtained using summary stats for 10,000 SNPs from a larger sample including 23&Me data

Genome-wide association

No genome-wide significant SNP associations were identified. See Supplementary Information for Manhattan plots (Supplementary figures 7, 8).

Our stable emotional problems score showed significant genetic correlations with adult depression (case–control (0.48) and symptoms (0.64)), case–control anxiety (0.45), and wellbeing (−0.36). The latter did not remain significant after multiple testing correction, and genetic correlations with nine other phenotypes were non-significant (see Supplementary Table 9).

Discussion

This study found that a phenotype capturing the pervasive stability of emotional problems across childhood and adolescence showed higher heritability than individual measures. Twin heritability increased from 45% on average for individual measures to 76% (se = 0.02; 95% CI = 0.72–0.81) for stable emotional problems. SNP heritability rose from 5% on average to 14% (se = 0.05; 95% CI = 0.04–0.24; $p = 0.002$) by capturing common variance across ages and raters. These findings were consistent for a common factor based on scale-level and item-level data for the 12 measures. The findings also held for a simple, non-latent variable approach. This simple approach is easier to calculate but does not account for measurement error. Additionally, polygenic scores for adult anxiety and depression significantly explained variance in stable emotional problems, and the variance explained (0.4% ($p = 0.0001$)) was higher than in most individual measures. Stable emotional problems showed significant genetic correlation with adult depression and anxiety (average = 52%), mirroring previous findings in adults^{34,46}. Together, the polygenic score and genetic correlation results demonstrate that a significant proportion of common SNPs influencing stable emotional problems in childhood also influence adult anxiety and depression.

This research has several limitations. Our hypothesis was concerned with capturing stability in emotional problems, and not with finding the best-fitting model explaining stability and change across age, and the structure of age-, scale- and rater-specific influences. Consequently, we did not explicitly separate out transient and enduring factors, but acknowledge that the presence of genes driving stable emotional problems does not preclude age effects. We also note the contribution of stability in measurement, not only of longitudinal stability, to increased heritability, although heritability remains higher for stable within-rater composites (Supplementary Figure 4). Structural equation models such as the Trait–State–Occasion (TSO) model⁴⁷ can be used to explicitly separate these effects. However, our inclusion of multi-rater data is an advantage of this study. The common variance that has been extracted is free of rater bias

or rater-specific views. Our factor scores therefore reflect a core emotional problems trait: longitudinally stable and agreed upon by three raters.

Our findings suggest that common genetic variants have stable influences on emotional problems throughout childhood and adolescence, which extend into later life. Additional research from large longitudinal studies is needed to determine whether the SNP heritability of emotional problems is more accurately estimated by capturing pervasive stability beyond age 16, or even throughout the life course. This is likely, given the evidence from twin research that genetic influences on emotional problems at age 3 remain influential well into adulthood¹⁸. Further research could also investigate whether extracting a measure of stability increases the heritability of other less developmentally stable and less heritable traits.

Future genomic studies of emotional problems could benefit from adopting a lifelong approach, using measures of adult case/control status as well as childhood dimensions. Our results indicate that stable emotional problems offer a more useful phenotype than individual measures for: finding variants predisposing to early emotional problems; creating polygenic scores for emotional problems in childhood and adolescence; and for using as a target phenotype for polygenic prediction. Subsequent research should examine how polygenic risk for stable child and adolescent emotional problems is expressed throughout development and into adulthood, as well as the mediators, moderators and multivariate outcomes of this polygenic risk.

The young age of onset of emotional problems, and their persistent, wide-ranging negative outcomes, mean that prediction and prevention should be prioritised. The key contribution of genomic research into early emotional problems is likely to be the predictive value of polygenic scores. The predictive accuracy of polygenic scores is increasing with gains in power for genome-wide association, thanks to collaborative consortia and large national projects with homogenous phenotyping such as the UK Biobank. Researchers are working towards the much-needed large-scale genomic study of emotional problems in childhood itself. The present research underlines the utility of extracting a more stable emotional problems phenotype. This phenotype, theoretically grounded in evidence from decades of twin research, is more reliable, heritable and useful for prediction studies.

Acknowledgements

We gratefully acknowledge the ongoing contribution of the participants in the Twins Early Development Study (TEDS) and their families. TEDS is supported by a program grant to R.P. from the UK Medical Research Council (MR/M021475/1 and previously G0901245), with additional support from the US National Institutes of Health (AG046938). The research leading to these results has also received funding from the European Research Council under the European Union's Seventh Framework Programme (FP7/2007–2013)/ grant agreement

no. 602768 and ERC grant agreement no. 295366). R.P. is supported by a Medical Research Council Professorship award (G19/2). This research has been conducted using the UK Biobank Resource, under application 16577. This study represents independent research part funded by the National Institute for Health Research (NIHR) Biomedical Research Centre at South London and Maudsley NHS Foundation Trust and King's College London. The views expressed are those of the author(s) and not necessarily those of the NHS, the NIHR or the Department of Health. High-performance computing facilities were funded with capital equipment grants from the GSTT Charity (TR130505) and Maudsley Charity (980). R.C. is supported by an ESRC studentship. K.L.P. acknowledges funding from the Alexander von Humboldt Foundation. J.-B.P. is a fellow of MQ: Transforming Mental Health (MQ16IP16).

Author details

¹Social, Genetic and Developmental Psychiatry Centre, Institute of Psychiatry, Psychology and Neuroscience, King's College London, London, UK. ²Division of Psychology and Language Sciences, University College London, London, UK. ³NIHR Biomedical Research Centre for Mental Health; South London and Maudsley NHS Trust, London SE5 8AF, UK

Conflict of interest

The authors declare that they have no conflict of interest.

Publisher's note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Supplementary Information accompanies this paper at (<https://doi.org/10.1038/s41398-018-0269-5>).

Received: 25 March 2018 Revised: 10 July 2018 Accepted: 10 September 2018

Published online: 17 October 2018

References

- Merikangas, K. R., Nakamura, E. F. & Kessler, R. C. Epidemiology of mental disorders in children and adolescents. *Dialog. Clin. Neurosci.* **11**, 7–20 (2009).
- Kessler, R. C. et al. Lifetime prevalence and age-of-onset distributions of DSM-IV disorders in the National Comorbidity Survey Replication. *Arch. Gen. Psychiatry* **62**, 593–602 (2005).
- Beesdo-Baum, K. & Knappe, S. Developmental epidemiology of anxiety disorders. *Child Adolesc. Psychiatr. Clin. N. Am.* **21**, 457–478 (2012).
- Copeland, W. E., Angold, A., Shanahan, L. & Costello, E. J. Longitudinal patterns of anxiety from childhood to adulthood: the Great Smoky Mountains Study. *J. Am. Acad. Child Adolesc. Psychiatry* **53**, 21–33 (2014).
- Cohen, D. et al. Absence of cognitive impairment at long-term follow-up in adolescents treated with ECT for severe mood disorder. *Am. J. Psychiatry* **157**, 460–462 (2000).
- Eley, T. C. et al. A twin study of anxiety-related behaviours in pre-school children. *J. Child Psychol. Psychiatry* **44**, 945–960 (2003).
- Saudino, K. J., Carter, A. S., Purper-Ouakil, D. & Gorwood, P. The etiology of behavioral problems and competencies in very young twins. *J. Abnorm. Psychol.* **117**, 48–62 (2008).
- Van Hulle, C. A., Lemery-Chalfant, K. & Goldsmith, H. H. Genetic and environmental influences on socio-emotional behavior in toddlers: an initial twin study of the infant-toddler social and emotional assessment. *J. Child Psychol. Psychiatry* **48**, 1014–1024 (2007).
- Hallett, V., Ronald, A., Rijdsdijk, F. & Eley, T. C. Phenotypic and genetic differentiation of anxiety-related behaviors in middle childhood. *Depress Anxiety* **26**, 316–324 (2009).
- Waszczuk, M. A., Zavos, H. M. S., Gregory, A. M. & Eley, T. C. The phenotypic and genetic structure of depression and anxiety disorder symptoms in childhood, adolescence, and young adulthood. *JAMA Psychiatry* **71**, 905–916 (2014).
- Cheesman, R. et al. Childhood behaviour problems show the greatest gap between DNA-based and twin heritability. *Transl. Psychiatry* **7**, 1284 (2017).
- Trzaskowski, M. et al. First genome-wide association study on anxiety-related behaviours in childhood. *PLoS ONE* **8**, e58676 (2013).
- Mick, E. et al. Genome-wide association study of the child behavior checklist dysregulation profile. *J. Am. Acad. Child Adolesc. Psychiatry* **50**, 807–817.e8 (2011).
- Benke, K. S. et al. A genome-wide association meta-analysis of preschool internalizing problems. *J. Am. Acad. Child Adolesc. Psychiatry* **53**, 667–676.e7 (2014).
- Krapohl, E. et al. Phenome-wide analysis of genome-wide polygenic scores. *Mol. Psychiatry* **21**, 1188–1193 (2016).
- Jones, H. J. et al. Phenotypic manifestation of genetic risk for schizophrenia during adolescence in the general population. *JAMA Psychiatry* **73**, 221–228 (2016).
- Nivard, M. G., Gage, S. H., Hottenga, J. J., van Beijsterveldt, C. E. M., Abdellaoui, A., Bartels, M., Baselmans, B. M. L., Ligthart, L., Pourcain, B. S., Boomsma, D. I., Munafò, M. R. & Middeldorp, C. M. Genetic overlap between schizophrenia and developmental psychopathology: longitudinal and multivariate polygenic risk prediction of common psychiatric traits during development. *Schizophrenia Bulletin* **43**, 1197–1207 (2017).
- Nivard, M. G. et al. Stability in symptoms of anxiety and depression as a function of genotype and environment: a longitudinal twin study from ages 3 to 63 years. *Psychol. Med.* **45**, 1039–1049 (2015).
- Trzaskowski, M., Zavos, H. M. S., Haworth, C. M. A., Plomin, R. & Eley, T. C. Stable genetic influence on anxiety-related behaviours across middle childhood. *J. Abnorm. Child Psychol.* **40**, 85–94 (2012).
- Waszczuk, M. A., Zavos, H. M. S., Gregory, A. M. & Eley, T. C. The stability and change of etiological influences on depression, anxiety symptoms and their co-occurrence across adolescence and young adulthood. *Psychol. Med.* **46**, 161–175 (2016).
- Hannigan, L. J., Walaker, N., Waszczuk, M. A., McAdams, T. A. & Eley, T. C. Aetiological influences on stability and change in emotional and behavioural problems across development: a systematic review. *Psychopathology review.* **4**, 52–108 (2017).
- Kendler, K. S. et al. A longitudinal twin study of fears from middle childhood to early adulthood: evidence for a developmentally dynamic genome. *Arch. Gen. Psychiatry* **65**, 421–429 (2008).
- Lau, J. Y. F. & Eley, T. C. Changes in genetic and environmental influences on depressive symptoms across adolescence and young adulthood. *Br. J. Psychiatry* **189**, 422–427 (2006).
- Zavos, H. M. S., Gregory, A. M. & Eley, T. C. Longitudinal genetic analysis of anxiety sensitivity. *Dev. Psychol.* **48**, 204–212 (2012).
- Lubke, G. H. et al. A powerful phenotype for gene-finding studies derived from trajectory analyses of symptoms of anxiety and depression between age seven and 18. *Am. J. Med. Genet. B. Neuropsychiatr. Genet.* **171**, 948–957 (2016).
- Hatoun, A. S., Rhee, S. H., Corley, R. P., Hewitt, J. K. & Friedman, N. P. Etiology of stability and growth of internalizing and externalizing behavior problems across childhood and adolescence. *Behav. Genet.* **48**, 298–314 (2018).
- Bartels, M. et al. Disentangling genetic, environmental, and rater effects on internalizing and externalizing problem behavior in 10-year-old twins. *Twin. Res.* **7**, 162–175 (2004).
- Wesseldijk, L. W. et al. Psychopathology in 7-year-old children: Differences in maternal and paternal ratings and the genetic epidemiology. *Am. J. Med. Genet. B. Neuropsychiatr. Genet.* **174**, 251–260 (2016).
- Fedko, I. O. et al. Heritability of behavioral problems in 7-year olds based on shared and unique aspects of parental views. *Behav. Genet.* **47**, 152–163 (2017).
- Eley, T. C. & Stevenson, J. Using genetic analyses to clarify the distinction between depressive and anxious symptoms in children. *J. Abnorm. Child Psychol.* **27**, 105–114 (1999).
- Laurin, C. A., Hottenga, J.-J., Willemsen, G., Boomsma, D. I. & Lubke, G. H. Genetic analyses benefit from using less heterogeneous phenotypes: an illustration with the hospital anxiety and depression scale (HADS). *Genet. Epidemiol.* **39**, 317–324 (2015).
- Neumann, A. et al. Single nucleotide polymorphism heritability of a general psychopathology factor in children. *J. Am. Acad. Child Adolesc. Psychiatry* **55**, 1038–1045.e4 (2016).
- Van den Berg, S. M., Glas, C. A. W. & Boomsma, D. I. Variance decomposition using an IRT measurement model. *Behav. Genet.* **37**, 604–616 (2007).
- Purves, K. L. et al. The common genetic architecture of anxiety disorders. *BioRxiv*. 2017.
- Wray, N. R., Ripke, S., Mattheisen, M., Trzaskowski, M. et al. Genome-wide association analyses identify 44 risk variants and refine the genetic architecture of major depression. *Nature Genetics*. **50**, 668–681 (2018).

36. Haworth, C. M. A., Davis, O. S. P. & Plomin, R. Twins Early Development Study (TEDS): a genetically sensitive investigation of cognitive and behavioral development from childhood to young adulthood. *Twin. Res. Hum. Genet.* **16**, 117–125 (2013).
37. Goodman, R. The strengths and difficulties questionnaire: a research note. *J. Child Psychol. Psychiatry* **38**, 581–586 (1997).
38. Angold, A., Costello, E. J., Messer, S. C. & Pickles, A. The development of a questionnaire for use in epidemiological studies of depression in children and adolescents. *Int. J. Methods Psychiatr. Res.* **5**, 237–249 (1995).
39. Silverman, W. K., Fleisig, W., Rabian, B. & Peterson, R. A. Childhood anxiety sensitivity index. *J. Clin. Child Psychol.* **20**, 162–168 (1991).
40. Boker, S. et al. Openmx: an open source extended structural equation modeling framework. *Psychometrika* **76**, 306–317 (2011).
41. Yang, J., Lee, S. H., Goddard, M. E. & Visscher, P. M. GCTA: a tool for genome-wide complex trait analysis. *Am. J. Hum. Genet.* **88**, 76–82 (2011).
42. Euesden, J., Lewis, C. M. & O'Reilly, P. F. PRSice: Polygenic Risk Score software. *Bioinformatics* **31**, 1466–1468 (2015).
43. Purcell, S. et al. PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am. J. Hum. Genet.* **81**, 559–575 (2007).
44. Bulik-Sullivan, B., Finucane, H. K., Anttila, V., Gusev, A., Day, F. R., Loh, P.-R. ReproGen Consortium, Psychiatric Genomics Consortium, Genetic Consortium for Anorexia Nervosa of the Wellcome Trust Case Control Consortium 3, Duncan, L., Perry, J.R.B., Patterson, N., Robinson, E.B., Daly, M.J., Price, A.L. and Neale, B.M. 2015. An atlas of genetic correlations across human diseases and traits. *Nature Genetics* **47**, pp. 1236–1241 (2015).
45. Zheng, J. et al. LD Hub: a centralized database and web interface to perform LD score regression that maximizes the potential of summary level GWAS data for SNP heritability and genetic correlation analysis. *Bioinformatics* **33**, 272–279 (2017).
46. Brainstorm Consortium, Anttila, V., Bulik-Sullivan, B., Finucane, H. K., et al. Analysis of shared heritability in common disorders of the brain. *Science* **360**, 1–12 (2015).
47. Cole, D. A. Coping with longitudinal data in research on developmental psychopathology. *Int. J. Behav. Dev.* **30**, 20–25 (2006).

Chapter 3. Familial Influences on Neuroticism and Education in the UK Biobank

This chapter is an exact copy of a peer-reviewed publication:

Cheesman, R., Coleman, J., Rayner, C., Purves, K.L., Morneau-Vaillancourt, G., Glanville, K., Choi, S.W., Breen, G. and Eley, T.C. 2019. Familial influences on neuroticism and education in the UK biobank. *Behavior Genetics*. doi:10.1007/s10519-019-09984-5

Supplementary Information is included in Appendix 2.



Familial Influences on Neuroticism and Education in the UK Biobank

R. Cheesman¹ · J. Coleman^{1,2} · C. Rayner¹ · K. L. Purves¹ · G. Morneau-Vaillancourt³ · K. Glanville¹ · S. W. Choi¹ · G. Breen^{1,2} · T. C. Eley^{1,2}

Received: 21 March 2019 / Accepted: 20 November 2019
© The Author(s) 2019

Abstract

Genome-wide studies often exclude family members, even though they are a valuable source of information. We identified parent–offspring pairs, siblings and couples in the UK Biobank and implemented a family-based DNA-derived heritability method to capture additional genetic effects and multiple sources of environmental influence on neuroticism and years of education. Compared to estimates from unrelated individuals, total heritability increased from 10 to 27% and from 17 to 56% for neuroticism and education respectively by including family-based genetic effects. We detected no family environmental influences on neuroticism. The couple similarity variance component explained 35% of the variation in years of education, probably reflecting assortative mating. Overall, our genetic and environmental estimates closely replicate previous findings from an independent sample. However, more research is required to dissect contributions to the additional heritability by rare and structural genetic effects, assortative mating, and residual environmental confounding. The latter is especially relevant for years of education, a highly socially contingent variable, for which our heritability estimate is at the upper end of twin estimates in the literature. Family-based genetic effects could be harnessed to improve polygenic prediction.

Keywords Genomics · Family data · Heritability · Neuroticism · Education

Introduction

Heritability measures the proportion of individual differences in a trait explained by inherited genetic variation, and is traditionally estimated by comparing the resemblance of identical and non-identical twins (Knopik et al. 2017).

Researchers can also now estimate single nucleotide polymorphism (SNP) heritability, the variance explained by the additive effects of common genetic variants tagged by a genotyping array (Yang et al. 2010, 2011). SNP heritability is expected to be less than twin and family-based heritability, since the former only estimates the additive effects of measured common variants, plus variants that are correlated (i.e. in linkage disequilibrium) with them, and ignores influences of DNA sequence differences that are rare and/or not well tagged by genotyping arrays. Since genome-wide association studies (GWAS) also only consider the additive effects of common variants, SNP heritability provides an estimate of the total genetic effect that could be identified with well-powered association studies of a given phenotype in a given population. Given the importance of SNP heritability, researchers have investigated approaches to maximise the accuracy of estimates, beyond increasing sample sizes or denser genotyping (van den Berg et al. 2014; Laurin et al. 2015; Cheesman et al. 2018; van der Sluis et al. 2010).

The dominant method for the estimation of SNP heritability, Genomic-RElatedness-based restricted Maximum-Likelihood (GREML), takes advantage of small genetic differences among many unrelated individuals to predict trait

Handling Editor: Meike Bartels, PhD

Electronic supplementary material The online version of this article (<https://doi.org/10.1007/s10519-019-09984-5>) contains supplementary material, which is available to authorized users.

✉ R. Cheesman
rosa.cheesman@kcl.ac.uk

✉ T. C. Eley
thalia.eley@kcl.ac.uk

¹ Social Genetic & Developmental Psychiatry Centre, Institute of Psychiatry, Psychology & Neuroscience, King's College London, 16 de Crespigny Park, Denmark Hill, London SE5 8AF, UK

² NIHR Biomedical Research Centre for Mental Health, South London and Maudsley NHS Trust, London, UK

³ Research Unit on Child Psychosocial Maladjustment, Laval University, Quebec City, Canada

similarity (Yang et al. 2011). The effects of genetic variants that are not in linkage disequilibrium with common genotyped SNPs will not be captured using this method. However, when GREML is applied to family data, the higher genetic relatedness among relatives increases the correlation between genotyped SNPs and causal variants, because they are more likely to be inherited together (Zaitlen et al. 2013; Xia et al. 2016). This increase in linked variants helps to capture additional genetic variation not normally picked up in population studies of unrelated people, such as rare single nucleotide variants, copy number variants, and structural variants.

An extension of the method that uses family data, GREML-KIN, allows the estimation of two categories of genetic influence: population-level common variant heritability, plus additional heritability that is associated with kinship (Xia et al. 2016; Zaitlen et al. 2013). The first estimate is similar to that derived from GREML using unrelated individuals. The latter heritability estimate captures additional family-based effects, due to the increased correlation between genotyped SNPs and causal variants among relatives. The GREML-KIN method also allows for effects of environment sharing amongst family members, siblings and couples. This inclusion is important as it attempts to remove confounding that results from people who are more genetically related having more similar environments and higher phenotypic resemblance than people who are less related.

One study applied GREML-KIN to neuroticism and years of education (Hill et al. 2018) in Generation Scotland, a large family-based study (Smith et al. 2006). Neuroticism is a personality trait characterised by readily experiencing negative emotions. It is a strong predictor of common mental disorders, occupational attainment and mortality (Ormel et al. 2013). Years of education is also a complex trait that shows significant associations with diverse socioeconomic and health outcomes (Mackenbach et al. 2008). Twin studies have repeatedly demonstrated substantial broad sense heritability for neuroticism (40–60%) (Hettema et al. 2006) and educational attainment (40–50%) (Branigan et al. 2013). Extended twin family designs estimate the narrow sense heritability of neuroticism to be ~25% (Coventry and Keller 2005). These designs provide a better benchmark for additive genetic influence than the classical twin design, since they incorporate relatives of twin pairs to allow explicit separation additive genetic influences from non-additivity, shared environment, assortative mating, and gene-environment correlation. We are not aware of applications of the design to educational attainment, but one study estimated the narrow sense heritability of adult intelligence, a comparable phenotype, to be 44% (Vinkhuyzen et al. 2012). Neuroticism and educational attainment both index a range of important traits, and are available in numerous large datasets. As a result, they have been subject to some of the largest GWA

studies of psychological traits (N = 449,484 for neuroticism (Nagel et al. 2018; Luciano et al. 2018); N = 1.1 million for years of education (Lee et al. 2018)). Nonetheless, only ~10% of the variance in neuroticism can be explained by the additive effects of common SNPs (Luciano et al. 2018; Nagel et al. 2018). Estimates of SNP heritability for years of education also fall substantially short of twin and pedigree estimates (14.7%; Lee et al. 2018).

GREML-KIN analyses in Generation Scotland revealed a large increase in heritability compared to a standard GREML analysis of unrelated individuals (Hill et al. 2018). For neuroticism, the total heritability from the best-fitting model was 30%, primarily accounted for by kin-based genetic effects (19%), as well as common variant effects tagged in studies of unrelated individuals (11%—akin to SNP heritability). They detected no family environment effects. For years of education, there was a strong kin genetic component (28%) in addition to common genetic influence (16%), plus substantial variance explained by sibling and couple similarity (11% and 31%, respectively). The findings align well with evidence from twin studies that the family environment influences education-related outcomes (~36% of the variance; Branigan et al. 2013) but not neuroticism (Polderman et al. 2015). If these results are replicated, then the total DNA-based heritability of neuroticism and education would be close to twin and pedigree study estimates. Moreover, a replication of these results would suggest that most of the variance in education (86%) can be captured with measured parameters. Notably, for each, the larger component of the genetic contributions results from less common variants not identified in genomic studies of unrelated individuals. The authors also found that rarer variants (0.1–1% in frequency) explained 12% of the variance in education, but did not influence variation in neuroticism (Hill et al. 2018).

Our study aimed to estimate familial influences on neuroticism and years of education in the UK Biobank, using GREML-KIN. We capitalise on the presence of thousands of family members in the UK Biobank to shed light on the genetic and environmental architecture of these two phenotypes. To robustly replicate the previous Generation Scotland study, we ensured that phenotype definitions were as similar as possible, and that there was no sample overlap. Based on previous research, we hypothesised that neuroticism and years of education would show increased heritability by exploiting the higher linkage disequilibrium within families. Our secondary analyses aimed to validate our kin-based estimates and specifically quantify the contribution of rarer genetic variants to the heritability of neuroticism and years of education. For this, we used the LDMS-I method, and stratified imputed genetic variants by their individual level of linkage disequilibrium and allele frequency to allow estimation of the variance explained by rarer variants (Evans et al. 2018).

Methods

Sample

Analyses were conducted using the UK Biobank, a resource containing rich phenotype and genotype data on ~500,000 individuals aged between 40 and 70 (Allen et al. 2014). To minimise confounding from population stratification, analyses were limited to white British individuals. We identified families and restricted heritability analyses to individuals with at least one family member in the UK Biobank, as well as phenotype data on neuroticism and/or years of education. Previous publications suggested a sample size of ~40,000 pairs of family members (parent-offspring, siblings, and couples) within the full dataset (Bycroft et al. 2018).

Genotyping

Genome-wide genetic data from the full release of the UK Biobank data were collected and processed according to the quality control pipeline (Bycroft et al. 2018). For primary GREML-KIN analyses, we used genotyped or imputed SNPs with minor allele frequency > 0.01 and imputation confidence (INFO) score > 0.4, indicating well imputed variants. Due to computing memory constraints, we used PLINK2 to prune down to 241,678 variants in approximate linkage equilibrium using an r^2 threshold of 0.2 (Chang et al. 2015) before calculating genetic relatedness matrices.

Measures

Neuroticism was measured as a continuous trait, captured with 12 questionnaire items such as “Does your mood often go up and down?”, “Would you call yourself tense or ‘highly strung’?”. This trait was defined previously in the UK Biobank (Smith et al. 2013, 2016).

The years of education variable was defined according to ISCED categories, as in previous genomic studies in the UK Biobank and other samples (Hill et al. 2018; Lee et al. 2018). The six response categories were: none of the above (no qualifications) = 7 years of education; CSEs or equivalent = 10 years; O levels/GCSEs or equivalent = 10 years; A levels/AS levels or equivalent = 13 years; other professional qualification = 15 years; NVQ or HNC or equivalent = 19 years; college or university degree = 20 years of education. To test whether the number of response categories affected heritability estimates, as has been shown previously in the UK Biobank (Lee et al. 2018), we ran

sensitivity analyses using a ‘coarsened’ years of education variable, plus a binary variable reflecting college completion (see Supplementary Fig. 1).

In all analyses the following covariates were included: age, sex, the first 40 ancestry principal components from the UK Biobank (Bycroft et al. 2018), genotyping batch, and assessment centre.

Analyses

Identification of family members

Sibling and parent-offspring pairs were identified using relatedness files (KING n.d.) received with the UK Biobank data. Relatedness between two individuals is summarised by a kinship coefficient, which is defined as the probability that a random allele from an individual is identical by descent (IBD) with an allele at the same locus from the other individual (i.e. identical and inherited from a common ancestor). For example, in parent-offspring duos, kinship is ~0.25, as it is the probability that a random allele in a child is from one specific parent (0.5 since humans are diploid) multiplied by the probability that the parental allele from that parent is passed to the child (0.5; independent to the first probability). To allow for normal variation in within-pair similarity, first-degree relatives are therefore defined as pairs that have a pairwise kinship coefficient of ≥ 0.177 and ≤ 0.354 .

To distinguish parent-offspring pairs from sibling pairs, we plotted the proportion of SNPs with *zero* identity-by-state (IBS0) within the kinship bounds of 0.177–0.354 (Fig. 1). IBS describes the probability that alleles are the same regardless of common ancestry. When comparing two individuals, variants are termed *IBS0* if neither allele is shared by the pair. Parent-offspring pairs have $IBS0 \approx 0$ since they share one allele inherited by descent (IBD) in all positions on autosomes. In other words, an individual is unlikely to share *zero* variants with one of their parents, unless for example both copies come from the other parent (uniparental disomy), or unless there are genotyping errors meaning that shared variation is not called. In contrast, siblings have a higher pairwise *IBS0*.

Couples were identified as pairs of unrelated opposite-sex individuals matching exactly on a string of household variables: social deprivation (Townsend Deprivation Index), assessment centre, income, time at address, smoker in household, type of accommodation, relatives in household, number in household. This approach of matching on household variables was used in a recent study of assortative mating in the UK Biobank (Yengo et al. 2018). We note that there is potential for type 1 error: it is possible, especially in densely populated areas, that people could match on all eight variables by chance.

Kin-based SNP heritability method accounting for environmental similarity: GREML-KIN

GREML requires the calculation of genetic similarity for each pair of individuals across genotyped variants. This matrix of genomic similarity is compared to a matrix of pairwise phenotypic similarity using a random-effects mixed linear model, such that the variance of a trait can be decomposed into genetic and residual components, using maximum likelihood. Ordinarily, GREML is applied in samples of unrelated individuals and has a single common genetic variance component.

GREML-KIN is an extension of GREML that estimates the variance explained by multiple genetic and non-genetic sources. The method uses a linear mixed model to fit five matrices: G = common genotyped SNP effects; K = kin genetic effects; F = nuclear family (siblings, parent-offspring, and couples); S = siblings; C = couples. For the G matrix, we calculated genetic similarity for all possible pairs of individuals across all genotyped SNPs. As GREML-KIN allows for effects of the family environment, no relatedness cut-off was applied to the G matrix (unlike the standard GREML model applied only to unrelated individuals, where a cut-off of <0.025 is typically used). The K matrix is a modified G matrix, containing only information on relatives (cut-off >0.025), since values for unrelated pairs are set to 0. Family, sibling and couple (F, S and C) similarity matrices were created in the format required for GCTA. Elements in the genomic relatedness matrix were replaced by 0 if a pair did not have the specific relationship; and 1 if a pair do have the relationship, or for elements representing individuals' relatedness to themselves.

Importantly, the variance components are not purely 'genetic' and 'environmental'. The sibling and couple environment sharing matrices likely pick up variance due to other processes that inflate covariance between relatives, including dominance and assortative mating, respectively.

Assortative mating refers to greater similarity between partners than is expected by chance. This can result from multiple mechanisms, including direct choice based on phenotype, social homogamy, and convergence over time due to shared environments. Assortative mating amongst couples in the UK Biobank sample will be captured by fitting the couple similarity matrix (C). However, to the extent that phenotypic similarity among the parents of the UK Biobank participants reflects their genetic similarity, it is also likely that assortative mating in their parents will contribute to the additive genetic variance in our estimates (G + K). This is because assortative mating induces a positive correlation between trait-increasing alleles ('gametic phase disequilibrium'), which elevates trait-specific genetic and phenotypic variance (Peyrot et al. 2016).

The genetic variance components are also likely to include some bias from the indirect effects of genetic variants shared with relatives. Genetic variants in the parents do not only have direct effects on offspring traits by being transmitted, but they also have indirect influences on offspring traits through the environment that they provide for their children. This can bias SNP-based heritability estimates (Young et al. 2018).

The residual component includes sources of variance that are not captured by the G, K, F, S or C matrices, particularly other environmental influences (idiosyncratic, individual-specific environments or perceptions that are not shared by family members) and error.

To identify the best-fitting model for each trait, we ran a model for every possible combination of variance components (31 models), and compared them with backwards stepwise likelihood ratio testing, starting with the full model and dropping non-significant parameters.

We compared GREML-KIN results against those from a standard GREML model in a subset of unrelated individuals from the family-based analyses. The standard GREML model uses a single genomic relatedness matrix with a cut-off to exclude one from each pair of related individuals (cut-off >0.025). This approach therefore only detects population-level additive genetic effects tagged by common genotyped SNPs, plus potential confounding, for example from gene-environment correlation and population stratification. The residual component contains other sources of variance: gene-environment interaction, error, plus all of the environmental influence, rare variant effects that are not captured when using an unrelated population sample, and non-additive genetic effects.

GREML-LDMS-I to investigate the effects of less common variants

In our GREML-LDMS-I analyses, we started with whole genome data imputed from the HRC panel (93,095,623 autosomal variants; see Bycroft et al. 2018). We ran quality control to include variants across the allele frequency spectrum that were imputed with high confidence (INFO score >0.80), and removed multiallelic variants. Three allele frequency bins were made, containing variants with minor allele frequency ranges of: 0.001–0.01, 0.01–0.1, 0.1–0.5, respectively. SNPs in each bin were split into high versus low linkage disequilibrium categories. We stratified by individual (rather than regional) SNP LD scores, since this has been shown to yield SNP heritability estimates that are more robust across different genetic architectures than estimates from other approaches (Evans et al. 2018). This led to six genome-wide genetic relatedness matrices, one for each allele frequency and LD bin (non-overlapping). All matrices

included the same number of unrelated individuals (cut-off 0.025) with phenotype data and with at least one family connection as in standard GREML. The matrices were simultaneously fitted using a linear mixed model, and estimates were allowed to be negative. In supplementary analyses we explored whether the variance explained by the rarest alleles could be underestimated because their imputation quality was lower. Hence, we checked how many SNPs in each minor allele frequency bin were dropped when applying the INFO > 0.8 cutoff.

Sample independence

To ensure that our results were independent from the previous Generation Scotland study, we compared checksums for both samples to identify and remove overlapping participants. A checksum is the sum of nine numbers taken from binary genotype files. Checksums were obtained from Generation Scotland without accessing genotype data directly. We then ran checksums in the UK Biobank (after ensuring quality control of genomic data was the same), using a script from the Broad Institute, which is available online: https://personal.broadinstitute.org/sripke/share_links/checksums_download/outdated_readme/id_gen_checksum.v2.

We note that the number of relationships per individual in the UK Biobank is lower than in Generation Scotland. Their sample was selected to capture dense kinships, where many individuals have siblings, parents and spouses who are also study participants. This may result in lower power in the UK Biobank to detect influences of family similarity, especially if small in magnitude, and reduces power to separate confounding factors, as in biometric designs (McAdams et al. 2018).

Software

We used the following software in our analyses: identification of family members was performed using R; construction of genomic relationship matrices was done in GCTA; family, sibling and couple similarity matrices were made in bash; GREML analyses were conducted in GCTA. Scripts are available from the lead author on request. The UK Biobank is a public dataset available to all bona fide researchers (with funds to pay the access fee).

Results

Identification of family members

Columns 2–4 of Table 1 (bold) show how many pairs of the three types of family members we identified with available data on neuroticism and years of education. The numbers of

family relationships closely matched findings from previous publications (Yengo et al. 2018; Bycroft et al. 2018). Column 5 contains the sample sizes of family pairs (the total of couple, sibling and parent–offspring pairs). From the number of family pairs, we derived the number of families (column 6), or in other words, the number of discrete sets of individuals who have at least one connection. The number of unique individuals (column 7) represents the total number of participants with at least one family member, after removing double-counted individuals who have multiple connections.

The discrepancies between the numbers of individuals, families and nuclear pairs reflects that most people only have one family member in the study. This contrasts to samples with dense kinship networks such as Generation Scotland, where many individuals have siblings, parents and spouses who are also study participants. As described in the Methods, we distinguished parent–offspring and sibling pairs according to their IBS0 (Fig. 1). We chose a threshold of IBS0 > 0.001 to define siblings (blue) separately from parent–offspring pairs (yellow).

Phenotypic correlations for neuroticism were 0.03 for couples, 0.14 for siblings, and 0.13 for parent–offspring pairs. Phenotypic correlations for years of education were 0.38 for couples, 0.30 for siblings, and 0.26 for parent–offspring pairs.

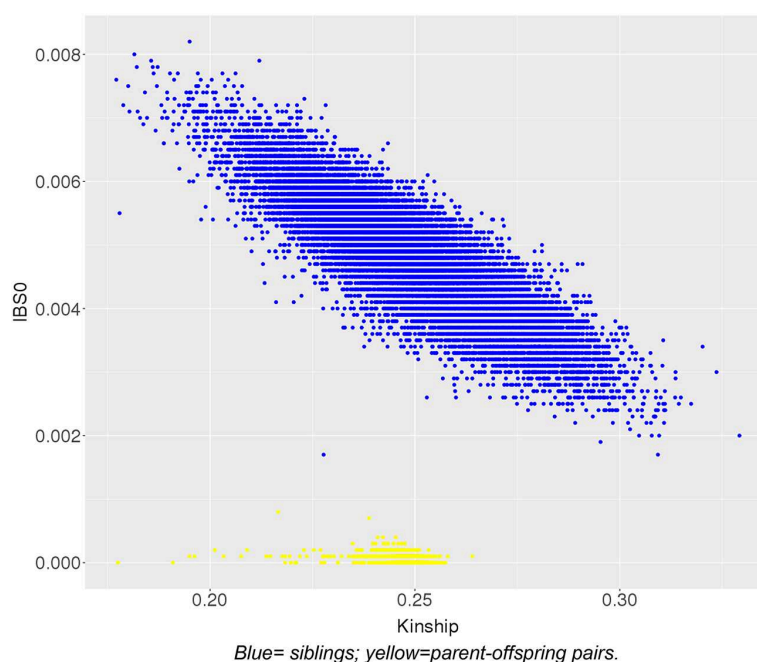
Kin-based SNP heritability method accounting for environmental similarity: GREML-KIN

Figure 2 shows results for the full GREML-KIN models for neuroticism and education years. For neuroticism, the full model indicated that the variance explained by common genetic and kin-based variants is 11% (se = 0.01) and 20% (se = 0.09), respectively (31% in total, first two bars). Bars 3–5 demonstrate that there is no significant influence of family, sibling or couple similarity. For neuroticism, the selected model gives similar results and contains common SNP and kin-based genetic influences (11% (se = 0.01) and 16% (0.02) respectively). Compared to this best-fitting model, the inclusion of matrices to control for the influence of family environments (in the full model) does not reduce heritability. The total additive heritability of neuroticism when including relatives (27%, selected model) is substantially higher than in our analysis of unrelated individuals: 10% (se = 0.01; N = 44,694 unrelated individuals; Supplementary Table 1).

Figure 2 also shows the results for years of education from our full model. Due to computational memory limitations, it was necessary to run models for education in two parts (explained below). We report meta-analysed (inverse variance weighted means) results here. The results for each part are in Supplementary Tables 2 and 3. The heritability of years of education in our selected model was 56%, made up of 12% common SNPs at the population level

Table 1 Sample sizes for different family relationships in the UK Biobank

Phenotype	Couple (pairs)	Sib (pairs)	Parent-offspring (pairs)	Nuclear (pairs)	Families	Unique individuals
Neuroticism	16,451	14,562	4004	35,017	31,369	65,361
Education	23,201	21,564	5912	50,677	44,316	93,737

Fig. 1 Kinship plotted against IBS0 for all first-degree relatives in the UK Biobank. Blue = siblings; yellow = parent-offspring pairs

($se = 0.01$) and 44% kin-based effects ($se = 0.02$). This increases heritability considerably from 17% ($se = 0.01$) in unrelated individuals. The final model for education also contains a couple similarity effect of 35% ($se = 0.01$) in addition to the common and kin-based genetic influences.

See Supplementary Tables 1–3 for full model-fitting results, including all sub-models and fit statistics. As noted above, we ran two sets of GREML-KIN models in independent samples to reduce the computational burden. The first used the same matrices as in analyses of neuroticism. The second used new matrices including individuals who have education data and at least one family member, and who were not included in the neuroticism matrices. In defining these groups, we ensured that individuals in the same family were kept together.

Supplementary Fig. 1 gives GREML-KIN results for alternative education phenotypes (years of education with fewer categories, and degree/college completion). Estimates differed only slightly, and the conclusions remain the same.

GREML-LDMS-I to investigate the effects of less common variants

Table 2 shows our estimates of the contribution of variants of different allele frequencies and individual linkage disequilibrium levels to neuroticism and years of education. For neuroticism, there is no evidence of a contribution of SNPs in the lowest minor allele frequency (MAF) bin (0.001–0.01), but SNPs of MAF 0.01–0.1 explained 3% of the phenotypic variance. All variants explained 11% ($se = 0.02$) of phenotypic variation. For education, variants tagged by low frequency SNPs (MAF between 0.001–0.01, and 0.01–0.1), particularly those in lower linkage disequilibrium, make a modest contribution to phenotypic variation, and all variants explained 21% of the phenotypic variance. Supplementary Table 4 shows that the lower the allele frequency bin, the more SNPs were dropped due to low imputation confidence.

Checksum analyses indicated that there were no family members with phenotype data in the UK Biobank who were

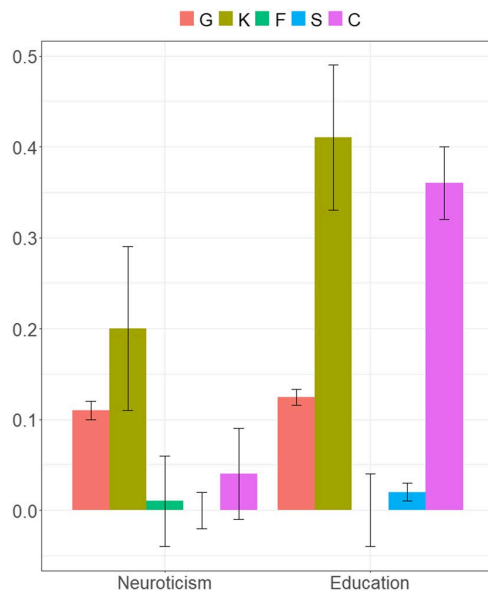


Fig. 2 Variance component estimates for neuroticism and education, plus standard errors. *G* population-level effects of common genotyped SNPs; *K* kin-based genetic effects; *F* effects of nuclear family (siblings, parent-offspring, couple) similarity; *S* effects of sibling similarity; *C* effects of couple similarity. Note that for neuroticism, the estimates from our full model for the *F*, *S*, and *C* components are all non-significant, and standard errors cross zero. For education, the *F* and *S* components are non-significant

also in Generation Scotland. We can therefore be confident that our results are independent of the previous study (Hill et al. 2018).

Discussion

In this study, we capitalised on the family information and genome-wide genotypes in the UK Biobank to estimate genetic and family environmental influences on neuroticism and years of education. Inclusion of family-based genetic effects substantially increases the heritabilities of neuroticism and education from 10 and 17% in standard analyses

of unrelated individuals, to 27% and 56% respectively. Our estimates closely replicate previous findings from an independent sample (Generation Scotland; Hill et al. 2018). The additional family-based influences likely include rare variants, copy number variants, and structural variants. Turning to our non-genetic findings, we did not detect any influence of nuclear family, sibling or couple similarity on variation in neuroticism. This is consistent with evidence against shared environmental contributions to personality (Nivard et al. 2015; Hettema et al. 2006; Coventry and Keller 2005), and with the previous study using the same method (Hill et al. 2018). For education years, we found that 35% of the variance was explained by couple similarity. This couple effect likely captures assortative mating (education years is likely to have occurred before they became couples) and family environment effects on education.

At least three potential biases mean that the higher GREML-KIN heritability may not only be explained by ‘revealing additional genetic effects’. First, confounding is likely to remain in the heritability estimates due to indirect effects of alleles shared by parents and offspring (passive gene-environment correlation). Simulations in an Icelandic dataset suggest that although GREML-KIN gives unbiased heritability estimates when family environmental influences are present, it over-estimates heritability if phenotypes are also substantially influenced by passive gene-environment correlation (Young et al. 2018). The influence of passive gene-environment correlation for educational attainment in the UK Biobank is suggested by the lower SNP heritability for adoptees, whose rearing environments are less correlated with their genotypes (Cheesman et al., 2019). In the future, methods that can distinguish direct from indirect influences should be applied to neuroticism, education and other complex, socially-contingent traits (Eaves et al. 2014; Visscher et al. 2006; Young et al. 2018).

Second, geographic population differences complicate our interpretations of heritability estimates, particularly for education. One study found that controlling for urban childhood residence, a proxy for a range of environmental factors, reduced the SNP heritability of education (but not the heritabilities of height or body mass index) above and beyond other controls for population stratification (Conley et al. 2014). Stratification in the UK Biobank is present even at a

Table 2 Results of GREML-LDMS-I variance components analyses for neuroticism and education using six minor allele frequency and LD bins

	MAF	0.001–0.01	0.001–0.01	> 0.01–0.1	> 0.01–0.1	> 0.1–0.5	> 0.1–0.5	Total
	LD	Lower	Higher	Lower	Higher	Lower	Higher	
No. SNPs		1,772,407	1,772,399	1,756,294	1,756,290	2,105,014	2,104,972	11,267,376
Neuroticism	<i>h</i> ²	0.00 (0.02)	0.00 (0.01)	0.02 (0.01)	0.01 (0.03)	0.05 (0.01)	0.04 (0.004)	0.11 (0.02)
Education	<i>h</i> ²	0.06 (0.02)	0.00 (0.05)	0.03 (0.01)	0.01 (0.01)	0.07 (0.01)	0.05 (0.03)	0.21 (0.02)

MAF minor allele frequency, *LD* linkage disequilibrium, *h*² (*se*) variance explained by SNPs in MAF and LD bin, plus standard error

fine scale, is unlikely to be completely removed by controlling for principal components, and is potentially more severe in rare variant analyses (Haworth et al. 2019; Young 2019).

A third key issue is assortative mating. There is clear evidence for assortment on educational attainment, including in the UK Biobank (Hugh-Jones et al. 2016; Robinson et al. 2017; Yengo et al. 2018). Our estimates of additive genetic variance are likely to be overestimated due to assortative mating among the parents of the UK Biobank participants (see Methods). Even beyond this, assortative mating may inflate the estimated additive genetic variance because it increases phenotypic covariance among relatives and the covariance among specific trait-associated loci, but it is not captured by genome-wide genetic relatedness (the predictor in our GREML models). Assortative mating may also exacerbate the bias due to gene-environment correlation (Keller et al. 2009; Kong et al. 2018). Specifically, the effects of variants influencing educational attainment in offspring may be magnified because variants are also present in the parents, influencing parental education, parents' choice of a mate with similar attainment, and the rearing environment they provide. The effects of assortment on variance component estimates from genomic heritability methods are likely to be complicated and are an important area for future research.

Although the potential biases mean that interpretations should be cautious, our heritability estimates are notably closer to twin and extended twin family study estimates than are those from standard GREML analysis. The latter design is able to differentiate the effects of additive genetics common environment, assortative mating and gene-environment correlation, and the best-powered heritability estimate for neuroticism of ~25% (Coventry and Keller 2005) is close to ours (27%). On the other hand, our total heritability estimate for educational attainment (56%) is notably high and likely to be more strongly biased. It is at the upper end of twin estimates in the literature (40–50%), even though twin heritability does not only reflect additive genetic effects. The increase beyond twin heritability is likely to be partly due to passive gene-environment correlation, which inflates the shared environment component rather than heritability in twin studies.

Our allele frequency and LD partitioned heritability results (GREML-LDMS-I) give mixed support for the notion that the contribution of family-associated genetic influence is explained by rarer variants. For education, some effects of low allele frequency, low LD variants are captured by our partitioned genomic relatedness matrices. For neuroticism, the same GREML-LDMS-I heritability analyses indicated less rare genetic influence than expected based on the large kin-genetic component from the GREML-KIN model-fitting. For neuroticism, the gain in heritability from GREML-KIN despite the lack of variance in the 0.001–0.01 frequency bins is perplexing but may be partly explained

by poor imputation, very rare and structural variation, and population stratification. With regard to imputation quality, the rarest SNPs were more likely to be dropped prior to analysis (see Supplementary Table 4). Since rare SNPs tag other rare SNPs poorly, dropping rare ones would lead to more dramatic underestimation of their effects.

Evidence from other studies suggests an important role for rarer variants, which will be further elucidated with the increasing availability of whole genome sequence data. Multifactorial traits such as personality and educational attainment are thought to have complex genetic aetiologies, consisting of interplay between common and rare variation. Recent exome sequencing work has demonstrated rare coding variant effects on many psychiatric traits [that are closely related to neuroticism; (Ganna et al. 2018)]. Other evidence for interactions between common variants across the genome and rare copy number variants in schizophrenia (Bergen et al. 2018) shows the importance of examining different types of genetic variation together. Importantly, pedigree estimates of the heritability of height and BMI can be recovered by using whole genome sequence data (rather than imputed data) in GREML analyses (Wainschtein et al. 2019). However, much larger sample sizes are needed to detect rare genetic influences, and confounding factors such as stratification, indirect effects and assortative mating are still present.

The additional influences correlated with kinship detected in this study indicate a chance to improve prediction using polygenic scores. Prediction accuracy is a function of the SNP heritability of the target sample, and of the SNP heritability of, and genetic correlation with, the GWAS discovery phenotype (de Vlaming et al. 2017). Polygenic scores for neuroticism currently explain maximum 4.2% of the phenotypic variance in independent samples, which is a fraction of the common SNP heritability (~10%; Nagel et al. 2018). For years of education, polygenic scores explain >10% of the phenotypic variance, close to the SNP heritability of 15%. Increasing sample sizes and phenotypic homogeneity for GWAS of unrelated population samples will continue to narrow this gap between polygenic prediction and SNP heritability. However, the population-level common variant approach may be limited for educational attainment, given the relatively high variance already explained by polygenic scores. For both traits, genomic prediction could be improved by leveraging family information (Lee et al. 2017). Inclusion of relatives could help to capture additional effects of typically untagged variants that are rare and possibly even family-specific. Importantly, for the purpose of prediction, genetic scores need not be, and already are not, 'pure' indices of individual genetic propensity. It may help to tag influential aspects of the environment, for example by combining parent and child polygenic scores within a single model.

In summary, we provide evidence for substantial family-based and common genetic effects on neuroticism and years of education in the UK Biobank. These results motivate the recruitment of samples with dense kinships and methods to leverage genomic family information, whilst understanding gene-environment interplay.

Acknowledgements We would like to thank the scientists involved in the construction of the UK Biobank and all of the participants who have shared their life experiences with investigators in the UK Biobank. This research has been conducted using the UK Biobank Resource, under the application 18177 (with thanks to Paul F. O'Reilly). This study represents independent research part funded by the National Institute for Health Research (NIHR) Biomedical Research Centre at South London and Maudsley NHS Foundation Trust and King's College London. The views expressed are those of the author(s) and not necessarily those of the NHS, the NIHR or the Department of Health and Social Care. High performance computing facilities were funded with capital equipment grants from the GSTT Charity (TR130505) and Maudsley Charity (980). T.C. Eley is part funded by a program grant from the UK Medical Research Council (MR/M021475/1). R. Cheesman is supported by an ESRC studentship. C. Rayner is supported by a grant from Fondation Peters to T.C. Eley and G. Breen. K.L. Purves is part supported by a grant from the Alexander Von Humboldt Foundation and UK Medical Research Council (MR/M021475/1). G. Morneau-Vaillancourt is supported by a studentship from the Quebec Network on Suicide, Mood Disorders and Related Disorders. S.W. Choi. is funded from the UK Medical Research Council (MR/N015746/1). KG is supported by a PhD studentship awarded from the UK Medical Research Council. We thank David M. Howard for assistance with checksums for Generation Scotland, and Loic Yengo for sharing information on identification of couples.

Compliance with ethical standards

Conflict of interest G. Breen is a consultant for Eli Lilly. R. Cheesman, J. Coleman, C. Rayner, K. L. Purves, G. Morneau-Vaillancourt, K. Glanville, S. W. Choi and T. C. Eley declare no conflicts of interest.

Human and Animal Rights and Informed Consent All participants gave full informed written consent for participation in the UK Biobank. This study was performed in accordance with the criteria defined by the rules of the UK Biobank.

Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

References

- Allen NE, Sudlow C, Peakman T, Collins R, UK Biobank (2014) UK biobank data: come and get it. *Sci Transl Med* 6(224):224ed4
- Bergen SE, Ploner A, Howrigan D et al (2018) Joint contributions of rare copy number variants and common snps to risk for schizophrenia. *Am J Psychiatry* 176(1):29–35
- Branigan AR, McCallum KJ, Freese J (2013) Variation in the heritability of educational attainment: an international meta-analysis. *Soc Forces* 92(1):109–140
- Bycroft C, Freeman C, Petkova D et al (2018) The UK Biobank resource with deep phenotyping and genomic data. *Nature* 562(7726):203–209
- Chang CC, Chow CC, Tellier LC, Vattikuti S, Purcell SM, Lee JJ (2015) Second-generation PLINK: rising to the challenge of larger and richer datasets. *GigaScience* 4:7
- Cheesman R, Major Depressive Disorder Working Group of the Psychiatric Genomics Consortium, Purves KL et al (2018) Extracting stability increases the SNP heritability of emotional problems in young people. *Transl Psychiatry* 8(1):223
- Cheesman R, Hunjan A, Coleman JR, Ahmadzadeh Y, Plomin R, McAdams TA, Eley TC, Breen G (2019) Comparison of adopted and non-adopted individuals reveals gene-environment interplay for education in the UK Biobank. *bioRxiv*, p.707695
- Conley D, Siegal ML, Domingue BW, Mullan Harris K, McQueen MB, Boardman JD (2014) Testing the key assumption of heritability estimates based on genome-wide genetic relatedness. *J Hum Genet* 59(6):342–345
- Coventry WL, Keller MC (2005) Estimating the extent of parameter bias in the classical twin design: a comparison of parameter estimates from extended twin-family and classical twin designs. *Twin Res Hum Genet* 8(3):214–223
- de Vlaming R, Okbay A, Rietveld CA et al (2017) Meta-GWAS accuracy and power (MetaGAP) calculator shows that hiding heritability is partially due to imperfect genetic correlations across studies. *PLoS Genet* 13(1):e1006495
- Eaves LJ, Pourcain BS, Smith GD, York TP, Evans DM (2014) Resolving the effects of maternal and offspring genotype on dyadic outcomes in genome wide complex trait analysis (“M-GCTA”). *Behav Genet* 44(5):445–455
- Evans LM, Tahmasbi R, Vrieze SI et al (2018) Comparison of methods that use whole genome data to estimate the heritability and genetic architecture of complex traits. *Nat Genet* 50(5):737–745
- Ganna A, Satterstrom FK, Zekavat SM et al (2018) Quantifying the impact of rare and ultra-rare coding variation across the phenotypic spectrum. *Am J Hum Genet* 102(6):1204–1211
- Haworth S, Mitchell R, Corbin L, Wade KH, Dudding T, Budu-Aggrey A, Carslake D, Hemani G, Paternoster L, Smith GD, Davies N (2019) Apparent latent structure within the UK Biobank sample has implications for epidemiological analysis. *Nat Commun* 10(1):333
- Hettema JM, Neale MC, Myers JM, Prescott CA, Kendler KS (2006) A population-based twin study of the relationship between neuroticism and internalizing disorders. *Am J Psychiatry* 163(5):857–864
- Hill WD, Arslan RC, Xia C et al (2018) Genomic analysis of family data reveals additional genetic effects on intelligence and personality. *Mol Psychiatry* 23(12):2347
- Hugh-Jones D, Verweij KJ, Pourcain BS, Abdellaoui A (2016) Assortative mating on educational attainment leads to genetic spousal resemblance for polygenic scores. *Intelligence* 59:103–108
- Keller MC, Medland SE, Duncan LE, Hatemi PK, Neale MC, Maes HH, Eaves LJ (2009) Modeling extended twin family data I: description of the Cascade model. *Twin Res Hum Genet* 12(1):8–18
- KING KING: Relationship Inference Software [Online]. Available at: <http://people.virginia.edu/~wc9c/KING/>
- Knopik VS, Neiderhiser JM, DeFries JC, Plomin R (eds) (2017) *Behavioral genetics*, 7th edn. Worth, New York
- Kong A, Thorleifsson G, Frigge ML, Vilhjalmsdottir BJ, Young AI, Thorgerdsson TE, Benonisdottir S, Oddsson A, Halldorsson BV, Masson G, Gudbjartsson DF (2018) The nature of nurture: effects of parental genotypes. *Science* 359(6374):424–428

- Laurin CA, Hottenga J-J, Willemsen G, Boomsma DI, Lubke GH (2015) Genetic analyses benefit from using less heterogeneous phenotypes: an illustration with the hospital anxiety and depression scale (HADS). *Genet Epidemiol* 39(4):317–324
- Lee SH, Weerasinghe WMSP, Wray NR, Goddard ME, van der Werf JHJ (2017) Using information of relatives in genomic prediction to apply effective stratified medicine. *Sci Rep* 7:42091
- Lee JJ, Wedow R, Okbay A et al (2018) Gene discovery and polygenic prediction from a genome-wide association study of educational attainment in 11 million individuals. *Nat Genet* 50(8):1112–1121
- Luciano M, Hagenaars SP, Davies G et al (2018) Association analysis in over 329,000 individuals identifies 116 independent variants influencing neuroticism. *Nat Genet* 50(1):6–11
- Mackenbach JP, Stirbu I, Roskam A-JR et al (2008) Socioeconomic inequalities in health in 22 European countries. *N Engl J Med* 358(23):2468–2481
- McAdams TA, Hannigan LJ, Eilertsen EM, Gjerde LC, Ystrom E, Rijdsdijk FV (2018) Revisiting the children-of-twins design: improving existing models for the exploration of intergenerational associations. *Behav Genet* 48(5):397–412
- Nagel M, Jansen PR, Stringer S et al (2018) Meta-analysis of genome-wide association studies for neuroticism in 449,484 individuals identifies novel genetic loci and pathways. *Nat Genet* 50(7):920–927
- Nivard MG, Middeldorp CM, Dolan CV, Boomsma DI (2015) Genetic and environmental stability of neuroticism from adolescence to adulthood. *Twin Res Hum Genet* 18(6):746–754
- Ormel J, Jeronimus BF, Kotov R et al (2013) Neuroticism and common mental disorders: meaning and utility of a complex relationship. *Clin Psychol Rev* 33(5):686–697
- Peyrot WJ, Robinson MR, Penninx BW, Wray NR (2016) Exploring boundaries for the genetic consequences of assortative mating for psychiatric traits. *JAMA Psychiatry* 73(11):1189–1195
- Polderman TJC, Benyamin B, de Leeuw CA et al (2015) Meta-analysis of the heritability of human traits based on fifty years of twin studies. *Nat Genet* 47(7):702–709
- Robinson MR, Kleinman A, Graff M et al (2017) Genetic evidence of assortative mating in humans. *Nat Hum Behav* 1(1):0016
- Smith BH, Campbell H, Blackwood D et al (2006) Generation Scotland: the Scottish Family Health Study; a new resource for researching genes and heritability. *BMC Med Genet* 7:74
- Smith DJ, Nicholl BI, Cullen B et al (2013) Prevalence and characteristics of probable major depression and bipolar disorder within UK biobank: cross-sectional study of 172,751 participants. *PLoS ONE* 8(11):e75362
- Smith DJ, Escott-Price V, Davies G et al (2016) Genome-wide analysis of over 106 000 individuals identifies 9 neuroticism-associated loci. *Mol Psychiatry* 21(6):749–757
- van den Berg SM, de Moor MHM, McGue M et al (2014) Harmonization of neuroticism and extraversion phenotypes across inventories and cohorts in the genetics of personality consortium: an application of item response theory. *Behav Genet* 44(4):295–313
- van der Sluis S, Verhage M, Posthuma D, Dolan CV (2010) Phenotypic complexity, measurement bias, and poor phenotypic resolution contribute to the missing heritability problem in genetic association studies. *PLoS ONE* 5(11):e13929
- Vinkhuyzen AA, Van Der Sluis S, Maes HH, Posthuma D (2012) Reconsidering the heritability of intelligence in adulthood: taking assortative mating and cultural transmission into account. *Behav Genet* 42(2):187–198
- Visscher PM, Medland SE, Ferreira MAR et al (2006) Assumption-free estimation of heritability from genome-wide identity-by-descent sharing between full siblings. *PLoS Genet* 2(3):e41
- Wainschtein P, Jain DP, Yengo L et al (2019) Recovery of trait heritability from whole genome sequence data. *BioRxiv*. <https://doi.org/10.1101/588020>
- Xia C, Amador C, Huffman J et al (2016) Pedigree- and SNP-associated genetics and recent environment are the major contributors to anthropometric and cardiometabolic trait variation. *PLoS Genet* 12(2):e1005804
- Yang J, Benyamin B, McEvoy BP et al (2010) Common SNPs explain a large proportion of the heritability for human height. *Nat Genet* 42(7):565–569
- Yang J, Lee SH, Goddard ME, Visscher PM (2011) GCTA: a tool for genome-wide complex trait analysis. *Am J Hum Genet* 88(1):76–82
- Yengo L, Robinson MR, Keller MC et al (2018) Imprint of assortative mating on the human genome. *Nat Hum Behav* 2(12):948–954
- Young AI (2019) Solving the missing heritability problem. *PLoS Genet* 15(6):e1008222
- Young AI, Frigge ML, Gudbjartsson DF et al (2018) Relatedness disequilibrium regression estimates heritability without environmental bias. *Nat Genet* 50(9):1304–1310
- Zaitlen N, Kraft P, Patterson N et al (2013) Using extended genealogy to estimate components of heritability for 23 quantitative and dichotomous traits. *PLoS Genet* 9(5):e1003520

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Chapter 4. Comparison of adopted and non-adopted individuals reveals gene-environment interplay for education in the UK Biobank

This chapter presents page proofs for an accepted peer-reviewed article, which is scheduled to appear in *Psychological Science*:

Cheesman, R., Hunjan, A., Coleman, J., Ahmadzadeh, Y., Plomin, R., McAdams, T., Eley, T.C. Breen, G. 2019. Comparison of adopted and non-adopted individuals reveals gene-environment interplay for education in the UK Biobank. In press *Psychological Science*.

Supplementary Information is included in Appendix 3.

Comparison of Adopted and Nonadopted Individuals Reveals Gene–Environment Interplay for Education in the UK Biobank

Psychological Science
 1–10
 © The Author(s) 2020
 Article reuse guidelines:
sagepub.com/journals-permissions
 DOI: 10.1177/0956797620904450
www.psychologicalscience.org/PS


Rosa Cheesman¹ , Avina Hunjan^{1,2}, Jonathan R. I. Coleman^{1,2},
 Yasmin Ahmadzadeh¹, Robert Plomin¹, Tom A. McAdams¹,
 Thalia C. Eley^{1,2}, and Gerome Breen^{1,2} **[AQ: 1][AQ: 2][AQ: 3]**

¹Social, Genetic & Developmental Psychiatry Centre, Institute of Psychiatry, Psychology & Neuroscience, King's College London, and ²National Institute for Health Research (NIHR) Biomedical Research Centre for Mental Health, South London and Maudsley National Health Service (NHS) Trust, London, United Kingdom

Abstract

Polygenic scores now explain approximately 10% of the variation in educational attainment. However, they capture not only genetic propensity but also information about the family environment. This is because of passive gene–environment correlation, whereby the correlation between offspring and parent genotypes results in an association between offspring genotypes and the rearing environment. We measured passive gene–environment correlation using information on 6,311 adoptees in the UK Biobank. Adoptees' genotypes were less correlated with their rearing environments because they did not share genes with their adoptive parents. We found that polygenic scores were twice as predictive of years of education in nonadopted individuals compared with adoptees (R^2 s = .074 vs. .037, $p = 8.23 \times 10^{-24}$). Individuals in the lowest decile of polygenic scores for education attained significantly more education if they were adopted, possibly because of educationally supportive adoptive environments. Overall, these results suggest that genetic influences on education are mediated via the home environment.

Keywords

educational attainment, polygenic scores, gene–environment interplay, adoption

Received 8/2/19; Revision accepted 12/30/19

An important process by which genes and environments work together to influence behavior is gene–environment correlation (Plomin, DeFries, Knopik, & Neiderhiser, 2016). Gene–environment correlation refers to the association between the genotype that individuals inherit from their parents and the environment in which they are raised (Plomin, DeFries, & Loehlin, 1977). Three forms of gene–environment correlation are typically distinguished: passive, active, and evocative. An example of passive gene–environment correlation is that more educated parents are likely to provide both beneficial genes and educationally supportive family environments, such as books in the home, for their children. Therefore, shared genes confound associations between putative environmental variables and child attainment. Active and evocative gene–environment correlations

reflect how genotypes lead to phenotypes: Individuals select and evoke environments on the basis of their genetically influenced traits.

It is essential to investigate gene–environment interplay in educational attainment for several reasons. First, educational attainment is an important trait for individuals

Corresponding Authors:

Thalia C. Eley, King's College London, Social, Genetic and Developmental Psychiatry Centre, Institute of Psychiatry, Psychology & Neuroscience, London, SE5 8AF, United Kingdom
 E-mail: thalia.eley@kcl.ac.uk

Gerome Breen, King's College London, Social, Genetic and Developmental Psychiatry Centre, Institute of Psychiatry, Psychology & Neuroscience, London, SE5 8AF, United Kingdom
 E-mail: gerome.breen@kcl.ac.uk

and society. Second, gene–environment correlation clearly matters for educational attainment. Adoption, twin, and instrumental-variables research suggest that shared genes largely explain associations between parent and child attainment (Holmlund, Lindahl, & Plug, 2011). Third, polygenic scores, which index the genetic liability that each individual carries for a specific trait, are notably powerful for educational attainment and now predict approximately 10% of the variation in years of education (Lee et al., 2018), **IAQ: 4** a finding with potential social implications (Plomin & von Stumm, 2018). However, it has been shown only recently that this prediction includes not only direct genetic effects on an individual's own education but also indirect genetic effects through relatives, that is, predicting the family environment (Bates et al., 2018; Kong et al., 2018).

Behavior-genetic study designs are needed to disentangle causal processes affecting educational attainment. Adoption studies do this by removing overlapping genetic and environmental influences (passive gene–environment correlation). This is achieved by measuring the resemblance of adopted children to their birth parents and to their adoptive parents. The former gives an estimate of direct genetic influence independent of passively correlated environmental effects. The latter gives an estimate of shared environmental influence, free of correlated genetic effects. Passive gene–environment correlation may be estimated as the extent to which genes contribute more to the covariation between measures of the family environment and offspring traits in nonadoptive than adoptive families (Plomin, Loehlin, & DeFries, 1985). Notably, other forms of gene–environment correlation are still present in adoptees, because heritable proclivities lead them to select and evoke experiences.

More recently, researchers have applied genomic tools to family data to estimate direct and indirect effects on educational attainment (Bates et al., 2018; Domingue, Belsky, Conley, Harris, & Boardman, 2015; Kong et al., 2018; Selzam et al., 2019; Wertz et al., 2018; Young et al., 2018). These designs are conceptually related to adoption designs, because they account for shared genes between parents and offspring. For example, genetic variants that were not passed on by parents can have only indirect effects on offspring traits, through genetically influenced parental behavior (Bates et al., 2018; Kong et al., 2018). When indirect effects were controlled for with a polygenic score for education based on nontransmitted variants, the variance explained by the transmitted score shrank from 5% to 2% (Kong et al., 2018). The nontransmitted score also independently predicted attainment. The family environment is an important contributor to polygenic-score prediction because it is adding to estimates of genetic influence and because parents still influence their offspring after analyses control for shared (transmitted) genes.

The present study drew on an unusually large and relatively unexplored sample of adoptees and harmonized a traditional quantitative genetic approach with modern genomic tools. Our main aim was to use the natural experiment created by adoptive placement to measure the importance of passive gene–environment correlation for educational attainment. When children are adopted by nonrelatives, the indirect genetic path between the rearing environment and their traits is not present because adoptive parents are not genetically related to adopted children. Three hypotheses follow. First, the phenotypic variance should be lower in adoptees compared with nonadopted individuals, because adoptees do not have the additional source of variance of passive gene–environment correlation (Loehlin & DeFries, 1987; Plomin, 1994). **IAQ: 5** This discrepancy could also be due to adoptive families varying less in socioeconomic status or being selected for perceived parenting ability (Natsuaki et al., 2019; Rutter, 2006). Second, if passive gene–environment correlation inflates heritability estimates, then heritability should be lower in adoptees than in nonadopted individuals, because adoptees are reared in environments that are less correlated with their genotypes. Third, for the same reason, the variance explained by polygenic scores will be lower in adoptees and may be closer to the direct genetic effect of an individual's own DNA.

Method

Sample, genotype quality control, and phenotype definition

The UK Biobank is an epidemiological resource including British individuals from the ages of 40 to 70 at recruitment (Allen, Sudlow, Peakman, Collins, & UK Biobank, 2014). UK Biobank participants were asked, “Were you adopted as a child?” In total, 7,407 individuals said “yes,” and 495,209 individuals said “no.” No additional information was collected on factors that are understood to reduce the representativeness of adoptees as a study sample: the age of adoption, whether the adoption was domestic or international, or whether individuals were adopted by biological relatives. Genome-wide genetic data came from the full release of the UK Biobank data and were collected and processed according to the quality-control pipeline (Bycroft et al., 2018). We restricted analyses to individuals with full phenotypic data for education, who also passed genotype quality-control criteria. This left 6,311 adopted and 375,343 nonadopted individuals for analysis.

Genotype quality-control criteria were common genetic variants of minor allele frequency greater than 0.01 that were directly genotyped or imputed with high confidence (INFO metric > 0.4), **IAQ: 6** individuals

with a genotype call rate greater than 98% who had concordant phenotypic and genetic gender information, and individuals who were unrelated to others in the data set (less than third-degree relatives). We removed relatives using a greedy algorithm to minimize the exclusion of adoptees. To reduce confounding from population stratification, we limited all analyses to individuals of European ancestry, as defined by four-means clustering on the first two genetic principal components provided by the UK Biobank. We also controlled for 10 ancestry principal components of the European sample in all genomic analyses.

Years of education, a proxy for educational attainment, was defined according to International Standard Classification of Education categories, as in previous genomic studies of the phenotype in UK Biobank and other samples (Lee et al., 2018). The response categories were as follows: none of the above (no qualifications) = 7 years of education; Certificate of Secondary Education (CSE) or equivalent = 10 years, O level/General Certificate of Secondary Education (GCSE) or equivalent = 10 years, A level/AS level or equivalent = 13 years, other professional qualification = 15 years, National Vocational Qualification (NVQ), Higher National Diploma (HNC), or equivalent = 19 years; college or university degree = 20 years.

Statistical analyses

Phenotypic comparisons. First, we formally tested the hypothesis that nonadopted individuals show greater phenotypic variance than adopted individuals because of the presence of an additional source of variance (passive gene–environment correlation). A nonparametric test was used given the nonnormal distribution of the years-of-education variable (Brown & Forsythe, 1974). This test is based on absolute deviations from the median rather than the group mean. We also tested for differences in years of education, age, and sex between the two groups using a Wald test, z test, and Wilcoxon test, respectively.

Single-nucleotide-polymorphism (SNP) heritability estimation. Second, to test the hypothesis that heritability is lower in adoptees, whose rearing environments are less correlated with their genotypes, we estimated the variance explained by common genetic variants for years of education in adoptees using genomic-relatedness-based restricted maximum likelihood (GREML; Yang, Lee, Goddard, & Visscher, 2011) and compared this with the heritability estimate for nonadopted individuals. This method estimates heritability as the extent to which genetic similarity among unrelated individuals can predict their trait similarity. In GREML, a matrix of genomic similarity for each pair of unrelated individuals across genotyped variants is compared with a matrix of their pairwise phenotypic similarity

using a random-effects mixed linear model. This allows the variance of a trait to be decomposed into genetic and residual components using maximum likelihood. We used two genetic-relatedness matrices: one for adopted individuals and a second for a subset of 6,500 nonadopted individuals. This was to enable comparison of two similarly sized samples and to reduce the computational burden that results from scaling GREML to a sample as large as the UK Biobank. For both genomic matrices, we used a relatedness cutoff of 0.025. Subsamples were made using the “sample_n” function in the *dplyr* package in the R programming environment (Version 3.5; R Core Team, 2019). We compared these results with heritability estimates derived from a second method, linkage-disequilibrium score regression (LDSC; Bulik-Sullivan et al., 2015). Unlike GREML, LDSC does not require individual-level data, allowing it to be computationally feasible to estimate the heritability of education in the full sample of nonadopted individuals. LDSC also enabled us to estimate genetic correlations (see below).

Polygenic scoring. Third, we tested whether the power of polygenic scores was greater for individuals who were reared with their biological relatives than for adoptees. The sample of nonadopted individuals was subdivided into three independent groups for polygenic-score analyses. Our first sample consisted of 318,843 nonadopted individuals for genome-wide-association (GWA) analysis. The purpose of this was to estimate the effect sizes of associations between genome-wide genetic variants and years of education for the creation of individual-level polygenic scores. We derived our base summary statistics file for years of education by meta-analyzing summary statistics from our own GWA analysis in this subsample with independent summary statistics obtained from the Social Science Genomics Consortium (excluding UK Biobank and 23andMe; Lee et al., 2018). The sample size for these external summary statistics was 324,160, leading to a total sample size of 643,003 individuals in our GWA meta-analysis.

Our second independent sample included 50,000 individuals for training our polygenic scores for years of education, that is, identifying the optimal p -value threshold for inclusion of SNPs. The standard set of p values in PRSice 2 software was tested: .001, .05, .1, .2, .3, .4, .5, 1 (Choi & O'Reilly, 2019).

Our third independent sample contained 6,500 individuals, to match the sample size of adopted individuals, on which to run polygenic-prediction models. In these prediction models, we regressed the years-of-education phenotype in the UK Biobank on polygenic scores for years of education in adoptees and then repeated the analysis in the 6,500 nonadopted individuals. In this third set of analyses, we used a set p -value threshold obtained from the training step. This exact

sample was the same as the one used to estimate SNP heritability of years of education. Notably, this polygenic-score analysis is better powered than the SNP heritability analysis, because it capitalizes on the power of the large discovery sample ($N = 643,003$).

As a negative control, we tested the polygenic-prediction comparison between adoptees and nonadopted individuals for height, which has not shown evidence of passive gene–environment correlation in previous studies (Kong et al., 2018; Selzam et al., 2019). As with the education analysis, we trained the polygenic score on the basis of the largest independent association study (Wood et al., 2014) in the sample of 50,000 individuals; we then tested the prediction at the best p -value threshold in our two independent and similarly sized samples of adopted and nonadopted individuals.

Supplemental analyses.

Heritability of adoption status. Substantial heritability of our environmental moderator might have affected the interpretation of our main results. To explore this, we also estimated the heritability of adoption status using LDSC in the full sample ($N = 381,654$). The genetic “influence” on adoption largely arises in the biological-parent generation because heritable traits influence the likelihood of adoption of their child.

Polygenic-score-by-adoption interaction analyses. Differences in genetic influences on the same trait across contexts—in this case, adoption—can also be conceptualized as gene–environment interaction, whereby the impact of genes on educational attainment may be contingent on adoption status. We aimed to further explore our main results by testing a formal polygenic-score-by-adoption interaction regression model. The model contained main effects for polygenic score for years of education, adoption, and covariates, plus the interaction term, as well as interaction terms for polygenic score and adoption with each covariate (Keller, 2014). We tested a linear model for additive interaction and a logistic model for multiplicative interaction. To visualize any interaction, we plotted the regression slopes for polygenic prediction of educational attainment for adopted and nonadopted individuals (with both variables standardized to have a mean of 0 and a standard deviation of 1). Additionally, we stratified polygenic scores for adopted and nonadopted individuals overall ($N = 12,811$) into deciles and tested for mean differences in years of education between adopted and nonadopted groups in each decile.

Qualitative differences in the genetic influence on education by adoption status. We assessed whether education is driven by the same set of genetic influences in adopted and nonadopted individuals. First, we estimated the genetic correlation between years of education in our

samples. For this, we ran GWA analyses of years of education in the full sample of nonadopted individuals ($N = 375,343$) and in the sample of adoptees, and then estimated the genetic correlation between them using LDSC. Second, we tested whether education is genetically linked to different traits between adoptees and nonadopted individuals. To this end, we estimated genetic correlations between years of education and 247 traits, available on LD Hub (Zheng et al., 2017), for both adopted and nonadopted individuals. We compared the magnitudes of genetic correlations between years of education and other variables between the adopted and nonadopted samples with z tests.

Birth-year-related differences in genetic influence. During the period when UK Biobank participants were growing up (1930s–1970s), access to education increased, and there was great change in the norms and regulations surrounding reproduction, contraception, and adoption. Previous studies have found that genetic influence on years of education increased in this period in the United Kingdom, because environmental differences between people had less influence on whether they stayed in education (Lee et al., 2018). We investigated temporal change in patterns of genetic influence on education in adopted compared with nonadopted individuals by stratifying polygenic-prediction analyses according to year of birth. Specifically, all individuals were split into seven mutually exclusive birth-year groups, each with a range of 5 years, and polygenic-score analyses were conducted separately for each of the year groups.

All analyses (SNP heritability, polygenic scoring, GWA) controlled for the following covariates: sex, age, 10 ancestry principal components, and factors capturing genotyping batch and center. The majority of the analyses were completed in R Version 3.5. GREML was performed in the GCTA software (Yang et al., 2011). GWA meta-analysis was performed in METAL (Willer, Li, & Abecasis, 2010). Polygenic-score analyses were performed in PRSice 2 (Choi & O'Reilly, 2019). To compare polygenic-score results between adopted and nonadopted individuals, we obtained bootstrapped standard errors for the R^2 statistics using the boot package in R, with 1,000 replications. Genome-wide genetic correlations were estimated using LDSC (Bulik-Sullivan et al., 2015) and LD Hub (Zheng et al., 2017). The UK Biobank is a controlled-access public data set available to all bona fide researchers.

Results

Sample analyzed

The total sample of individuals with education-phenotype data and quality-controlled genotype data was

Table 1. Descriptive Statistics for Adopted and Nonadopted Individuals Included in the Study

Variable	Adopted (<i>n</i> = 6,311)	Nonadopted (<i>n</i> = 375,343)	Nonadopted (Group d; <i>n</i> = 6,500)
Age	<i>M</i> = 56.4, <i>SD</i> = 8.53	<i>M</i> = 56.7, <i>SD</i> = 8.01	<i>M</i> = 56.7, <i>SD</i> = 8.06
Sex	48% male (<i>n</i> = 3,010)	46% male (<i>n</i> = 172,706)	46% male (<i>n</i> = 2,978)
Years of education			
7	1,209 (19%)	62,651 (17%)	1,064 (16%)
10	1,780 (28%)	100,210 (27%)	1,709 (26%)
13	749 (12%)	43,448 (12%)	755 (12%)
15	350 (6%)	19,428 (5%)	354 (5%)
19	433 (7%)	24,300 (6%)	428 (7%)
20	1,790 (28%)	125,306 (33%)	2,190 (34%)

Note: For years of education, both *ns* and percentages of the subsample are given. Adoptees were compared with the full sample of nonadopted individuals and with our smaller subsample used for genomic analyses (group d). **[AQ: 7]**

381,654. As described in the Method section, individuals were split into four mutually exclusive groups: (a) 6,311 individuals adopted as children, (b) 318,843 nonadopted individuals for GWA analysis, (c) 50,000 nonadopted individuals for training of polygenic scores, and (d) 6,500 nonadopted individuals for genomic analyses for comparison with adoptees. Nonadopted individuals were randomly placed into groups b, c, and d.

Phenotypic results

Phenotypic differences between adoptees and non-adopted individuals were generally modest in size but because of the large sample size in this study, several were statistically significant (see Table 1). We found that there was significantly greater variance in years of education for nonadopted individuals than for adoptees (26.2 vs. 25.8, respectively; $p = .002$ compared with 6,500 nonadopted individuals in group d; $p = 3.2 \times 10^{-5}$ compared with all nonadopted individuals). Table 1 provides descriptive statistics for years of education, age, and sex in the two groups. Adoptees in the UK Biobank were significantly younger on average ($p = .026$ compared with group d; $p = .009$ compared with all nonadopted individuals), although point estimates were similar (56.4 vs. 56.7). There were significantly more males in the adopted group ($p = .033$ compared with group d; $p = .008$ compared with all nonadopted individuals), but the magnitude of the difference was small (48% vs. 46% male). Adoptees had significantly fewer years of education ($p = 3.3 \times 10^{-11}$ compared with group d; $p < 2.2 \times 10^{-16}$ compared with all nonadopted individuals in the UK Biobank). This is also reflected in the lower percentage of college attendees (20 years of education in Table 1) in the adopted group (28% compared with 33% in the nonadopted group). All

comparison results were consistent between the large and small samples of nonadopted individuals.

SNP heritability estimates

Figure 1 shows GREML-derived SNP heritability estimates for years of education in both adopted individuals and nonadopted individuals (left-hand bars). The estimate of heritability was larger in individuals reared with their relatives (0.29, $SE = 0.079$), compared with adopted individuals (0.23, $SE = 0.079$). However, confidence intervals were wide and overlapped, so the difference in heritability was not significant.

It was not computationally feasible to estimate the heritability of education using all nonadopted individuals with GREML. Notably, though, the LDSC-derived heritability was 0.17 ($SE = 0.005$) in the full sample of nonadopted individuals ($N = 375,343$) and 0.14 ($SE = 0.073$) for adoptees, corroborating the pattern of results found using GREML. LDSC estimates are typically lower than GCTA and GREML-derived estimates (Evans & Keller, 2018).

Polygenic-prediction results

Figure 1 also shows that twice as much phenotypic variance in years of education was explained by polygenic scores for years of education in nonadopted individuals (0.074) as in adoptees (0.037). This difference was highly significant ($p = 8.23 \times 10^{-24}$). The optimal significance threshold (p) for inclusion of SNPs was 1 (see Table S1 in the Supplemental Material available online). Table S2 in the Supplemental Material shows the full results from the polygenic-prediction analyses.

For our negative control analysis of height, as expected, the variance explained by the polygenic

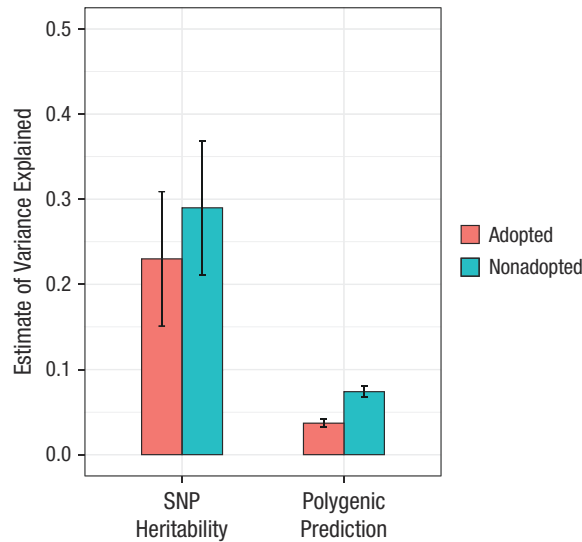


Fig. 1. Estimates of the variance explained by common single-nucleotide polymorphisms (SNPs) for years of education and by polygenic scores for education, separately for adoptees and individuals reared with their relatives. Error bars show 95% confidence intervals (CIs). Sample sizes for polygenic-prediction analyses were 6,311 and 6,500 for adopted and nonadopted individuals, respectively; sample sizes for genomic-relatedness-based restricted-maximum-likelihood (GREML) heritability analyses were lower (6,227 for adopted and 6,362 for nonadopted individuals) because relatives were removed at a cutoff of > 0.025 . For polygenic-score results, CIs were obtained by bootstrapping with 1,000 replications. **[AQ: 8]**

score in adoptees (0.127 , $SE = 0.008$) versus nonadopted individuals (0.134 , $SE = 0.008$) was not significantly different ($p = .62$). The optimal significance threshold (p) for inclusion of SNPs in the polygenic score was .001.

Supplemental analyses

Heritability of being adopted. We found a liability-scale SNP heritability of being adopted of 0.059 ($SE = 0.004$), assuming the population prevalence of adoption is identical to the sample prevalence (1.7%). If the actual population prevalence differed and was, for example, 0.7% or 2.7%, the liability-scale SNP heritability would become 0.047 ($SE = 0.002$) or 0.066 ($SE = 0.005$), respectively. Adoption status showed significant genetic correlations with education, age at first birth, depression, and obesity after analyses corrected for multiple testing (see Table S5 in the Supplemental Material), although these correlations should be viewed with caution given the low SNP heritability of adoption. Adoption status could be significantly predicted by the polygenic score for years of education ($R^2 = .008$, $p < 2 \times 10^{-16}$). The heritability of adoption is low but may have confounded our between-groups comparisons.

Polygenic-score-by-adoption interaction. We tested a formal interaction model to further examine the finding that genetic influences on education are weaker in the sample of adoptees. The interaction between polygenic score and adoption status in predicting years of education is shown in Figure 2. The regression slope is significantly steeper in the nonadopted group, indicating that years of education increases more as polygenic scores for education increase in this group compared with the adopted group. See Table S3 in the Supplemental Material for results of the full interaction model. Using linear regression, we confirmed that polygenic prediction of education interacts beyond additivity with adoption status (interaction estimate = -0.33 ; $p = 2.66 \times 10^{-4}$). This means that polygenic scores had a smaller association with education in adoptees. Then, using logistic regression instead of linear regression, we also found that the interaction exceeded multiplicativity (interaction estimate = -0.18 ; $p = .0009$). The finding of interaction exceeding both additive and multiplicative models means that the combined effect of polygenic score for education and adoption status is not scale dependent and is greater than either the sum or product of their individual effects, respectively.

To further explore the interaction, we plotted the mean years of education per decile of polygenic score for years of education, separately for adopted and nonadopted individuals. Figure 3 shows that for individuals

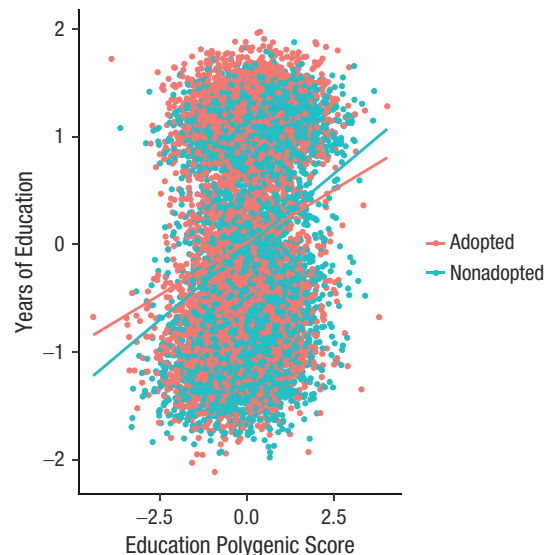


Fig 2. Results of the regression of years of education on polygenic score for education, separately for 6,311 adoptees and 6,500 nonadopted individuals. Best-fitting regression lines are shown for each group, and each data point represents one individual. The two clusters of data points reflect the distinct groups of individuals who did and did not attend university.

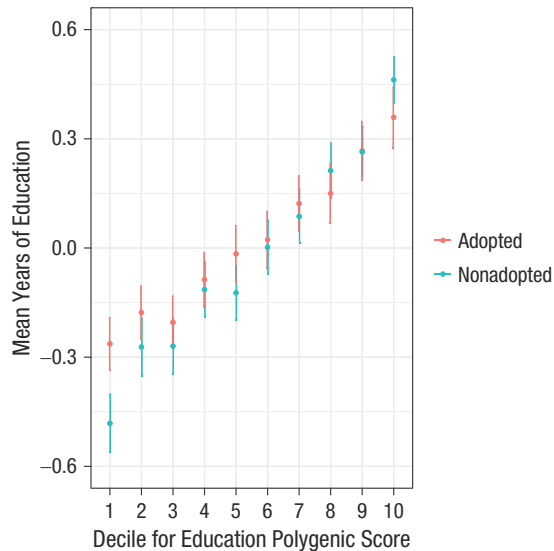


Fig. 3. Mean years of education (standardized) per decile of polygenic score for years of education (standardized), separately for adopted and nonadopted individuals. Error bars show 95% confidence intervals.

in the lowest decile of polygenic score for education, those who were adopted as children achieved substantially higher mean years of education (-0.24 , $SE = 0.03$; standardized) compared with nonadopted individuals (-0.40 , $SE = 0.03$). This mean difference between adopted and nonadopted individuals was significant in the bottom decile ($p = 7.05 \times 10^{-5}$) but not for other deciles of polygenic load. See Table S4 in the Supplemental Material for full results of the decile analysis.

Qualitative differences in genetic influences according to adoptee status. We found that largely the same genetic influences were operating on education regardless of adoption status. First, the genetic correlation between years of education in adopted and nonadopted individuals was not significantly different from 1, 0.81 ($SE = 0.21$). Second, we found no evidence that educational attainment was associated with different traits in individuals who were adopted. Figure S1 in the Supplemental Material presents estimates of genetic correlations between years of education and 247 external traits, separately for the adopted and nonadopted samples. None of these were significantly different between adoptees and nonadopted individuals after multiple-testing correction. Because of the relatively small sample of adopted individuals, these results should be interpreted with caution.

Year-of-birth stratification analysis. Our final sensitivity analysis tested for differences according to year of

birth in polygenic prediction from direct effects (indicated by the variance explained in the adoptees) versus from passive gene–environment correlation (indicated by the difference in variance explained between nonadopted individuals and adoptees). We found small, nonsignificant differences in the variance explained by polygenic scores for education depending on the year-of-birth group considered. Figure S2 in the Supplemental Material shows that polygenic prediction remained generally stable for the adoptees across generations at approximately 0.04 , and any differences between age strata were nonsignificant. We note that subsampling reduced the statistical power to detect differences within and between groups across time. See Table S6 in the Supplemental Material for sample sizes of each year-of-birth group.

Discussion

Cumulatively, our findings suggest that the family environment provided by relatives plays an important role in the manifestation of genetic effects on education. **[AQ: 9]** Polygenic scores explained significantly less variance in the educational attainment of individuals who were adopted away from their parents as children ($R^2 = .04$ vs. $.07$; difference test: $p = 8.23 \times 10^{-24}$). The variance explained by polygenic scores in years of education in adoptees ($.04$) **[AQ: 10]** approximated the prediction from the direct effects of individuals' own DNA. The difference between the variance explained in nonadopted individuals and adoptees suggests that about half of the predictive power of polygenic scores for educational attainment comes from passive gene–environment correlation. We also found that individuals in the lowest decile of polygenic score attained significantly more years of education if they were adopted.

By showing that polygenic scores for education are twice as powerful in nonadopted individuals compared with adoptees, we suggest that genetic influence on educational attainment is magnified when individuals are reared by their close genetic relatives, with whom they share both genes and environments. Our results agree with recent evidence showing that the effects of passive gene–environment correlation reduced the variance explained by polygenic scores by 30% to 50% (Kong et al., 2018; Selzam et al., 2019).

Following other recent research (Kong et al., 2018; Selzam et al., 2019), we use the term *direct genetic effect* to refer to the effect of the polygenic score among adoptees after analyses control for passive gene–environment correlation. However, it cannot be assumed that estimating the direct effect of a polygenic score is tantamount to isolating a straightforward purely genetic effect or a genetic propensity. Genetic effects are never truly direct but are always behaviorally mediated and

expressed in the context of an environment. Active and evocative gene–environment correlation mechanisms are essential in how genes influence traits in everyone, including adoptees (Plomin, 2014), and these are included in estimates of direct genetic influence.

Our observation that individuals in the lowest decile of polygenic score for education attain significantly more education if they are adopted could be due to educationally supportive adoptive environments. This agrees with previous evidence showing that adoptees had higher school achievement and intelligence-test scores than nonadopted siblings or peers who stayed with their birth family (van IJzendoorn, Juffer, & Poelhuis, 2005) and that such advantages are retained in their adult qualifications (Maughan, Collishaw, & Pickles, 1998). The specificity of this result to adoptees in the lowest decile of polygenic score links to previous evidence that the “boosting” effect may be stronger in higher-socioeconomic-status adoptive families and for children rescued from poverty (Duyme, Dumaret, & Tomkiewicz, 1999; Turkheimer, 1991). This environmental effect of adoptive parents might suggest that efforts to help individuals stay in education can be effective for those with less genetic propensity for education.

These results should be viewed in light of several limitations. First, interpreting genetic influence in adoptees as direct and free of passive gene–environment correlation requires that close biological relatives were not involved in the education of the adoptees. Unfortunately, the UK Biobank contains no information about the age of adoption beyond that it occurred in childhood, nor does it contain information about whether individuals were adopted by relatives or were able to identify and contact their biological parents. This knowledge would have allowed us to exclude individuals who were not solely socialized with adoptive families and therefore to make a precise comparison with individuals who were reared with their birth parents. However, polygenic prediction of education still differed markedly between the two groups, even though adoptees may have been in contact with their biological relatives. Thus, the effects of passive gene–environment correlation may contribute even more than half of the predictive power of polygenic scores for education, as we estimated here.

A second caveat is lack of generalizability. The UK Biobank is not representative of the general population, because there is “healthy and wealthy” (AQ: 11) volunteer selection (Fry et al., 2017; Keyes & Westreich, 2019), and we have analyzed data only on individuals with European ancestry. Adoptees may not be random samples of the population. Indeed, adoption status is not a purely random environmental exposure but is

slightly heritable (our estimate is 6%), and this may confound our between-groups comparisons. Moreover, because detailed adoption data are lacking in the UK Biobank, we were unable to conduct sensitivity analyses adjusting for aspects of adoption that have known associations with important outcomes, including school performance. For example, it would have been useful to know whether adoptions were domestic or international or occurred in the context of childhood adversity or institutionalization (Howard, Smith, & Ryan, 2004).

Similarly, adoptive parents tend to differ systematically from other parents: They are likely to be more educated, more socially advantaged, and to have lower rates of psychopathology (Rutter, 2006). A recent U.S. study found that children are adopted into households that differ in average parental education compared with biological children (Domingue & Fletcher, 2019). This probably applies to the UK Biobank, although we cannot be certain because of the lack of parental data. If adoptive families are more homogeneous with respect to these characteristics, environmental variance may contribute less to differences in educational attainment among adoptees, and trait heritability estimates are consequently likely to be higher. However, the fact that lower environmental variance may act to inflate genetic influence in adoptees compared with nonadopted individuals makes our finding of significantly higher polygenic prediction in nonadopted individuals more striking. Again, the effects of passive gene–environment correlation for education may be even greater than we estimate.

There are several advantages of using the present adoption design for distinguishing direct genetic influence from passive gene–environment correlation. Unlike other methods, our approach does not require intergenerational data, which is valuable but has its own issues, such as cohort differences in genetic effects. Analyzing the adoptees in the UK Biobank also bypasses several limitations of traditional adoption studies, including low sample size and reliance on weak indirect proxies for inherited load for specific traits (birth-parent trait status rather than individual-level polygenic scores). However, future progress in understanding the mechanisms driving the transmission of educational attainment will require intergenerational, longitudinal, and genetically informative data sets, including detailed characterization of the home environment. A developmental approach is useful, because gene–environment correlations likely arise early in childhood, and there will be complex reciprocal effects across time. Researchers have already started to pinpoint genetically influenced aspects of families that are associated with children’s polygenic scores for education (Krapohl et al., 2017; Wertz et al., 2019; Wertz et al., 2018).

The evidence presented in this article highlights the importance of the family environment to causal mechanisms influencing individual differences in educational attainment. These [AQ: 12] can be through possessing genes that shape the educational environment provided for offspring that also directly influence attainment in the child or through providing an educationally supportive environment for an adopted child.

Transparency

Action Editor: Steven W. Gangestad

Editor: D. Stephen Lindsay

Author Contributions

T. C. Eley and G. Breen are joint senior authors of this article. R. Cheesman and G. Breen conceived and designed the study. Genotype-data quality control was conducted by A. Hunjan and J. R. I. Coleman. R. Cheesman analyzed the data and wrote the manuscript. All of the authors provided critical revisions and approved the final version of the manuscript for submission.

Declaration of Conflicting Interests

The author(s) declared that there were no conflicts of interest with respect to the authorship or the publication of this article. [AQ: 13]

Open Practices

All data used in the present study are available to qualified researchers from the UK Biobank at <http://www.ukbiobank.ac.uk/>. The design and analysis plans for this study were not preregistered. [AQ: 14]

Funding

This research was conducted using the UK Biobank resource under Application 18177 (with thanks to P. F. O'Reilly). [AQ: 15] This study represents independent research partially funded by the National Institute for Health Research (NIHR) Biomedical Research Centre at the South London and Maudsley National Health Service (NHS) Foundation Trust and by King's College London. The views expressed are those of the authors and not necessarily those of the NHS, the NIHR, or the Department of Health and Social Care. High-performance computing facilities were funded with capital equipment grants from the Guy's and St. Thomas' Charity (TR130505) and Maudsley Charity (980). T. C. Eley was partially funded by a program grant from the UK Medical Research Council (MR/M021475/1). R. Plomin was supported by a Medical Research Council Professorship award (G19/2). T. A. McAdams and Y. Ahmadzadeh were funded by a Sir Henry Dale Fellowship awarded to T. A. McAdams, jointly funded by the Wellcome Trust and the Royal Society (Grant No. 107706/Z/15/Z). R. Cheesman was supported by an Economic and Social Research Council studentship.

ORCID iD

Rosa Cheesman  <https://orcid.org/0000-0002-6543-0402>

Acknowledgments

We thank the scientists involved in the construction of the UK Biobank and all of the participants who shared their life

experiences with investigators in the UK Biobank. We thank Aysu Okbay and the Social Science Genetic Association Consortium for providing genome-wide-association summary statistics for educational attainment excluding UK Biobank.

Supplemental Material

Additional supporting information can be found at <http://journals.sagepub.com/doi/suppl/10.1177/0956797620904450>

References

- Allen, N. E., Sudlow, C., Peakman, T., & Collins, R. UK Biobank. (2014). UK Biobank data: Come and get it. *Science Translational Medicine*, 6(224), Article 224ed4. doi:10.1126/scitranslmed.3008601
- Bates, T. C., Maher, B. S., Medland, S. E., McAloney, K., Wright, M. J., Hansell, N. K., . . . Gillespie, N. A. (2018). The nature of nurture: Using a virtual-parent design to test parenting effects on children's educational attainment in genotyped families. *Twin Research and Human Genetics*, 21(2), 73–83. doi:10.1017/thg.2018.11
- Brown, M. B., & Forsythe, A. B. (1974). Robust tests for the equality of variances. *Journal of the American Statistical Association*, 69, 364–367. doi:10.1080/01621459.1974.10482955
- Bulik-Sullivan, B., Finucane, H. K., Anttila, V., Gusev, A., Day, F. R., Loh, P.-R., . . . Neale, B. M. (2015). An atlas of genetic correlations across human diseases and traits. *Nature Genetics*, 47, 1236–1241. doi:10.1038/ng.3406
- Bulik-Sullivan, B. K., Loh, P.-R., Finucane, H. K., Ripke, S., & Yang, J., Schizophrenia Working Group of the Psychiatric Genomics Consortium, . . . Neale, B. M. (2015). LD score regression distinguishes confounding from polygenicity in genome-wide association studies. *Nature Genetics*, 47, 291–295. doi:10.1038/ng.3211
- Bycroft, C., Freeman, C., Petkova, D., Band, G., Elliott, L. T., Sharp, K., . . . Marchini, J. (2018). The UK Biobank resource with deep phenotyping and genomic data. *Nature*, 562, 203–209. doi:10.1038/s41586-018-0579-z
- Choi, S. W., & O'Reilly, P. F. (2019). PRSice-2: Polygenic risk score software for biobank-scale data. *GigaScience*, 8(7), Article giz082. doi:10.1093/gigascience/giz082 [AQ: 16]
- Domingue, B. W., & Fletcher, J. (2019). Separating measured genetic and environmental effects: Evidence linking parental genotype and adopted child outcomes. *BioRxiv*. doi:10.1101/698464
- Domingue, B. W., Belsky, D., Conley, D., Harris, K. M., & Boardman, J. D. (2015). Polygenic influence on educational attainment: New evidence from the National Longitudinal Study of Adolescent to Adult Health. *AERA Open*, 1(3). doi:10.1177/2332858415599972
- Duyme, M., Dumaret, A.-C., & Tomkiewicz, S. (1999). How can we boost IQs of “dull children”? A late adoption study. *Proceedings of the National Academy of Sciences, USA*, 96, 8790–8794.
- Evans, L. M., & Keller, M. C. (2018). Using partitioned heritability methods to explore genetic architecture. *Nature Reviews Genetics*, 19, Article 185. doi:10.1038/nrg.2018.6
- Fry, A., Littlejohns, T. J., Sudlow, C., Doherty, N., Adamska, L., Sprosen, T., . . . Allen, N. E. (2017). Comparison of

- sociodemographic and health-related characteristics of UK Biobank participants with those of the general population. *American Journal of Epidemiology*, 186, 1026–1034.
- Holmlund, H., Lindahl, M., & Plug, E. (2011). The causal effect of parents' schooling on children's schooling: A comparison of estimation methods. *Journal of Economic Literature*, 49, 615–651. doi:10.1257/jel.49.3.615
- Howard, J. A., Smith, S. L., & Ryan, S. D. (2004). A comparative study of child welfare adoptions with other types of adopted children and birth children. *Adoption Quarterly*, 7(3), 1–30.
- Keller, M. C. (2014). Gene $\times$ Environment interaction studies have not properly controlled for potential confounders: The problem and the (simple) solution. *Biological Psychiatry*, 75, 18–24. doi:10.1016/j.biopsych.2013.09.006
- Keyes, K. M., & Westreich, D. (2019). UK Biobank, big data, and the consequences of non-representativeness. *The Lancet*, 393, 1297. doi:10.1016/S0140-6736(18)33067-8
- Kong, A., Thorleifsson, G., Frigge, M. L., Vilhjalmsdottir, B. J., Young, A. I., Thorgeirsson, T. E., . . . Stefansson, K. (2018). The nature of nurture: Effects of parental genotypes. *Science*, 359, 424–428. doi:10.1126/science.aan6877
- Krapohl, E., Hannigan, L. J., Pingault, J. B., Patel, H., Kadeva, N., Curtis, C., . . . Plomin, R. (2017). Widespread covariation of early environmental exposures and trait-associated polygenic variation. *Proceedings of the National Academy of Sciences, USA*, 114, 11727–11732. doi:10.1073/pnas.1707178114
- Lee, J. J., Wedow, R., Okbay, A., Kong, E., Maghzian, O., Zacher, M., . . . Karlsson Linnér, R. (2018). Gene discovery and polygenic prediction from a genome-wide association study of educational attainment in 1.1 million individuals. *Nature Genetics*, 50, 1112–1121. doi:10.1038/s41588-018-0147-3
- Loehlin, J. C., & De Fries, J. C. (1987). Genotype-environment correlation revisited. *Behavior Genetics*, 17, 731–732. **[AQ: 17]**
- Maughan, B., Collishaw, S., & Pickles, A. (1998). School achievement and adult qualifications among adoptees: A longitudinal study. *Journal of Child Psychology and Psychiatry, and Allied Disciplines*, 39, 669–685. doi:10.1111/1469-7610.00367
- Natsuaki, M. N., Neiderhiser, J. M., Harold, G. T., Shaw, D. S., Reiss, D., & Leve, L. D. (2019). Siblings reared apart: A sibling comparison study on rearing environment differences. *Developmental Psychology*, 55, 1182–1190. doi:10.1037/dev0000710
- Plomin, R. (1994). *Genetics and experience: The interplay between nature and nurture*. Thousand Oaks, CA: SAGE.
- Plomin, R. (2014). Genotype-environment correlation in the era of DNA. *Behavior Genetics*, 44, 629–638. doi:10.1007/s10519-014-9673-7
- Plomin, R., DeFries, J. C., Knopik, V. S., & Neiderhiser, J. M. (2016). Top 10 replicated findings from behavioral genetics. *Perspectives on Psychological Science*, 11, 3–23. doi:10.1177/1745691615617439
- Plomin, R., DeFries, J. C., & Loehlin, J. C. (1977). Genotype-environment interaction and correlation in the analysis of human behavior. *Psychological Bulletin*, 84, 309–322. doi:10.1037/0033-2909.84.2.309
- Plomin, R., Loehlin, J. C., & DeFries, J. C. (1985). Genetic and environmental components of “environmental” influences. *Developmental Psychology*, 21, 391–402. doi:10.1037/0012-1649.21.3.391
- Plomin, R., & von Stumm, S. (2018). The new genetics of intelligence. *Nature Reviews Genetics*, 19, 148–159. doi:10.1038/nrg.2017.104
- R Core Team. (2019). *R: A language and environment for statistical computing*. Retrieved from <http://www.R-project.org>
- Rutter, M. (2006). *Genes and behavior: Nature-nurture interplay explained*. Hoboken, NJ: Wiley.
- Selzam, S., Ritchie, S. J., Pingault, J.-B., Reynolds, C. A., O'Reilly, P. F., & Plomin, R. (2019). Comparing within- and between-family polygenic score prediction. *The American Journal of Human Genetics*, 105, 351–363. doi:10.1016/j.ajhg.2019.06.006 **[AQ: 18]**
- Turkheimer, E. (1991). Individual and group differences in adoption studies of IQ. *Psychological Bulletin*, 110, 392–405.
- van IJzendoorn, M. H., Juffer, F., & Poelhuis, C. W. K. (2005). Adoption and cognitive development: A meta-analytic comparison of adopted and nonadopted children's IQ and school performance. *Psychological Bulletin*, 131, 301–316. doi:10.1037/0033-2909.131.2.301
- Wertz, J., Belsky, J., Moffitt, T. E., Belsky, D. W., Harrington, H., Avinun, R., . . . Caspi, A. (2019). Genetics of nurture: A test of the hypothesis that parents' genetics predict their observed caregiving. *Developmental Psychology*, 55, 1461–1472. doi:10.1037/dev0000709
- Wertz, J., Moffitt, T. E., Agnew-Blais, J., Arseneault, L., Belsky, D. W., Corcoran, D. L., . . . Caspi, A. (2018). Using DNA from mothers and children to study parental investment in children's educational attainment. *Child Development*. Advance online publication. doi:10.1111/cdev.13329 **[AQ: 19]**
- Willer, C. J., Li, Y., & Abecasis, G. R. (2010). METAL: Fast and efficient meta-analysis of genomewide association scans. *Bioinformatics*, 26, 2190–2191. doi:10.1093/bioinformatics/btq340
- Wood, A. R., Esko, T., Yang, J., Vedantam, S., Pers, T. H., Gustafsson, S., . . . Kutalik, Z. (2014). Defining the role of common variation in the genomic and biological architecture of adult human height. *Nature Genetics*, 46, 1173–1186. doi:10.1038/ng.3097
- Yang, J., Lee, S. H., Goddard, M. E., & Visscher, P. M. (2011). GCTA: A tool for genome-wide complex trait analysis. *The American Journal of Human Genetics*, 88, 76–82. doi:10.1016/j.ajhg.2010.11.011
- Young, A. I., Frigge, M. L., Gudbjartsson, D. F., Thorleifsson, G., Bjornsdottir, G., Sulem, P., . . . Kong, A. (2018). Relatedness disequilibrium regression estimates heritability without environmental bias. *Nature Genetics*, 50, 1304–1310. doi:10.1038/s41588-018-0178-9
- Zheng, J., Erzurumluoglu, A. M., Elsworth, B. L., Kemp, J. P., Howe, L., Haycock, P. C., . . . Neale, B. M. (2017). LD Hub: A centralized database and web interface to perform LD score regression that maximizes the potential of summary level GWAS data for SNP heritability and genetic correlation analysis. *Bioinformatics*, 33, 272–279. doi:10.1093/bioinformatics/btw613

Chapter 5. How important are parents in the development of child anxiety and depression? A genomic analysis of parent-offspring trios in the Norwegian Mother Father and Child Cohort Study (MoBa)

This chapter is a manuscript that is in preparation for peer-review.

Supplementary Information is included in Appendix 4.

Rosa Cheesman,¹ Espen Moen Eilertsen,^{2,3} Yasmin I. Ahmadzadeh,¹ Line C. Gjerde,^{2,3} Laurie J. Hannigan,^{4,5} Alexandra Havdahl,^{2,3,4,5} Alexander I. Young,^{1,6} Thalia C. Eley,^{1,7} Eivind Ystrom,^{2,3,8*} Tom A. McAdams^{1,2*}

¹Social Genetic & Developmental Psychiatry Centre, Institute of Psychiatry, Psychology & Neuroscience, King's College London, UK

²PROMENTA Research Center, Department of Psychology, University of Oslo, Oslo, Norway

³Department of Mental Disorders, Norwegian Institute of Public Health, Oslo, Norway

⁴Nic Waals Institute at Lovisenberg Diaconal Hospital, Oslo, Norway

⁵MRC Integrative Epidemiology Unit, University of Bristol, Bristol, UK

⁶Center for Economic and Social Research, University of Southern California, Los Angeles, CA, USA

⁷NIHR Maudsley Biomedical Research Centre; South London and Maudsley NHS Trust, UK

⁸School of Pharmacy, University of Oslo

*Joint senior authors

Abstract

Background: Many studies detect associations between parent behaviour and child symptoms of anxiety and depression. However, most do not account for the fact that parents and children may resemble each other because they share genetic risk. Quantitative genetic designs provide a means of controlling for shared genes, but rely on observed putative exposure variables, and require data from highly specific family structures.

Methods: A new intergenerational genomic method, Relatedness Disequilibrium Regression (RDR), indexes environmental effects of parents on a child trait using measured DNA. RDR estimates how much the parent genome influences the child indirectly via the environment, over and above effects of genes acting directly in the child. This ‘genetic nurture’ effect is agnostic to parent phenotype and captures unmeasured heritable parent behaviours. We applied RDR in a sample of 11,598 parent-offspring trios from the Norwegian Mother, Father and Child Cohort Study (MoBa) to estimate indirect parent genetic effects (genetic nurture) separately from direct child genetic effects on anxiety and depression symptoms at age 8. We also tested for mediation of genetic nurture via maternal emotional symptoms. Results were compared to a complementary non-genomic pedigree model.

Results: Parental genetic nurture influenced depression symptoms at age 8, explaining 14% of the phenotypic variance. Subsequent analyses suggested that maternal anxiety and depression partially mediated the parental genetic nurture effect. The genetic nurture effect was mirrored by the finding of shared family environmental influence in our complementary pedigree model. In contrast, variance in anxiety symptoms was not significantly influenced by common genetic variation in children or parents, despite a moderate pedigree heritability.

Conclusions: Genomic methods like RDR represent new opportunities for genetically sensitive family research in humans, which until now has been largely confined to adoption, twin and other pedigree designs. Our results are relevant to debates about the role of parents in the development of emotional problems in children, and possibly where to intervene to reduce problems.

Introduction

Associations between parent behaviours and child emotional problems (anxiety, depression and related symptoms and traits) are well documented (Yap, Pilkington, Ryan, & Jorm, 2014). Parental factors, such as control and hostility, that associate with child anxiety and depression symptoms (McLeod, Weisz, & Wood, 2007; McLeod, Wood, & Weisz, 2007), *may* represent useful modifiable causes. However, it cannot be assumed that such associations necessarily index environmental effects of parents' behaviour. A commonly overlooked complication is that genes shared by parents and children may lead to spurious or inflated associations between aspects of the rearing environment (e.g. parent characteristics) and heritable offspring traits. For instance, findings suggest that the association between parent divorce and offspring emotional problems is due to a common genetic component influencing parent and child traits, rather than there being any causal environmental effect of parent divorce (D'Onofrio et al. 2007).

Another challenge to the notion that parents environmentally influence child emotional development is the apparent lack of empirical support from behaviour genetics research (Plomin et al. 2016). Twin studies indicate negligible shared environmental influence (in which effects of parenting would likely be captured) on most child outcomes, including emotional development. Anxiety phenotypes are moderately genetically influenced, and the remaining variation is non-shared environmental in origin (i.e., non-genetic effects that make siblings more different from one another) (Waszczuk, Zavos, Gregory, & Eley, 2014). For depression, shared environment effects are more commonly detected, but decline across development, and have little influence on stability in emotional symptoms across ages (Hannigan, Walaker, Waszczuk, McAdams, & Eley, 2017).

The absence of shared environmental influence can be interpreted as an argument against the existence of main effects of the kinds of parental behaviours that make useful modifiable exposures – i.e. those with systematic effects causing familial aggregation. However, twin and pedigree designs alone have limited utility for investigating parent effects on child anxiety and depression. For instance, interactions between shared environmental and genetic influences will load onto estimates of genetic influence, thus potentially masking shared environmental effects (Rijsdijk & Sham 2002). Also, estimates of shared environmental variance include effects of other experiences that cause sibling resemblance (e.g. common friends and teachers), not only parents. Further, these designs assume that genes and environments are uncorrelated, which can lead to biased estimates of shared environmental variance (Coventry & Keller, 2005; Neiderhiser et al., 2004).

Adoption, children of twins (CoT), and sibling comparison designs provide stronger tests for environmental effects of parents on children than twin studies do, because they combine data on parents and children to correct for shared genes. Intergenerational genetically sensitive research on child anxiety is scarce, but two studies suggest that parent-child associations for anxiety remain after controlling for genetic relatedness (Ahmadzadeh et al., 2019; Eley et al., 2015). The evidence base is larger and more consistent for child depression. Numerous CoT and adoption studies support the idea of an environmental influence of parent depression after accounting for genetic confounding (Kendler,

Ohlsson, Sundquist, & Sundquist, 2018; Silberg, Maes, & Eaves, 2010; Singh et al., 2011; Natsuaki et al., 2014). Beyond parent depression, there are plausible environmental effects on child depression such as marital instability and harsh physical punishment, which survive correction for genetic relatedness (D’Onofrio et al., 2005; Lynch et al., 2006). Other putative effects on child depression, such as parent education (Torvik et al., 2020), may be better explained by shared genes.

While adoption and CoT designs are valuable, their utility is limited in three ways. First, these studies focus on how *specific* measured parent phenotypes relate to child outcomes. Such bivariate associations can be useful for identifying or disqualifying specific sources of potential causal influence. For example, it is useful to know that maternal depression during pregnancy does not predict offspring emotional problems in childhood after accounting for genetic relatedness (Hannigan et al., 2018), but that maternal postnatal depression does (Gjerde et al., 2019). However, it is not always possible to know *a priori* how to identify and effectively measure the parent traits that matter for children, or at what point in childhood to measure them. A second limitation of current genetically-informative parenting research is that the direction of effects is often not – or cannot be – assessed. We know that child-to-parent effects exist (Ahmadzadeh et al., 2019; Eley, Napolitano, Lau, & Gregory, 2010; McAdams et al., 2015; Narusyte et al., 2008), so where studies do not distinguish parent-to-child from child-to-parent effects, associations cannot be interpreted as environmental effects of parents. Third, reliance on highly specific samples (i.e. adoptees and twins) might preclude large-scale data collection and reduce the generalisability of findings.

Novel intergenerational genomic designs present opportunities to address these limitations and extend insights from adoption and CoT research. One method, Relatedness Disequilibrium Regression (RDR), was introduced as a technique for estimating genetic influence without bias from the family environment (Young et al., 2018). A less emphasised attribute of this method is that RDR can estimate environmental effects of parents. RDR measures the effect of parent behaviour indexed by their genome – termed ‘genetic nurture’ (Kong et al., 2018). In RDR, estimates of genetic nurture capture all behaviours of both parents that are influenced by their common genetic variation, and that affect the child’s trait. This is the first advantage of RDR: it removes the reliance on identifying and measuring specific parent traits. Unlike those estimated in adoption and CoT designs, the parent effect is global and atemporal. Genetic nurture indexes any lasting parent effect from any parent phenotype, up to the time-point at which the child’s phenotype was measured. Additional advantages of RDR are that it allows child- versus parent-driven effects to be disentangled from each other and uses scalable population-based family samples rather than adoptive or twin families. A previous study detected genetic nurture for child depression using a similar method (Jami et al. *in press*). However, the analysis was underpowered, maternal and paternal effects were not modelled simultaneously, and mediating parent phenotypes were not explored.

Here, we apply RDR to 11,598 parent offspring trios in the Norwegian Mother Father and Child Cohort Study (MoBa) to index environmental effects of parents on children’s anxiety and depression symptoms using the parental genome. Alongside parental genetic nurture, RDR estimates direct child genetic effects, and the covariance between direct and nurturing effects. We use available maternal

trait measures to test whether genetic nurture is mediated by self-reported maternal depression and anxiety. Additionally, we directly compare RDR results against traditional pedigree estimates, using 26,086 pairs of relatives (cousins, siblings, half-siblings and twins) from the child generation of MoBa.

Methods

Sample

The Norwegian Mother, Father and Child Cohort Study (MoBa; Magnus et al., 2016) is a prospective population-based pregnancy cohort study conducted by the Norwegian Institute of Public Health. Pregnant women were recruited from across Norway from 1999 to 2008. The women consented to initial participation in 41% of the pregnancies. The total cohort includes 114,500 children, 95,200 mothers and 75,200 fathers. To date, 98,110 individuals who are part of a trio (both parents and a child) from MoBa have been genotyped (Supplementary Table 1). As part of the Intergenerational Transmission of Risk (ITOR) subproject, MoBa has also been linked to Norwegian registry data containing pedigree and twin zygosity information. Version 11 of the quality-assured MoBa data files were used, released in 2018. Written informed consent was obtained from all participants upon recruitment. The establishment and data collection in MoBa was previously based on a license from the Norwegian Data Protection Agency and approval from The Regional Committee for Medical Research Ethics, and it is now based on regulations related to the Norwegian Health Registry Act. The current study was approved by The Regional Committee for Medical Research Ethics.

RDR and pedigree analyses required slightly different subsamples of MoBa families. For RDR, we used genotyped trios with available child anxiety and depression data. Siblings in the child generation were removed so that trios were independent families. For pedigree analyses, we retained all available cousins, siblings, half-siblings and twins in the child generation with anxiety and depression data, without requiring the presence of genotype data. Since only a subset of the families in MoBa were genotyped, all individuals used for genomic analyses were also within the sample used for pedigree analyses.

Measures

Our two outcome variables are well-validated and reliable quantitative measures of childhood anxiety and depression symptoms at age 8. Anxiety was measured using the 5-item version of the Screen for Child Anxiety Related Disorders (SCARED) (Birmaher et al., 1999). Depression was measured using the 13-item Short Moods and Feelings Questionnaire (SMFQ) (Angold, A., Costello, E. J., Messer, S. C., & Pickles, A., 1995). Both questionnaires were rated by mothers using three-point Likert response scales. Phenotypes were regressed on child sex prior to analyses.

To test whether measured parent emotional phenotypes explained any possible genetic nurture effects, we used the mother-rated Hopkins Symptoms Checklist (SCL-8) (Tambs & Røysamb, 2014). The SCL-8 assesses self-report anxiety and depression symptoms experienced during the last two weeks. To obtain a reliable measure of maternal emotions and behaviours that children are consistently exposed to, we capitalised on the longitudinal data in MoBa and constructed a common factor of SCL-8 scores at five time-points: 15 weeks of pregnancy, 30 weeks of pregnancy, child age 6 months, 18 months, and 8 years. As has been demonstrated with childhood data (Cheesman et al., 2018), a stable factor composed of measurements at multiple time points is more likely than individual scores at a single time point to capture a heritable core trait. Applying this approach to assess maternal

symptoms leads to a measure that captures stable exposure to maternal depression symptoms, rather than exposure associated with temporary fluctuations. Mediation of any genetic nurture effect by stable maternal emotional symptoms would be indicated by observed change in the genetic nurture point estimate when using the factor score as a covariate.

Genotype quality control

The current MoBa genomic dataset comprises imputed genetic data for 98,110 individuals (~32,000 parent offspring trios), derived from nine batches of participants, who make up four study cohorts. Within each batch, parent and offspring genetic data were quality controlled separately. Quality control exclusion criteria for individuals were: genotyping call rate <95%, or autosomal heterozygosity >4 standard deviations from the sample mean. Quality control exclusion criteria for SNPs were: ambiguous (A / T and C / G), genotyping call rate <98%, minor allele frequency <1%, or Hardy-Weinberg equilibrium P-value $<1 \times 10^{-6}$. Population stratification was assessed, using the HapMap phase 3 release 3 as a reference, by principal component analysis using EIGENSTRAT version 6.1.4. Visual inspection identified a homogenous population of European ethnicity and individuals of non-European ethnicity were removed. The parents and offspring datasets were then merged into one dataset per genotyping batch, keeping only the SNPs that passed quality control in both datasets. Phasing was conducted using Shapeit 2 release 837 and the duoHMM approach was used to account for the pedigree structure. Imputation was conducted using the Haplotype reference consortium (HRC) release 1-1 as the genetic reference panel. The Sanger Imputation Server was used to perform the imputation with the Positional Burrows-Wheeler Transform (PBWT). The phasing and imputation were conducted separately for each genotyping batch. Supplementary Table 1 contains details of the numbers of SNPs and individuals in each batch. More detailed information about the cohorts, quality control and imputation can be found at <https://github.com/folkehelseinstituttet/mobagen>.

We conducted post-imputation quality control, selecting SNPs meeting the follow criteria: high imputation confidence scores (INFO >0.8), minor allele frequency >0.05, Hardy-Weinberg equilibrium $p < 1 \times 10^{-6}$, non-multiallelic, non-duplicated.

Before calculating relatedness matrices for RDR, we removed individuals who were not part of a complete genotyped trio, and pruned down to 451,442 variants in approximate linkage equilibrium using an r^2 threshold of 0.5 (to reduce the size of the matrices and therefore the computational burden of the analyses).

After data management, a final sample of 25,828 genotyped parent child trios were identified. Almost 11,600 of these trios also had child emotional symptom data.

Statistical analyses

1. Genomic analysis of trios: Relatedness Disequilibrium Regression (RDR)

The RDR method (Young et al., 2018) allows the estimation of parent and child genetic effects on traits. This is achieved by extending a commonly used genomic method for estimating heritability – single-component GREML (Genomic-Relatedness based restricted Maximum-Likelihood) (Yang, Lee,

Goddard, & Visscher, 2011) – to include individuals' parents. The standard GREML method estimates the variance explained by common single nucleotide polymorphisms (SNPs) by comparing a matrix of pairwise genomic similarity for unrelated individuals across genotyped SNPs to a matrix of their pairwise phenotypic similarity, using a random-effects mixed linear model. Instead of using the random variation in genetic similarity among unrelated individuals, RDR capitalises on the random variation in genetic similarity between parents and offspring, which arises due to random segregation, and is independent of environmental factors.

There are two versions of RDR: one uses identity-by-descent (IBD) relatedness, which distinguishes parts of the genome that are inherited from common ancestors, and the other uses common SNP-based relatedness. We used the SNP version since it has similar properties to the IBD version but is more comparable to the standard single-component GREML approach and has greater statistical power. Rather than estimating a single genetic variance component, RDR estimates three. The first estimates the *direct* effect of children's own DNA on their trait. This variance component is independent of the effect of being reared by biological parents. Importantly, a direct genetic effect is only direct in the sense that it does not stem from another individual's genotype. It is also not necessarily 'purely' genetic, since genotypes are behaviourally mediated and expressed in the context of environments. Notably, mechanisms by which individuals evoke and select environments based on their genotype, are essential in how genes lead to phenotypes (Plomin, 2014), and these are included in estimates of direct genetic influence.

The second variance component estimates the effect of parent genetics on the child trait, controlling for child genetic effects: 'genetic nurture'. Any parent genetic effect over and above child-driven direct effects must be an *indirect* genetic effect, where parents' genes affect child traits by influencing parent behaviours and the rearing environment they provide.

The third component captures variance in the offspring phenotype attributable to covariance between the direct and nurturing genetic effects, known as 'passive gene-environment correlation' (see Supplementary Figure 1 for detail on this concept). This somewhat abstract variance component is easier to understand when considering the conditions for the estimate to be zero. Specifically, the direct-nurturing genetic covariance would not explain any phenotypic variation if only one generation contributes genetic effects to the child trait, or if different SNPs contribute to child and parent genetic influences, such that loci have only either direct or indirect effects.

Finally, the residual component captures environmental effects on the trait of interest that are not correlated with measured parent genetic variation, the effects of variants not tagged by genotyped SNPs (e.g. rarer), and measurement error.

In practice, the variance components are estimated by regressing phenotypic similarity/dissimilarity on three genomic relatedness matrices simultaneously. The first is similar to the matrix used in standard GREML: the genome-wide genetic relatedness of the children. The second and third represent the genetic relatedness of the parents and the genetic covariance between children and parents.

Notably, the genotypes of mothers and fathers are combined to allow estimation of the average effect of *both* parents, rather than only mothers as in a similar model, M-GCTA (Eaves, Pourcain, Smith, York, & Evans, 2014), which estimates the effects of the mother and child genomes and of their covariance. Therefore, before constructing the parent genomic relatedness matrix, maternal and paternal genotypes are summed and standardised to have mean zero and variance two (Young et al., 2018).

All RDR analyses included 10 ancestry principal components and genotyping batch, both derived from the child generation, as covariates. Analyses were performed in the GCTA software using the `--reml-no-constrain` flag to allow variance components to take positive and negative values (Yang et al. 2011).

To test whether any genetic nurture effect was partially explained by parent emotional symptoms, we re-ran the RDR models adding a measure of stable maternal emotional symptoms as a covariate. We also tested the individual time-specific maternal measures as covariates.

To test whether any of the variance components, particularly the effect of covariance between direct and nurturing effects, were biased because some children were genotyped in a different batch to their parents, we re-ran the RDR models using only individuals who were genotyped as a complete trio (90% of the analysis sample).

II. Classical pedigree modelling

To compare RDR results to a traditional quantitative genetic (non-genomic) design, we implemented a univariate pedigree model (Neale & Cardon, 1992). As in the classic twin design, this model allows estimation of genetic, shared environmental, and non-shared environmental (residual) influences on anxiety and depression symptoms at age 8. The model used phenotypic correlations among cousins, siblings, half-siblings and twins in the child generation, to derive estimates based on the following specifications: genetic correlations are 1.00, 0.50, 0.25, and 0.125 for identical twins, non-identical twins/siblings, half-siblings and cousins respectively, and shared environmental correlations are 1.00 for all siblings and 0.00 for cousins.

Comparison of variance components from RDR and pedigree models

Before comparing RDR and pedigree results, it is essential to note the differences between variance components derived from the models. First, with respect to child genetic effects, the pedigree genetic component denotes a broader effect of any DNA variation, not only common additive effects tagged by measured variants, as in RDR applied to common SNPs. Second, regarding family environmental effects, the shared environment component in the pedigree model is broader than the genetic nurture estimate using RDR. The shared environment captures at least three sources of variance that genetic nurture does not: parent behaviours not tagged by common SNPs, shared environments *beyond* parent behaviours, and effects of any correlation between the shared environment and genetic components (passive gene-environment correlation). Third, whereas pedigree and twin models assume passive gene-environment correlation is absent (i.e. that genetic and environmental effects are

uncorrelated), RDR can directly estimate it as the covariance between genetic and environmental (genetic nurturing) effects.

Supplementary Figure 1 and its accompanying text explore overlaps and distinctions between the concepts of genetic nurture, passive gene-environment correlation, and shared environment.

Software

Genome-wide relatedness matrices were constructed using python, and RDR models were run using GCTA (Yang et al. 2011). Pedigree analyses were conducted in R using the structural equation modelling package OpenMx (Neale et al. 2016).

Results

RDR model

Results from our RDR analyses using genomic data to disentangle child and parent genetic effects on anxiety and depression symptoms at age 8 are shown in Figure 2. Direct effects of children's own common genetic variation (yellow bar) explained 5% of the variance in anxiety symptoms ($se=0.07$), and 19% ($se=0.07$) of the variance in depressive symptoms. Genetic nurture (dark green bars) had a negligible influence on variation in child anxiety, but explained 14% ($se=0.07$) of the variance in child depression. For depression, the estimate of the phenotypic variance explained by covariance between direct and indirect genetic effects (i.e. passive gene-environment correlation) was negative (-16%; $se=0.07$; pale green bars).

See Supplementary Table 2 for full RDR results including sensitivity analyses exploring mediators, batch effects, and the effect of constraining components to take positive values. In the sensitivity analysis exploring mediation, maternal emotional symptoms partially explained the genetic nurture effect on childhood depression. This manifested as the genetic nurture effect, but not the other variance components, being substantially attenuated when including a covariate capturing stable maternal anxiety and depression symptoms. The genetic nurture effect on child depression dropped from 14% to 5%. Most individual time-specific maternal measures also led to attenuation of the genetic nurture effect when included as covariates, even some that were measured during pregnancy.

Our results were robust to any effects of children being genotyped in a different batch to their parents. Variance components were unchanged after restricting the sample to trios who were genotyped together.

Supplementary Figure 2 visualises the RDR results in a path diagram, with estimates of paths for direct and nurturing effects, and of the correlation between them.

Pedigree model

Our key finding of genetic nurture for child depression but not anxiety was supported by our pedigree analysis (Figure 3). Shared environmental influences contributed 13% ($se = 0.05$) of the variation in depression, but we observed no robust evidence of shared environmental effects on anxiety. The pedigree heritability estimates were similar for anxiety and depression (~37%). This contrasts to the pattern of heritability estimates from RDR, in which child genetic effects explained more variance in depression than anxiety (19% versus 5%, respectively). See Supplementary Figure 3 for a path diagram version of the pedigree results.

Figure 2. Variance components from the RDR model. Labels= variance explained by direct child genetic effects (yellow), parent genetic effects (i.e. genetic nurture; dark green), and covariance between the direct and nurturing genetic effects ('Both'; pale green). Bars=standard errors. Not shown: residual variance affecting the phenotypes, including parent genetic effects not tagged by common SNPs, individual-specific environmental effects, chance factors and error.

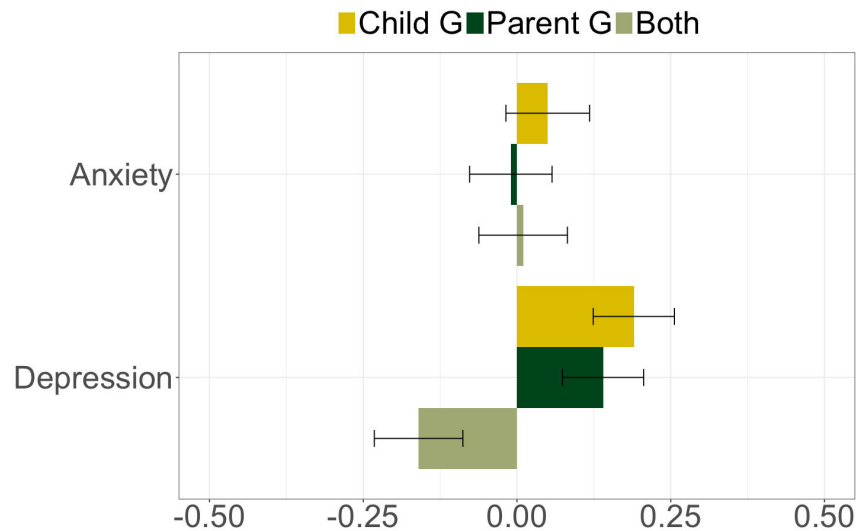
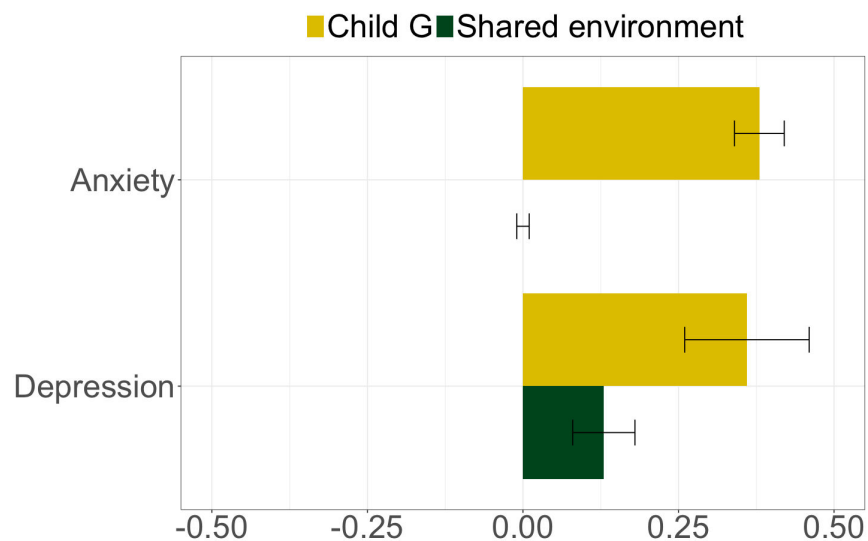


Figure 3. Pedigree variance components. Labels= variance explained by genetic effects (i.e pedigree heritability; yellow), variance explained by environmental influences that increase resemblance among siblings in the same family (dark green). Bars = standard errors. Not shown: Residual variance, which includes non-shared environmental effects and error. The effect of any covariance between genetic and shared environmental influences is not estimated in this model.



Discussion

We used genomic data from parent-offspring trios to estimate indirect parent genetic effects (genetic nurture) separately from direct child genetic effects on anxiety and depression symptoms at age 8. Results from the RDR model suggested that depressive symptoms were influenced by both genetic nurture and direct genetic effects. Genetic nurture represents an environmentally mediated effect of genetic origin, indexing the behaviours of both parents that are influenced by their common DNA variation and that influence child depression, over and above direct child genetic effects. We found that the genetic nurture effect was partially mediated via an observed measure of maternal anxiety and depression, which explained 64% of the original genetic nurture effect (14%). In contrast, individual differences in anxiety symptoms were not significantly influenced by parental genetic nurture, and direct effects of common genetic variants explained just 5% of the phenotypic variance.

Genetic nurture for child depression symptoms. Our genomic evidence of an environmental effect of parents on child depression, but not anxiety, was supported by pedigree-based evidence for shared environmental influence on depression, but not anxiety. This aligns well with findings from epidemiological and twin studies. Unlike anxiety, depressive symptoms are uncommon in pre-pubertal children, but when they do occur, are typically a response to psychosocial risks, such as perinatal issues, abuse, and bullying (Jaffee et al., 2002; Maughan, Collishaw, & Stringaris, 2013). Twin data also suggest a greater role for family environmental factors in child depression than anxiety. In one twin study, shared environmental effects contributed 18% of the variance in depressive symptoms in middle childhood, but none for anxiety (Cheesman et al., 2018). The use of RDR adds to our knowledge of parent effects from pedigree methods, because genetic nurture is simultaneously specific to parents and global. In contrast, estimates of parental influence from twin studies (i.e. the shared environment) are not specific to parents, and those from the children of twins (CoT) method rely on phenotypic measurement of specific parent traits.

Mediation by maternal anxiety and depression. Our analysis adjusting for parent emotional symptoms suggested that maternal anxiety and depression mediate the genetic nurture effect on childhood depression. This is compatible with adoption- and CoT-based evidence for phenotypic effects of maternal depression (Natsuaki et al., 2014). This environmental influence could reflect a learning mechanism directly related to maternal emotional symptoms, or could be mediated through secondary parenting-related behaviours such as disorganized home environments (Rapee, Schniering, & Hudson, 2009; Natsuaki et al., 2014).

Negative covariance between direct and nurturing genetic effects on child depression. This negative component suggests that genetic variation that is associated with higher depression risk when present in children, has an opposing effect (*reducing* child depression risk) when present in the parents. On the face of it, this result is difficult to explain and should be interpreted with caution. Notably, covariance components were also negative in RDR analyses of multiple phenotypes including creatinine levels and age of first child in men (Young et al. 2018), and in a Maternal-GCTA analysis of gestational weight gain (Warrington et al. 2018). Our finding could be explained by emotional sensitivity. In other

words, genes that lead mothers to identify depressive symptoms in their children could be associated with parent behaviours that reduce those same symptoms.

Pedigree-genomic heritability gap. For child anxiety, the pedigree heritability estimate was moderate, despite the small genomic signal from RDR. This is compatible with findings from other samples (Cheesman et al., 2017). However, it is surprising that the SNP-based effects are lower for anxiety than depression, given that they are similar traits, with similar pedigree heritability estimates. One possible explanation is that anxiety, more than depression, is influenced by gene-by-shared environment interactions. These interactions may be captured in pedigree heritability (genes make identical twins more similar than non-identical twins; shared environmental factors make both sets of twins more similar, so the interaction may further increase identical twin similarity) but not RDR heritability estimates, since the latter does not compare children who share a family environment.

Limitations. We acknowledge potential sources of bias affecting the RDR results. First, although supportive of the notion of parental influence on child depressive symptoms, the genetic nurture estimate may also capture residual population stratification, assortative mating, and indirect genetic effects from siblings.

Second, the genetic nurture effect, and the mediation effect, could be partly generated by the use of maternal ratings of child outcomes. Maternal reports of child depression symptoms may reflect their own emotional symptoms. As a result, an apparent parental genetic nurture effect may in fact be a direct genetic effect on parents' own traits. This phenomenon has been evidenced in developmental twin studies, in which rater bias may cause shared environmental variance to be overestimated (Bartels et al., 2004). However, if rater bias explained the environmental effect for depression, then we might expect to see some evidence for this with anxiety too. The lack of a genetic nurture effect for anxiety in our RDR and pedigree models suggests that maternal report is still a measure of child and not entirely maternal phenotype. Measurement issues will be clarified with the upcoming availability of child, teacher, and clinician reports in MoBa.

Third, selective participation and attrition might have reduced coverage of families experiencing more severe emotional symptoms. Depression-linked attrition has been demonstrated in the Norwegian Twin Panel (Tambs et al. 2009). Stronger parent effects might occur at the tails of the distribution of child emotional problems. Future linkage of MoBa with Norwegian registry data will help to investigate and control for participation biases affecting the genotyped trios.

Future directions. Future studies could seek to identify behaviours other than maternal emotional symptoms that explain why the parent genome independently influences child depression. Since common genetic variants only capture a subset of the genetic component of complex traits involved in parenting, our observed genetic nurture effect on childhood depression is likely to be larger than we estimate here. This implies substantial scope for identifying heritable parent phenotypes with effects on child depression symptoms. However, individual parenting phenotypes may have small effect sizes, and are usually only modestly heritable (~20%)(Klahr & Burt 2014). Researchers should therefore continue to explore wider social factors that might affect parenting, and in turn, childhood anxiety and

depression. Mediators of genetic nurture might depend on socioeconomic context. For example, poverty increases the magnitude of the effect of maternal depression on diverse child outcomes (Goodman et al., 2011). In the future, RDR could be adapted to allow moderation effects to be tested.

The RDR method could be broadened to address multivariate and longitudinal questions, as has been done with twin designs (Boomsma, Busjahn, & Peltonen, 2002). For example, does genetic nurture persistently influence depressive symptoms across development? Twin research suggests that the effects of parents (and passive gene-environment correlation) will be stronger earlier in life (Hannigan, McAdams, & Eley, 2017; Rice, 2010; Thapar & McGuffin, 1994). This is partly because with increasing age less time is spent in the family environment, and individuals have greater capacity to choose their experiences and express their genetic predispositions. Genomic tools could be used to strengthen and develop findings from traditional designs.

Summary. Our genomic approach strengthens the evidence base regarding parent effects on child emotional development, by directly accounting for shared DNA, and by indexing the child's environment using the parental genome. Quantifying the contribution of parents is crucial for the understanding of how anxiety and depression develop in childhood, and potentially how to target modifiable causes. Our results are supportive of parent effects on depressive symptoms at age 8. The genetic nurture effect on depressive symptoms was partially mediated by maternal anxiety and depression. Though the effect size is modest, this suggests that efforts to alleviate maternal emotional problems could help to prevent child depression.

Acknowledgements. The work was supported by the Norwegian Research Council (Grant Number 262177). TE is part-funded by a program grant from the UK Medical Research Council (MR/M021475/1), and by the National Institute for Health Research (NIHR), the Biomedical Research Centre at South London, Maudsley NHS Foundation Trust and King's College London. EY, and TAM are supported by the Norwegian Research Council (grant number 288083). AH and LJH are supported by the South-Eastern Norway Regional Health Authority (grant numbers 2018058 and 2018059). The views expressed are those of the author(s) and not necessarily those of the NHS, the NIHR or the Department of Health. TAM and YIA are supported by a Sir Henry Dale Fellowship jointly funded by the Wellcome Trust and the Royal Society to TAM (Grant Number 107706/Z/15/Z). RC is supported by an ESRC studentship. The Norwegian Mother, Father and Child Cohort Study is supported by the Norwegian Ministry of Health and Care Services and the Ministry of Education and Research. We are grateful to all the participating families in Norway who take part in this on-going cohort study.

References

- Ahmadzadeh, Y. I., Eley, T. C., Leve, L. D., Shaw, D. S., Natsuaki, M. N., Reiss, D., ... McAdams, T. A. (2019). Anxiety in the family: a genetically informed analysis of transactional associations between mother, father and child anxiety symptoms. *Journal of Child Psychology and Psychiatry, and Allied Disciplines*. doi:10.1111/jcpp.13068
- Angold, A., Costello, E. J., Messer, S. C., & Pickles, A. (1995). The Development of a Questionnaire for Use in Epidemiological Studies of Depression in Children and Adolescents. *International Journal of Methods in Psychiatric Research*, 5(4), 237-249.
- Bartels, M., Boomsma, D. I., Hudziak, J. J., Rietveld, M. J. H., van Beijsterveldt, T. C. E. M., & van den Oord, E. J. C. G. (2004). Disentangling genetic, environmental, and rater effects on internalizing and externalizing problem behavior in 10-year-old twins. *Twin Research: The Official Journal of the International Society for Twin Studies*, 7(2), 162–175. doi:10.1375/136905204323016140
- Birmaher, B., Brent, D. A., Chiappetta, L., Bridge, J., Monga, S., & Baugher, M. (1999). Psychometric properties of the Screen for Child Anxiety Related Emotional Disorders (SCARED): a replication study. *Journal of the American Academy of Child and Adolescent Psychiatry*, 38(10), 1230–1236. doi:10.1097/00004583-199910000-00011
- Boomsma, D., Busjahn, A., & Peltonen, L. (2002). Classical twin studies and beyond. *Nature Reviews. Genetics*, 3(11), 872–882. doi:10.1038/nrg932
- Cheesman, R., Major Depressive Disorder Working Group of the Psychiatric Genomics Consortium, Purves, K. L., Pingault, J.-B., Breen, G., Rijdsdijk, F., ... Eley, T. C. (2018). Extracting stability increases the SNP heritability of emotional problems in young people. *Translational Psychiatry*, 8(1), 223. doi:10.1038/s41398-018-0269-5
- Cheesman, R., Selzam, S., Ronald, A., Dale, P. S., McAdams, T. A., Eley, T. C., & Plomin, R. (2017). Childhood behaviour problems show the greatest gap between DNA-based and twin heritability. *Translational Psychiatry*.
- Coventry, W. L., & Keller, M. C. (2005). Estimating the extent of parameter bias in the classical twin design: a comparison of parameter estimates from extended twin-family and classical twin designs. *Twin Research and Human Genetics*, 8(3), 214–223. doi:10.1375/1832427054253121
- D’Onofrio, B. M., Turkheimer, E., Emery, R. E., Slutske, W. S., Heath, A. C., Madden, P. A., & Martin, N. G. (2005). A genetically informed study of marital instability and its association with offspring psychopathology. *Journal of Abnormal Psychology*, 114(4), 570–586. doi:10.1037/0021-843X.114.4.570
- D’Onofrio, B.M., Turkheimer, E., Emery, R.E., Maes, H.H., Silberg, J. and Eaves, L.J. 2007. A Children of Twins Study of parental divorce and offspring psychopathology. *Journal of Child Psychology and Psychiatry, and Allied Disciplines* 48(7), pp. 667–675.
- Eaves, L. J., Pourcain, B. S., Smith, G. D., York, T. P., & Evans, D. M. (2014). Resolving the effects of maternal and offspring genotype on dyadic outcomes in genome wide complex trait analysis (“M-GCTA”). *Behavior Genetics*, 44(5), 445–455. doi:10.1007/s10519-014-9666-6
- Eley, T. C., McAdams, T. A., Rijdsdijk, F. V., Lichtenstein, P., Narusyte, J., Reiss, D., ... Neiderhiser, J. M. (2015). The Intergenerational Transmission of Anxiety: A Children-of-Twins Study. *The American Journal of Psychiatry*, 172(7), 630–637. doi:10.1176/appi.ajp.2015.14070818
- Eley, T. C., Napolitano, M., Lau, J. Y. F., & Gregory, A. M. (2010). Does childhood anxiety evoke maternal control? A genetically informed study. *Journal of Child Psychology and Psychiatry, and Allied Disciplines*, 51(7), 772–779. doi:10.1111/j.1469-7610.2010.02227.x

- Gjerde, L. C., Eilertsen, E. M., Hannigan, L. J., Eley, T., Røysamb, E., Reichborn-Kjennerud, T., ... Ystrom, E. (2019). Associations between maternal depressive symptoms and risk for offspring early-life psychopathology: the role of genetic and non-genetic mechanisms. *Psychological Medicine*, 1–9. doi:10.1017/S0033291719003301
- Goodman, S. H., Rouse, M. H., Connell, A. M., Broth, M. R., Hall, C. M., & Heyward, D. (2011). Maternal depression and child psychopathology: a meta-analytic review. *Clinical Child and Family Psychology Review*, 14(1), 1–27. doi:10.1007/s10567-010-0080-1
- Hannigan, L. J., Walaker, N., Waszczuk, M. A., McAdams, T. A., & Eley, T. C. (2017). Aetiological influences on stability and change in emotional and behavioural problems across development: a systematic review. *Psychopathology Review*, 4(1), 52–108. doi:10.5127/pr.038315
- Hannigan, Laurie J, McAdams, T. A., & Eley, T. C. (2017). Developmental change in the association between adolescent depressive symptoms and the home environment: results from a longitudinal, genetically informative investigation. *Journal of Child Psychology and Psychiatry, and Allied Disciplines*, 58(7), 787–797. doi:10.1111/jcpp.12689
- Jaffee, S. R., Moffitt, T. E., Caspi, A., Fombonne, E., Poulton, R., & Martin, J. (2002). Differences in early childhood risk factors for juvenile-onset and adult-onset depression. *Archives of General Psychiatry*, 59(3), 215–222. doi:10.1001/archpsyc.59.3.215
- Jami, E. S., Eilertsen, E. M., Hammerschlag, A. R., Qiao, Z., Evans, D. M., Ystrom, E., Bartels, M., & Middeldorp, C. (*In press*). Maternal and paternal effects on offspring internalising problems: results from genetic and family-based analyses. *American Journal of Medical Genetics Part B: Neuropsychiatric Genetics*.
- Kendler, K. S., Ohlsson, H., Sundquist, K., & Sundquist, J. (2018). Sources of Parent-Offspring Resemblance for Major Depression in a National Swedish Extended Adoption Study. *JAMA Psychiatry*, 75(2), 194–200. doi:10.1001/jamapsychiatry.2017.3828
- Kong, A., Thorleifsson, G., Frigge, M. L., Vilhjalmsen, B. J., Young, A. I., Thorgeirsson, T. E., ... Stefansson, K. (2018). The nature of nurture: Effects of parental genotypes. *Science*, 359(6374), 424–428. doi:10.1126/science.aan6877
- Lynch, S. K., Turkheimer, E., D'Onofrio, B. M., Mendle, J., Emery, R. E., Slutske, W. S., & Martin, N. G. (2006). A genetically informed study of the association between harsh punishment and offspring behavioral problems. *Journal of Family Psychology*, 20(2), 190–198. doi:10.1037/0893-3200.20.2.190
- Magnus, P., Birke, C., Vejrup, K., Haugan, A., Alsaker, E., Daltveit, A. K., ... Stoltenberg, C. (2016). Cohort profile update: the norwegian mother and child cohort study (moba). *International Journal of Epidemiology*, 45(2), 382–388. doi:10.1093/ije/dyw029
- Maughan, B., Collishaw, S., & Stringaris, A. (2013). Depression in childhood and adolescence. *Journal of the Canadian Academy of Child and Adolescent Psychiatry*, 22(1), 35–40.
- McAdams, T. A., Rijdsdijk, F. V., Neiderhiser, J. M., Narusyte, J., Shaw, D. S., Natsuaki, M. N., ... Eley, T. C. (2015). The relationship between parental depressive symptoms and offspring psychopathology: evidence from a children-of-twins study and an adoption study. *Psychological Medicine*, 45(12), 2583–2594. doi:10.1017/S0033291715000501
- McAdams, Tom A, Hannigan, L. J., Eilertsen, E. M., Gjerde, L. C., Ystrom, E., & Rijdsdijk, F. V. (2018). Revisiting the Children-of-Twins Design: Improving Existing Models for the Exploration of Intergenerational Associations. *Behavior Genetics*, 48(5), 397–412. doi:10.1007/s10519-018-9912-4
- McLeod, B. D., Weisz, J. R., & Wood, J. J. (2007). Examining the association between parenting and childhood depression: a meta-analysis. *Clinical Psychology Review*, 27(8), 986–1003. doi:10.1016/j.cpr.2007.03.001

- McLeod, B. D., Wood, J. J., & Weisz, J. R. (2007). Examining the association between parenting and childhood anxiety: a meta-analysis. *Clinical Psychology Review*, 27(2), 155–172. doi:10.1016/j.cpr.2006.09.002
- Narusyte, J., Neiderhiser, J. M., D'Onofrio, B. M., Reiss, D., Spotts, E. L., Ganiban, J., & Lichtenstein, P. (2008). Testing different types of genotype-environment correlation: an extended children-of-twins model. *Developmental Psychology*, 44(6), 1591–1603. doi:10.1037/a0013911
- Natsuaki, M. N., Shaw, D. S., Neiderhiser, J. M., Ganiban, J. M., Harold, G. T., Reiss, D., & Leve, L. D. (2014). Raised by depressed parents: is it an environmental risk? *Clinical Child and Family Psychology Review*, 17(4), 357–367. doi:10.1007/s10567-014-0169-z
- Neale, M.C. and Cardon, L.R. 1992. Methodology for genetic studies of twins and families. Dordrecht: Springer Netherlands.
- Neale, M.C., Hunter, M.D., Pritikin, J.N., Zahery, M., Brick, T.R., Kirkpatrick, R.M., Estabrook, R., Bates, T.C., Maes, H.H. and Boker, S.M. 2016. Openmx 2.0: extended structural equation and statistical modeling. *Psychometrika* 81(2), pp. 535–549.
- Neiderhiser, J. M., Reiss, D., Pedersen, N. L., Lichtenstein, P., Spotts, E. L., Hansson, K., ... Ellhammer, O. (2004). Genetic and environmental influences on mothering of adolescents: a comparison of two samples. *Developmental Psychology*, 40(3), 335–351. doi:10.1037/0012-1649.40.3.335
- Plomin, R. (2014). Genotype-environment correlation in the era of DNA. *Behavior Genetics*, 44, 629–638.
- Plomin, R., & Daniels, D. (1987). Why are children in the same family so different from one another? *Behavioral and Brain Sciences*, 10(01), 1. doi:10.1017/S0140525X00055941
- Plomin, R., DeFries, J.C., Knopik, V.S. and Neiderhiser, J.M. 2016. Top 10 replicated findings from behavioral genetics. *Perspectives on psychological science : a journal of the Association for Psychological Science* 11(1), pp. 3–23.
- Rapee, R. M., Schniering, C. A., & Hudson, J. L. (2009). Anxiety disorders during childhood and adolescence: origins and treatment. *Annual Review of Clinical Psychology*, 5, 311–341. doi:10.1146/annurev.clinpsy.032408.153628
- Rice, F. (2010). Genetics of childhood and adolescent depression: insights into etiological heterogeneity and challenges for future genomic research. *Genome Medicine*, 2(9), 68. doi:10.1186/gm189
- Rijsdijk, F.V. and Sham, P.C. 2002. Analytic approaches to twin data using structural equation models. *Briefings in Bioinformatics* 3(2), pp. 119–133.
- Silberg, J. L., Maes, H., & Eaves, L. J. (2010). Genetic and environmental influences on the transmission of parental depression to children's depression and conduct disturbance: an extended Children of Twins study. *Journal of Child Psychology and Psychiatry, and Allied Disciplines*, 51(6), 734–744. doi:10.1111/j.1469-7610.2010.02205.x
- Singh, A. L., D'Onofrio, B. M., Slutske, W. S., Turkheimer, E., Emery, R. E., Harden, K. P., ... Martin, N. G. (2011). Parental depression and offspring psychopathology: a children of twins study. *Psychological Medicine*, 41(7), 1385–1395. doi:10.1017/S0033291710002059
- Tambs, K., Rønning, T., Prescott, C.A., Kendler, K.S., Reichborn-Kjennerud, T., Torgersen, S. and Harris, J.R. 2009. The Norwegian Institute of Public Health twin study of mental health: examining recruitment and attrition bias. *Twin Research and Human Genetics* 12(2), pp. 158–168.
- Tambs, K., & Røysamb, E. (2014). Selection of questions to short-form versions of original psychometric instruments in MoBa. *Norsk Epidemiologi*, 24(1-2). doi:10.5324/nje.v24i1-2.1822

- Thapar, A., & McGuffin, P. (1994). A twin study of depressive symptoms in childhood. *The British Journal of Psychiatry*, 165(2), 259–265. doi:10.1192/bjp.165.2.259
- Torvik, F. A., Eilertsen, E. M., McAdams, T. A., Gustavson, K., Zachrisson, H. D., Brandlistuen, R., ... Ystrom, E. (2020). Mechanisms linking parental educational attainment with child ADHD, depression, and academic problems: a study of extended families in The Norwegian Mother, Father and Child Cohort Study. *Journal of Child Psychology and Psychiatry*. doi:10.1111/jcpp.13197
- Warrington, N.M., Richmond, R., Fenstra, B., Myhre, R., Lawlor, D.A., et al. 2018. Maternal and fetal genetic contribution to gestational weight gain. *International Journal of Obesity* 42(4), pp. 775–784.
- Waszczuk, M. A., Zavos, H. M. S., Gregory, A. M., & Eley, T. C. (2014). The phenotypic and genetic structure of depression and anxiety disorder symptoms in childhood, adolescence, and young adulthood. *JAMA Psychiatry*, 71(8), 905–916. doi:10.1001/jamapsychiatry.2014.655
- Wilkinson, P. O., Trzaskowski, M., Haworth, C. M. A., & Eley, T. C. (2013). The role of gene-environment correlations and interactions in middle childhood depressive symptoms. *Development and Psychopathology*, 25(1), 93–104. doi:10.1017/S0954579412000922
- Yang, J., 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. doi:10.1016/j.ajhg.2010.11.011
- Yap, M. B. H., Pilkington, P. D., Ryan, S. M., & Jorm, A. F. (2014). Parental factors associated with depression and anxiety in young people: a systematic review and meta-analysis. *Journal of Affective Disorders*, 156, 8–23. doi:10.1016/j.jad.2013.11.007
- Young, A. I., Frigge, M. L., Gudbjartsson, D. F., Thorleifsson, G., Bjornsdottir, G., Sulem, P., ... Kong, A. (2018). Relatedness disequilibrium regression estimates heritability without environmental bias. *Nature Genetics*, 50(9), 1304–1310. doi:10.1038/s41588-018-0178-9

Chapter 6: Discussion

The over-arching aim of this thesis was to create a productive synthesis of quantitative genetics and genomics. This final chapter discusses the empirical findings of the thesis in relation to this aim. I will first present an overview of the main results, and then summarise each chapter. Next, I will discuss two key cross-cutting themes: missing heritability, and gene-environment interplay. I will then outline some general limitations common to the approaches taken across the four studies. Finally, I will draw out directions for future research, before summarising the conclusions.

Overview of findings

Genomic data is being used to answer old and new questions about how individual differences in complex traits arise. Traditionally, expected values of genetic similarity among relatives were used to estimate heritability. Now, it is possible to estimate realised genetic similarity between unrelated individuals and to estimate heritability by exploiting that information. As advances in genomics and statistical analysis have facilitated large-scale projects aimed at identifying specific genetic variants involved in complex traits, the continuing utility of family-based quantitative designs has been questioned. However, population genomic research faces problems that family-based approaches rooted in quantitative genetics can help to address. First, low SNP heritability estimates suggest that genome-wide association studies will struggle to give insights on the genetic architecture of complex traits, because both SNP heritability and genome-wide association studies are limited to the additive effects of measured common variants. Results from twin studies can be used to guide genomic studies towards more heritable phenotypic assessments and incorporating family data may help to tag additional effects of rarer genetic variants. Second, SNP heritability estimates and polygenic score associations are sometimes entangled with effects of the environment provided by relatives. Environmentally and genetically informative datasets such as those including adoptees and parent-offspring trios can be used to isolate direct genetic effects on complex traits. Research on genetic nurture is a good example of how integrating genomics with quantitative genetic knowledge can yield novel insights relevant to social epidemiology, about plausible causal effects of modifiable family environmental factors. Overall, the work presented in this thesis demonstrates that family-based genetic research is not only the foundation and justification for complex trait genomics, but is of enduring importance, especially when integrated with new genomic methods.

Chapter-specific findings

Each chapter in this thesis built on the preceding one, iteratively using new family-based approaches to overcome challenges in complex trait genomics.

Study 1: Extracting stability increases the SNP heritability of emotional problems in young people

This thesis began with a pressing issue in genomic studies of childhood psychopathology – negligible SNP heritability. Chapter 2 showed that genomic signal can be detected for childhood emotional problems when longitudinal twin data from TEDS are leveraged to extract a temporally stable and contextually pervasive core emotional trait. A pertinent analogy for this would be that a ‘fly on the wall’ observing a child’s development, at home and school, gathers more accurate and reliable information compared to any one-off measure. Informed by evidence from decades of twin research, this assessment is more heritable and useful for genomic prediction studies than individual measures.

Study 2: Familial influences on neuroticism and education in the UK Biobank

Although Chapter 2 addressed the problem of low SNP heritability for emotional problems, phenotypic measurement-based approaches cannot tackle the SNP-twin heritability gap per se, since twin heritability also scales up with increased accuracy of measurement. Chapter 3 directly investigated the extent that rarer genetic variation can account for this missing heritability, by applying a little-used family-based genomic heritability method (GREML-KIN) in the UK Biobank to decompose sources of variance in neuroticism and educational attainment. This method suggested that additional family-based genetic effects can recover pedigree-level heritability. However, results from a parallel method using unrelated individuals and imputed variants (GREML-LDMS-I) did not fully support the rare variant explanation for the heritability gap. A further limitation is that GREML-KIN-based heritability may be inflated due to passive gene-environment correlation. This occurs when direct genetic effects are correlated with environmentally mediated indirect genetic effects from participants’ biological relatives.

Study 3: Comparison of adopted and non-adopted individuals reveals gene-environment interplay for education in the UK Biobank

Chapter 4 presented a means of establishing the extent that heritability estimates based on UK Biobank data (as in Chapter 3) are inflated by passive gene-environment correlation. Adoptive placement creates an ideal context to assess how much individuals’ own inherited DNA explains variation in their phenotype. Since adoptees are unrelated to their parents, strong correlations between offspring genotype and rearing environment are ruled out. Polygenic scores explained only half the variance in educational attainment among adoptees as in non-adopted individuals, suggesting that half of their effect stems from environmentally mediated effects of participants’ genetic relatives. Clearly, making connections between new genomic tools and classic ‘natural experiments’ from quantitative genetics allows new insights about modes of action of genes and environments, and also about the nature of the new genomic tools.

Study 4: How important are parents in the development of child anxiety and depression? A genomic analysis of parent-offspring trios in the Norwegian Mother Father and Child Cohort Study

Chapter 5, like Chapter 4, focussed on distinguishing direct from indirect genetic effects on complex traits. Instead of using adoptees to eliminate environmental effects of genetic relatives, Chapter 5 used genotyped parent-offspring trios to isolate environmentally mediated genetic effects of parents (genetic nurture) from direct genetic effects. Having genotype data for complete trios does not only mean that genetics researchers can obtain estimates of direct genetic influence without environmental bias. It also means that ‘environmentalists’ can test for potential causal effects of heritable parent behaviours on child outcomes, controlling for child genetic effects. Relatedness Disequilibrium Regression results revealed a parental genetic nurture effect on depression symptoms in childhood. This effect appeared to be partially explained by maternal anxiety and depression symptoms, a potentially useful modifiable cause. This study used family genomic data to confirm and extend the insights from quantitative genetics that the environment is genetically influenced, and that the strongest evidence for environmental effects comes from genetically informative designs.

Conclusions about cross-cutting themes

The beginning of this chapter talked about the big picture findings from this thesis – namely, how family data can be used to understand missing heritability and gene-environment interplay in genomic research. It then summarised how each chapter contributed to these goals. The following sections will discuss general conclusions that emerge when thinking collectively about all four empirical chapters.

Missing heritability of complex traits

The Introduction examined possible reasons for missing heritability separately. However, since the genetic basis of complex traits is highly phenotype-specific, it is important to consider the missing heritability problem separately for each of the two phenotypic domains considered in the thesis.

Emotional problems including anxiety, depression and neuroticism

This thesis began and ended with studies reporting heritability estimates for childhood emotional problems. In Study 1, based on TEDS data, despite substantial twin heritabilities, SNP heritability estimates were only ~5%. Measurement is a particular problem for these anxiety and depression: in contrast to a trait such as educational attainment, which is an objective characteristic of an individual, emotional symptoms are subjective and fluctuate over time. By creating a better assessment using longitudinal twin data, three times as much genomic variance could be captured (SNP heritability = 14%). However, improved measurement could not account for the SNP-twin heritability gap.

The SNP-twin heritability gap is partly explained by rare variation for adult neuroticism, which is closely correlated phenotypically and genetically with anxiety and depression (Study 2). The role of rare variant effects specifically on anxiety and depression has yet to be investigated using methods such as GREML-KIN and LDMS-I. The use of whole genome sequence (WGS) data can capture rare variant contributions, and recover pedigree-level estimates of the heritability of height and BMI

(Wainschtein et al. 2019). The utility of WGS data for recovering heritability of more complex, socially influenced traits like emotional symptoms is not known.

Currently, the low SNP heritability of emotional problems in childhood remains a barrier to progress in identifying genetic associations and building polygenic predictors. However, this is not a universal problem. Some samples and measures already show stronger genomic signal for emotional problems. Indeed, the RDR-derived SNP heritability of childhood depressive symptoms in MoBa (the largest study to date; Chapter 5 of this thesis) was 19%. This is higher than the SNP heritability of anxiety symptoms in MoBa, and higher than SNP heritability estimates for individual anxiety and depression measures in TEDS and other samples (see Appendix 1). This estimate of 19% is more surprising given that RDR allows estimation of direct genetic effects without family environmental bias. Future research should not only continue to address the low SNP heritability of childhood emotional problems, but also seek to reconcile inconsistencies across studies. However, the context-specific nature of variance component estimates means that such inconsistencies are unlikely to be fully understood (see General Limitations).

Educational attainment

For educational attainment, twin- and extended twin family-based heritability estimates for education and cognitive ability are high (~40-50%), yet common SNPs explain little variance in population genomic studies. Notably, it is not straightforward to obtain a consistent picture of the heritability gap in the first place, because SNP heritability estimates vary depending on the method the sample, or sub-sample studied. Sub-sample variation is pertinent here. Estimates of the SNP heritability of education in the UK Biobank derived in Chapters 3 and 4 of the current thesis from standard methodologies are: 29% (se = 0.08) (GREML, N=6,500); 17% (se = 0.01) (LD score regression, N~375,000); and 17% (se = 0.01) (GREML, N=44,520). However, attention should be paid to the range of estimates, rather than specific values derived from individual studies. Reassuringly, 95% confidence intervals overlap for all these results. Taking the results together, common SNPs explain about 20% of the variance in educational attainment in unrelated participants of the UK Biobank study. This ~20% genetic variance is highly likely to be partly inflated due to passive gene-environment correlation. This is highlighted by the shrinkage in genetic variance among adoptees, also in the UK Biobank (Chapter 4).

The gap between <20% SNP heritability and ~50% twin heritability for educational attainment may be explained by rare variants, since GREML-KIN based heritability was 56%. Though, analysis of additional imputed SNPs did not validate this finding (Chapter 3). Again, whole-genome sequencing offers a way to examine the role of rarer genetic effects, but this has not yet been applied to educational attainment. Family-based designs like RDR could be used in combination with whole genome sequence data to rule out biasing factors. These could include assortative mating, population stratification, and passive gene-environment correlation. The relative contributions of individual biasing factors and the effects of interplay between these on genomic heritability estimates also have yet to be characterised.

In sum, although the missing heritability problem is still unresolved for emotional problems and educational attainment, this thesis has shown that family information provides numerous avenues for making incremental progress.

Gene-environment interplay

One central source of difficulty in studying the aetiology of complex human traits is that parents provide both environment and genes for their offspring. The resulting patterns of covariation between genes and environments lead to genetic confounding in studies of environmental effects, and environmental confounding in studies of genetic effects. Chapters 4 and 5 showed how family-based genomic methods can reveal environmental effects of parents on their children. The same methods can also help to rule out environmental confounding of estimates of genetic effects on phenotypes. This section discusses these findings in the context of other studies applying new genomic tools to family structures. These allow novel findings to be made about causal genetic and environmental effects, and the intertwining between genes and environments.

Environmental effects

To know whether putative exposure effects are truly explained by social mechanisms, research designs must consider that associations may reflect genetic inheritance. Powerful quantitative genetic tools such as the children-of-twins approach have allowed environmental routes of intergenerational transmission to be delineated from genetic routes. For example, smoking during pregnancy has an environmental influence on offspring birth weight, above and beyond genetic and shared environmental confounds (D'Onofrio et al. 2003). Until recently, this kind of genetically sensitive environmental research was largely limited to pedigree designs. Family-based genomic methods shed new light on environmental causation. In Chapter 5 it was emphasised that the use of genomic data circumvents difficulties related to identifying and measuring important parental behaviours. For example, RDR can estimate 'global' effects of parental behaviour (genetic nurture). While the classic twin design can estimate a latent environmental effect, it is not specific to parental behaviours and is not separable from passive gene-environment correlation. Without child and parent genotype data, genetic nurture effects over and above child genetic effects cannot be estimated. Chapter 4 also provided evidence of an environmental effect of educationally supportive adoptive parents on the attainment of individuals in the lowest decile of the educational attainment polygenic score. This phenomenon would not be detectable without being able to construct individual-level polygenic scores and identify strata of educational attainment polygenic scores.

Notably, other valuable areas of study harness genomic data on families to infer causal environmental effects. Mendelian randomisation (MR) is a useful method for testing causal relationships between exposures and outcomes. MR uses genetic variants as instrumental variables in a design analogous to randomised controlled trials. The random assignment of individual genotypes makes this possible. Genetic variants are valid for MR purposes if they directly influence the exposure, and only influence the outcome via the exposure. Using an MR framework, one study found that higher maternal BMI did not have a strong causal intrauterine effect on greater offspring adiposity over and above any effect of shared BMI-associated variants (Richmond et al. 2017). MR becomes even more useful when family-

based and rooted more closely in quantitative genetic principles. Within-family MR involves conditioning on parental genotypes – i.e. a genuine mendelian randomisation. This rules out bias from assortative mating, population stratification, and passive gene-environment correlation (which the authors call ‘dynastic effects’) (Davies et al. 2019). A second example is the MR-DoC twin model, which combines MR with the Direction of Causation twin model (Minică et al. 2018). Integrating the direction-of-causation twin model with MR allows testing of the assumption of no pleiotropy (the ‘exclusion restriction’). By accounting for pleiotropy, polygenic scores rather than individual variants can be used as instruments.

Inclusion of genetics in the context of research on environmental effects has relevance for policy and intervention. In the context of family-level exposures, genetically sensitive designs are needed to suggest whether policies or interventions that affect one generation will have consequences in the next. For example, if the intergenerational transmission of depression symptoms is partially socially mediated (as suggested in Chapter 5 of this thesis), then attempts to target modifiable aspects of parental depression will have implications for the distribution of depression symptoms in the next generation. If the intergenerational process is primarily due to the transmission of genes related to depression symptoms, then interventions may need to be applied across both generations (assuming constancy in the genotype–phenotype relationship).

Genetic effects

As much as genetic confounding needs to be accounted for when estimating environmental effects, environmental confounding needs consideration when estimating genetic effects on complex phenotypes. Some have drawn parallels between GWAS and studies estimating many small associations between traits and putative environments across the ‘exposome’ without any genetic controls (Turkheimer 2012). In genomics, samples are typically restricted to unrelated individuals and ancestry principal components are included. However, this might not be enough to eliminate environmental confounding, especially for highly socially contingent traits. As with the problem of genetic confounding in social science, quantitative genetic principles can take care of environmental confounding in genomics. In the largest genomic study of educational attainment, the authors presented results from a GWAS comparing 22,135 sibling pairs (Lee et al. 2018). Due to random segregation, the variability around the average genetic relationship between siblings of 50% is independent of passive gene-environment correlation and population stratification. Effect sizes from the sibling GWAS of education were ~40% smaller than effect sizes derived from standard analysis of an unrelated sample. This was not the case for their negative control variable, height. The effect magnitude is consistent with the findings from Chapter 4 of this thesis, in which polygenic prediction accuracy was 50% lower for adopted compared to non-adopted individuals.

Strong support for family environmental confounding of polygenic score associations with educational traits also comes from the use of DZ twin comparisons. This approach strengthens the sibling design, since many important environmental influences, such as pregnancy exposures, parental age, and family socio-economic status, are time-invariant between co-twins. Interestingly, attenuation of polygenic score associations in TEDS when moving from between- to within-family effect sizes was

large for cognitive traits (60%), but negligible for others including childhood behaviour problems and anthropometric traits (Selzam et al 2019). This mirrors the pattern of shared environmental variance when applying the classical twin design, and indeed passive gene-environment correlation loads as shared environment in such models. Hence, twin-based estimates of shared environmental influence may offer a guide as to the importance of environmental confounding in genomic studies of specific traits.

An important ongoing effort by the Within Family GWAS Consortium led by Bristol University is to collect sibling (including DZ) genotypes on a wide range of complex phenotypes. This will reveal phenotype-specific nuances of family environmental confounds in GWAS, and allow construction of polygenic scores with fewer sources of bias. This is not only an intrinsically important academic contribution, but also useful for other purposes: for example, it can help create better instruments for MR.

Developments in family genomic datasets and methodologies are socially as well as scientifically important. Interpretations about the origins of complex traits based on non-family-based designs, with no principled way to rule out environmental confounding, are potentially dangerous. This is most obvious when it comes to discussions about the use of polygenic scores for individual-level precision medicine and schooling. If the scores contain not only direct genetic effects but indirect effects of relatives' genomes, then there is potential for individuals to be penalised or rewarded for characteristics that are not personal but circumstantial.

If the goal is to explain phenotypic variation with respect to genetic and environmental causes and mechanisms, parsing out the relative roles of direct and indirect genetic effects is clearly important. Notably, for other purposes, it may not matter if family environmental effects are included in estimates of genetic variance. For example, as mentioned in Chapter 3, researchers interested in maximising prediction might not be concerned with drilling down to the direct mechanisms underpinning polygenic score prediction. Similarly, in studies aiming to precisely identify causal exposure effects, polygenic scores capturing direct and indirect pathways would make the strongest control variables. Furthermore, to estimate the magnitude of family confounding, between-family methods still need to be deployed for comparison to within-family designs.

Gene-environment correlation

It should be reiterated that environmental confounding is not the only important reason for genomic research to recognise the complexities of the social environment. Processes of gene-environment correlation need to be considered in order to build a full understanding of how genotypes lead to phenotypes across the life course. For complex behavioural traits, even if influential direct genetic effects with 'proximal' biological mechanisms are discovered, it is likely that this genetic variation still operates through human social contexts. This involves processes such as evocative and active gene-environment correlation (e.g. variants associated with lung cancer risk taking effect through smoking behaviour).

General limitations

This thesis has incorporated a range of study samples, models and phenotypes. This variety of approaches helps to mitigate some of the disadvantages of the individual studies. Nonetheless, some general limitations remain.

Comparability of results from the chapters

For practical reasons, the samples, methods and phenotypes used in this thesis varied depending on the research question, precluding close comparisons. For example, using TEDS data, it was possible to estimate the extent that more accurate measurement could increase the SNP heritability of childhood emotional problems. However, due to the lack of other kinds of family relationships, we could not also apply the GREML-KIN model to test whether including rarer variants increases SNP heritability for the same phenotype. It would also be valuable to compare results from using the adoption and RDR methods in the same sample in the same phenotype. Consistency across different methods for separating direct and indirect genetic effects – i.e. triangulation (Munafò and Davey Smith 2018) – would provide stronger evidence. Existing methods are not equivalent, partly because they involve different assumptions and biases, but these inconsistencies are useful for more precisely pinpointing the ratio of indirect to direct genetic effects and quantifying the contributions of population phenomena. For example, comparisons of effect sizes in adoptees and non-adopted individuals are robust to assortative mating and population stratification, assuming that these factors do not systematically differ across the two groups. RDR estimates are not. However, since MoBa is a birth cohort comprising mothers and their biologically related children, there is no way to employ the adoption design. Similarly, there are not enough genotyped trios in the UK Biobank to use RDR. In the future, if samples like MoBa and UK Biobank are extended to include additional genotyped family members, triangulation will become easier. Nevertheless, it is a strength that all Chapters of this thesis present multiple analytical strategies e.g. RDR and pedigree modelling in Chapter 5.

Generalisability of findings

Variance components estimated in all chapters must be interpreted as **population-specific and assumption bound**. No single heritability estimate can be considered definitive for a given trait. It is difficult to compare results even for subsamples of the same study. For example, it may seem surprising that the heritability of educational attainment in the UK Biobank differed depending on the subsample used. However, we have no reason to expect them to be equal, because they were derived from different groups, of different sizes. Similarly, we have no reason to assume that heritability estimates for childhood depression will be identical across different social contexts. Although the same instrument was used to measure depression symptoms in both TEDS and MoBa, and both in middle childhood, symptoms might be more heritable in MoBa due to factors that can be genetic or environmental. Reasons for heritability differences between different contexts can include: allele frequency differences (e.g. from selection or inbreeding); genetic innovation in one population (e.g., due to migration or mutation); gene-environment interactions meaning that genetic variation contributes more in the presence of certain social factors. It is possible that there is less environmental

variation impacting on childhood depression symptoms in MoBa than in TEDS, resulting in greater genetic variance in the former sample.

Importantly, there is much yet to learn about how genotype-environment interactions affect the generalisability of genomics findings across environments in the sense of **socio-cultural and political context**. There are some striking examples of how implications of genetic variants associated with complex traits only refer to the group of people studied and the particular time and place in which they lived. For instance, the heritability of educational attainment in Estonia was twice the magnitude for individuals brought up during the post-Soviet era compared to those growing up during the Soviet era, probably due to diminished equality of opportunity during the communist regime (Rimfeld et al. 2018). In this thesis, the environmentally mediated parental genetic effect on child depressive symptoms (Chapter 5) may be underestimated due to selection bias and attrition reducing the presence of less advantaged families in MoBa. This is because the effect of maternal depression on diverse child outcomes is likely to be stronger in the context of poverty (Goodman et al. 2011). Notably, cautious and socially-contextualised interpretations of estimates of 'direct' genetic effects based on adoptees or RDR are also still essential. Even though these are principled methods for estimating heritability free of many confounds, the above limitations regarding sample specificities still apply. A specific estimate is never the 'true' value of genetic variance for a trait in any general sense, and genetic effects cannot manifest independently of environmental context.

Clearly, the findings presented in this thesis may only be generalisable to the study populations at the specific ages tested, and may not replicate directly in other populations with different educational, social, or political settings. Importantly, all findings are based on white European individuals and may not be generalisable to other **ancestries**. Perhaps the single biggest issue in contemporary genomics is that only ~10% of individuals in GWAS and genotyped samples have non-European ancestries (Mills and Rahal 2019). The problem is particularly obvious with respect to polygenic scores, which give systematically worse predictions when applied to individuals from populations that are ancestrally different to European discovery samples. For example, the prediction accuracy of the years of education polygenic score was 10.6% for European-Americans but only 1.6% for African Americans (Lee et al., 2018). This lower accuracy is explained by population differences in allele frequencies and linkage disequilibrium. Although the variants influencing complex traits are largely the same across human populations, the allele frequencies can vary substantially, which is mainly due to genetic drift. Also, differences between populations in linkage disequilibrium arising from different demographic histories and recombination events means that the populations may differ in how causal sites correlate with other parts of the genome. This ancestry bias may lead to unfair benefits to people with white European ancestry at the expense of others, in terms of scientific understanding as well as clinical applications (Mostafavi et al. 2020).

Going forward, ancestry bias will hopefully be eliminated through the recruitment of large and ancestrally diverse genotyped samples. This is already underway, for example, the East London Genes & Health project, which currently includes genotype data for ~35,000 participants of South Asian ancestry with the goal to reach 100,000 (Finer et al. 2019). Countering ancestry bias in

genomics will also bear fruit beyond the representativeness of results. In particular, it will aid the identification of causal genetic variants, by a) characterising cross-population patterns of linkage disequilibrium at greater resolution, b) capturing variants that may be rare in European samples, and c) improving sample size (Claussnitzer et al. 2020). For example, a recent whole-exome sequencing study of schizophrenia in a sample of the Xhosa population of South Africa revealed new novel risk genes across the allele frequency spectrum. Power to detect rarer genetic effects on schizophrenia was aided by the great degree of genetic diversity in African populations, which exists because the majority of human genetic diversity evolved before the out of Africa migration 50,000 to 100,000 years ago (Gulsuner et al. 2020).

In sum, estimates of genetic variance always lack generalisability in the sense that they are population specific. Here, this lack of generalisability is compounded by ancestry and socio-economic status biases, which are likely to affect the present results in ways that are not easy to account for. However, these issues are general conundrums in genetics and are not more severe here than in other current work. Making use of data from classic quantitative genetic studies could help to alleviate selection biases. For example, many contemporary GWAS samples are drawn from clinical case-control samples, opt-in biobanks and direct-to-consumer companies, and thus tend to include wealthier and better educated people than the population average. In contrast, twin studies tend to be more representative of national populations (Mostafavi et al. 2020).

Unknown quantities in genomics

As well as the biases relating to generalisability mentioned above, there are sources of variation in traits which would be good candidates for understanding missing heritability and gene-environment interplay. These are: gene-environment interactions, assortative mating, non-additivity, and sibling interactions. However, these factors are currently relatively poorly understood, and their roles in shaping the current findings have not been examined. These effects differ in salience in general, as well as in their importance for specific phenotypes.

Gene-environment interactions

The prior plausibility of gene-by-environment interactions affecting complex traits is high, and such interactions are of interest for many reasons. For example, identification of specific genetic variants and environmental measures with proven interactive effects could provide hints to aid the design of effective stratified policies as well as mechanistic genetic research. Variants that are ‘environmentally sensitive’ would make invalid instruments in MR and so their removal would allow more accurate causal inference. Also, pertinent to this thesis, gene-by-shared environment interactions may inflate twin heritability estimates and may also affect SNP heritability estimates.

However, it has proven difficult in genome-wide research, and also in prior quantitative genetic and candidate gene work, to identify robust examples of specific *measured* environmental exposures interacting with genetic effects. For example, four studies assessing interactions between genetic predisposition to depression and reported trauma have yielded four different patterns of results (Coleman et al. 2020; Peyrot et al. 2018; Colodro-Conde et al. 2017; Mullins et al. 2016). Gene-

environment interaction research is hampered by many factors. First, as discussed in Chapter 5, identification and measurement of important environmental factors is difficult, often expensive, and heterogeneous across studies. Second, environmental measures are often themselves genetically influenced, and this non-independence of genes and environment confuses interpretations of interaction studies. Third, small effect sizes of common variants affecting complex traits mean that huge sample sizes are required to achieve adequate power to detect interaction effects with environmental measures. Other reasons include the use of small sample sizes, candidate genes, and incorrectly specified statistical models and publication bias (Keller 2014).

Nonetheless, exciting examples of gene-environment interactions have been identified using methods that circumvent some of the challenges described above. One example is that polygenic scores can be used in combination with econometrics-style regression discontinuity designs leveraging random environmental ‘shocks’, such as policy changes (Barcellos et al. 2018). This avoids the issue that environmental moderators are often not exogenous. A second example is that, researchers can search for SNPs that create trait variance (known as variance QTLs (vQTLs)). This circumvents the problem of measuring the environment. A recent study detected several vQTLs affecting obesity-related traits, including FTO, a locus with replicated environmental interactions (Wang et al. 2019). The Future Directions section discusses how family-based genomic approaches might also be used to investigate gene-environment interactions.

Assortative mating

Recent work has demonstrated that assortative mating affects variation in many complex phenotypes including psychiatric traits (Nordsletten et al. 2016). Innovative study designs have been used to quantify the signature of assortative mating on the genome. Properties of germline DNA suggest that genotype data can be useful for disentangling the reasons for couple similarity. For example, Yengo et al. (2018) detected correlations between polygenic scores for educational attainment constructed from odd and even chromosomes. This is indicative of positive assortative mating on education because without assortment, trait-associated alleles on different chromosomes are not expected to correlate, due to independent segregation of chromosomes. However, the overall importance of assortative mating, in terms of the variance in specific phenotypes it explains in genomic models such as GREML and RDR, remains unknown. Chapter 3 offered speculation on how assortative mating may affect variance components in the GREML-KIN model, yet more research is needed to quantify this.

Non-additivity and sibling interaction

Non-additivity and sibling interaction are two possible explanations for missing heritability because they are detected in twin but not SNP heritability estimates. However the quantitative genetic literature suggests that neither are likely to be important causes of complex trait variation or explanations of missing heritability. Methodological work in quantitative genetics suggests that, even if genetic variants act in complex, interactive, nonlinear ways, most variance is additive, so linear models work well (Hill et al. 2008). Competitive sibling interaction effects could partly explain why the RDR method yields systematically lower heritability estimates than the twin design. In the latter design, but not the former, such effects mimic genetic dominance and thus inflate heritability estimates (Eaves 1976). However,

quantitative genetic evidence suggests that competitive sibling interaction effects are uncommon. Specifically, competitive sibling interaction would lead DZ twin pairs reared apart to be more similar than DZ twin pairs reared together, as they do not interact and differentiate from each other. Yet correlations for DZ twins reared apart, at least for some traits, are equal or lower than correlations for DZ reared together (Pedersen et al. 1989).

Future directions

Several important directions for future research have already been highlighted, for example the development of samples with diverse ancestries, and the use of whole genome sequence data to capture the full allele frequency spectrum of genetic effects. The following section homes in on specific ways in which **family-based quantitative genetic research charts the direction for complex trait genomics**. Potential uses of more **comprehensive modelling of family relationships in genomics** are described.

Estimating various environmental effects and disentangling biases

Family-based genomic methods like RDR present exciting new opportunities to measure environmental effects of parents on offspring traits. Such methods could be extended to incorporate other family relationships, therefore allowing additional sources of environmental influence and gene-environment covariance to be estimated. Incorporating additional relationships into genomic methods will also help to disentangle genetic nurture effects from assortative mating and population stratification. Importantly, in the field of quantitative genetics, many extensions of the classical twin design already exist, which provide a wealth of inspiration for family-based genomic methods.

In the Introduction and in Chapter 5, the children-of-twins design was mentioned as a method to estimate environmental effects of parents, over and above the effects of genes that parents share with their children. The design uses data on adult twin pairs and their children. Offspring of ‘identical’ monozygotic twins are as genetically related to their parent’s twin as they are to their own parent (50%). However, only the parent provides the rearing environment. If the parent-child correlation is greater than the correlation between child and aunt/uncle, this indicates that living together contributes to the association between parent and offspring phenotypes above and beyond genetic confounding. In the context of intergenerational genetically sensitive studies, we know that density of relatives is important. For example, it has also been shown that adding siblings in the child generation to yield a ‘Multiple children of twins’ (MCoT) model results in large gains in statistical power – more so than is the case when adding a whole new family to the analysis (McAdams et al. 2018).

These ideas could be borrowed to design useful extensions of genomics methods. Genotyped siblings in the parent generation could be used as a new way to estimate genetic nurture, controlling for family and population biases. Specifically, phenotypic differences between cousins in the child generation could be regressed on genome-wide genetic differences between siblings in the parent generation. The latter genetic differences are due to random segregations of genetic material in the parents during

meiosis and are therefore uncorrelated with population stratification effects. This method would give estimates of genetic nurture that are robust to population stratification. We are testing this in MoBa, in which there is a growing sample of more than 1,200 genotyped siblings in the parent generation (15,000 sibling pairs when all are genotyped). With the presence of genotyped siblings and half-siblings in the parent generation, there is sufficient information to 'impute' unobserved grandparental genotypes (Hwang et al. 2020). This development opens up a range of exciting possibilities. These include estimating indirect genetic effects of siblings together with parental genetic nurture effects, assessing questions about intergenerational transmission across three generations, and measuring the importance of indirect genetic effects on adult outcomes, which is difficult since datasets with genotyped trios are generally young.

Estimating how individuals shape their experiences

Family-based genomic methods afford opportunities to measure not only effects of parents on offspring traits, but of the reverse: child-driven effects on parent traits. Examples of heritable child behaviour influencing parenting are abundant in the twin literature. If an environmental measure analysed as a child-specific trait is heritable, this is argued to be evidence of evocative genotype environment correlation, because the design reflects the influence of children's genes. For example, high child anxiety was found to evoke maternal control, which was partially influenced by the same genetic factors as the anxiety itself (Eley et al. 2010). We are currently combining ideas about evocative gene-environment correlation and genomics methods to study maternal depression during (also before and after) pregnancy. There is a large literature on how maternal perinatal depression affects the developing child, but few studies are genetically sensitive (Hannigan et al. 2018). What's more, no studies seem to examine the opposite direction of effects, whereby genetically influenced traits manifesting in the developing child might play a part in the causation of maternal perinatal depressive symptoms. The MCoTs model cannot rule out that phenotypic associations between parent and child remaining after controlling for shared genes are operating from child to parent, but genomic methods like RDR and m-GCTA *can* do this. In quantitative genetics, it is well understood that parent-child relationships are transactional (Deater-Deckard and O'Connor 2000; Deater-Deckard and Petrill 2004; Ahmadzadeh et al. 2019; Tucker-Drob and Harden 2012). However, this emphasis on bidirectional relationships over time has yet to emerge strongly in the genomics literature. Genomics methods are being developed by leveraging ideas and family data from quantitative genetics.

Investigating gene-environment interactions

In the Limitations section, novel methods to detect specific loci that interact with the environment were described. Instead of applying such methods in unrelated samples, an analogous way forward is to capitalise on phenotypic differences between monozygotic (MZ) twins. MZ pairs that are more phenotypically different may hold more environmentally sensitive genetic loci i.e. vQTLs. This approach has been used for a long time in quantitative genetics, for example with respect to cholesterol (Martin et al. 1983). In the genome-wide era, a polygenic score derived from a GWAS of MZ phenotypic differences has been shown to moderate the effects of parenting and treatment

response (Keers et al. 2016). The Within Family GWAS Consortium plan to build on these findings by aggregating samples together for a GWAS meta-analysis of MZ differences. This is another good example of family studies guiding the future of genomics. It would be interesting, when samples are large enough, to test whether stronger vQTL effects are detectable in older twins, who would have had more time to diverge phenotypically compared to younger twins. However, the MZ differences GWAS meta-analysis project will face challenges. First, MZ differences are usually small in magnitude and show weak temporal stability. Second, difference scores contain measurement error. Third, as has been argued previously (Conley 2016), studying gene-environment interactions is challenging in the meta-analysis context. When data are aggregated across cohorts, environments, and periods, there is less chance of detecting genetic associations that are contingent upon specific social environments. Therefore, gene-environment interaction studies could benefit from the collection of large homogeneous samples including extended families who share many aspects of their environment, such as MoBa. Clearly, going forward in genome-wide gene-environment interaction research, and beyond, family-based datasets will be extremely useful.

Answering longitudinal and multivariate questions with family genomic data

Currently, multi-generational genomic data is available in few samples, but such datasets tend to be large population-representative cohorts like MoBa, TEDS and ALSPAC. This kind of study has rich longitudinally assessed data on traits and environmental exposures, which should be leveraged as much as possible to obtain detailed understanding of aetiological influences on complex traits across the life-course. There is much yet to be understood about the 'developmentally-dynamic' genome (Kendler et al. 2008), including attenuation, amplification, and innovation in SNP effects on complex behaviours. Some exciting recent genomic research has suggested that certain patterns and trajectories of developmental mental health symptoms, such as early onset chronic depression, and joint emotional and conduct problems, are particularly strongly associated with polygenic risk for depression (Kwong et al. 2019; Hannigan, Pingault, et al. 2018). Future studies could test whether patterns and trends are differentially predicted by multiple genetic and environmental predictors, and their correlations and interactions. Robust within-family genomic methods will give the most convincing evidence concerning how environments and genes function together over development.

Conclusions

Genomic research is transforming our understanding of the influences on complex trait variation. Quantitative genetics charts the direction for future genomic research. Family datasets will be not only essential for deepening insights about effects of genes, environments and their interplay, but also for understanding sources of bias. This PhD has demonstrated that the most valuable research designs will be not only genetically but environmentally informative.

References

- Ahmadzadeh, Y.I., Eley, T.C., Leve, L.D., et al. 2019. Anxiety in the family: a genetically informed analysis of transactional associations between mother, father and child anxiety symptoms. *Journal of Child Psychology and Psychiatry, and Allied Disciplines*.
- Barcellos, S.H., Carvalho, L.S. and Turley, P. 2018. Education can reduce health differences related to genetic risk of obesity. *Proceedings of the National Academy of Sciences of the United States of America* 115(42), pp. E9765–E9772.
- Claussnitzer, M., Cho, J.H., Collins, R., et al. 2020. A brief history of human disease genetics. *Nature* 577(7789), pp. 179–189.
- Coleman, J.R.I., Peyrot, W.J., Purves, K.L., et al. 2020. Genome-wide gene-environment analyses of major depressive disorder and reported lifetime traumatic experiences in UK Biobank. *Molecular Psychiatry*.
- Colodro-Conde, L., Couvy-Duchesne, B., Zhu, G., et al. 2017. A direct test of the diathesis-stress model for depression. *Molecular Psychiatry* 23(7), pp. 1590–1596.
- Conley, D. 2016. Socio-Genomic Research Using Genome-Wide Molecular Data. *Annual review of sociology* 42(1), pp. 275–299.
- D’Onofrio, B.M., Turkheimer, E.N., Eaves, L.J., et al. 2003. The role of the children of twins design in elucidating causal relations between parent characteristics and child outcomes. *Journal of Child Psychology and Psychiatry, and Allied Disciplines* 44(8), pp. 1130–1144.
- Davies, N.M., Howe, L.J., Brumpton, B., Havdahl, A., Evans, D.M. and Davey Smith, G. 2019. Within family Mendelian randomization studies. *Human Molecular Genetics* 28(R2), pp. R170–R179.
- Deater-Deckard, K. and O’Connor, T.G. 2000. Parent-child mutuality in early childhood: two behavioral genetic studies. *Developmental Psychology* 36(5), pp. 561–570.
- Deater-Deckard, K. and Petrill, S.A. 2004. Parent-child dyadic mutuality and child behavior problems: an investigation of gene-environment processes. *Journal of Child Psychology and Psychiatry, and Allied Disciplines* 45(6), pp. 1171–1179.
- Eaves, L. 1976. A model for sibling effects in man. *Heredity* 36(2), pp. 205–214.
- Eley, T.C., Napolitano, M., Lau, J.Y.F. and Gregory, A.M. 2010. Does childhood anxiety evoke maternal control? A genetically informed study. *Journal of Child Psychology and Psychiatry, and Allied Disciplines* 51(7), pp. 772–779.
- Finer, S., Martin, H.C., Khan, A., et al. 2019. Cohort profile: East London genes & health (ELGH), a community-based population genomics and health study of British Bangladeshi and British Pakistani people. *International Journal of Epidemiology*.
- Goodman, S. H., Rouse, M. H., Connell, A. M., Broth, M. R., Hall, C. M., & Heyward, D. (2011). Maternal depression and child psychopathology: a meta-analytic review. *Clinical Child and Family Psychology Review*, 14(1), 1–27. doi:10.1007/s10567-010-0080-1
- Gulsuner, S., Stein, D.J., Susser, E.S., et al. 2020. Genetics of schizophrenia in the South African Xhosa. *Science* 367(6477), pp. 569–573.
- Hannigan, L.J., Eilertsen, E.M., Gjerd, L.C., et al. 2018. Maternal prenatal depressive symptoms and risk for early-life psychopathology in offspring: genetic analyses in the Norwegian Mother and Child Birth Cohort Study. *The Lancet. Psychiatry* 5(10), pp. 808–815.
- Hill, W.G., Goddard, M.E. and Visscher, P.M. 2008. Data and theory point to mainly additive genetic variance for complex traits. *PLoS Genetics* 4(2), p. e1000008.

- Hwang, L.-D., Tubbs, J.D., Luong, J., et al. 2020. Estimating indirect parental genetic effects on offspring phenotypes using virtual parental genotypes derived from sibling and half sibling pairs. *BioRxiv*.
- Keller, M.C. 2014. Gene $\times$ environment interaction studies have not properly controlled for potential confounders: the problem and the (simple) solution. *Biological Psychiatry* 75(1), pp. 18–24.
- Kendler, K.S., Gardner, C.O., Annas, P., Neale, M.C., Eaves, L.J. and Lichtenstein, P. 2008. A longitudinal twin study of fears from middle childhood to early adulthood: evidence for a developmentally dynamic genome. *Archives of General Psychiatry* 65(4), pp. 421–429.
- Lee, J.J., Wedow, R., Okbay, A., et al. 2018. Gene discovery and polygenic prediction from a genome-wide association study of educational attainment in 1.1 million individuals. *Nature Genetics* 50(8), pp. 1112–1121.
- McAdams, T.A., Hannigan, L.J., Eilertsen, E.M., Gjerd, L.C., Ystrom, E. and Rijdsdijk, F.V. 2018. Revisiting the Children-of-Twins Design: Improving Existing Models for the Exploration of Intergenerational Associations. *Behavior Genetics* 48(5), pp. 397–412.
- Mills, M.C. and Rahal, C. 2019. A scientometric review of genome-wide association studies. *Communications Biology* 2(1), p. 9.
- Minică, C.C., Dolan, C.V., Boomsma, D.I., de Geus, E. and Neale, M.C. 2018. Extending Causality Tests with Genetic Instruments: An Integration of Mendelian Randomization with the Classical Twin Design. *Behavior Genetics* 48(4), pp. 337–349.
- Mostafavi, H., Harpak, A., Agarwal, I., Conley, D., Pritchard, J.K. and Przeworski, M. 2020. Variable prediction accuracy of polygenic scores within an ancestry group. *eLife* 9.
- Mullins, N., Power, R.A., Fisher, H.L., et al. 2016. Polygenic interactions with environmental adversity in the aetiology of major depressive disorder. *Psychological Medicine* 46(4), pp. 759–770.
- Munafò, M.R. and Davey Smith, G. 2018. Robust research needs many lines of evidence. *Nature* 553(7689), pp. 399–401.
- Nordsletten, A.E., Larsson, H., Crowley, J.J., Almqvist, C., Lichtenstein, P. and Mataix-Cols, D. 2016. Patterns of nonrandom mating within and across 11 major psychiatric disorders. *JAMA psychiatry* 73(4), pp. 354–361.
- Pedersen, N.L., Gatz, M., Plomin, R., Nesselroade, J.R. and McClearn, G.E. 1989. Individual differences in locus of control during the second half of the life span for identical and fraternal twins reared apart and reared together. *Journal of Gerontology* 44(4), pp. P100–5.
- Peyrot, W.J., Van der Auwera, S., Milaneschi, Y., et al. 2018. Does Childhood Trauma Moderate Polygenic Risk for Depression? A Meta-analysis of 5765 Subjects From the Psychiatric Genomics Consortium. *Biological Psychiatry* 84(2), pp. 138–147.
- Richmond, R.C., Timpson, N.J., Felix, J.F., et al. 2017. Using genetic variation to explore the causal effect of maternal pregnancy adiposity on future offspring adiposity: A mendelian randomisation study. *PLoS Medicine* 14(1), p. e1002221.
- Rimfeld, K., Krapohl, E., Trzaskowski, M., et al. 2018. Genetic influence on social outcomes during and after the Soviet era in Estonia. *Nature human behaviour* 2(4), pp. 269–275.
- Tucker-Drob, E.M. and Harden, K.P. 2012. Early childhood cognitive development and parental cognitive stimulation: evidence for reciprocal gene-environment transactions. *Developmental Science* 15(2), pp. 250–259.

Turkheimer, E. 2012. Genome wide association studies of behavior are social science. In: Plaisance, K. S. and Reydon, T. A. C. eds. *Philosophy of behavioral biology*. Boston studies in the philosophy of science. Dordrecht: Springer Netherlands, pp. 43–64.

Wainschtein, P., Jain, D.P., Yengo, L., et al. 2019. Recovery of trait heritability from whole genome sequence data. *BioRxiv*.

Wang, H., Zhang, F., Zeng, J., et al. 2019. Genotype-by-environment interactions inferred from genetic effects on phenotypic variability in the UK Biobank. *Science Advances* 5(8), p. eaaw3538.

Yengo, L., Robinson, M.R., Keller, M.C., Kemper, K.E., Yang, Y., Trzaskowski, M., Gratten, J., Turley, P., Cesarini, D., Benjamin, D.J., Wray, N.R., Goddard, M.E., Yang, J. and Visscher, P.M. 2018. Imprint of assortative mating on the human genome. *Nature human behaviour* 2(12), pp. 948–954.

Appendices

Contents:

Appendix 1 – Supplementary Information for Chapter 2

Appendix 2 – Supplementary Information for Chapter 3

Appendix 3 – Supplementary Information for Chapter 4

Appendix 4 – Supplementary Information for Chapter 5

Appendix 1. Supplementary Information for Chapter 2 (Extracting stability increases the SNP heritability of emotional problems in young people).

Supplementary Table 1: Literature review: SNP heritability of child and adolescent emotional problems

Reference	Year	Trait	Sample	Measure	Age	Reporter	N	SNP h ²	SE	P
Benke et al.	2014	Emotional problems	Gen R	CBCL 1.5-5 internalising scale	3	Primary caregiver	2037	0.26	NA	0.07
			NTR	CBCL 2-3 internalising scale	3	Mother	1475	0.18	NA	0.3
			Raine	CBCL 2-3 internalising scale	2	Primary caregiver	1084	0.13	NA	0.33
Pappa et al.	2015	Emotional problems	NTR / Gen R	CBCL internalising	7	Parent	3175	0.12	0.1	
Sallis et al.	2017	Depressive Symptoms	ALSPAC	Short Moods and Feelings Questionnaire (SMFQ)	11	Child	5480	0.07	0.1	
					13	Child	5055	0.17	0.1	
					14	Child	4615	0.02	0.1	
					17	Child	3605	0.08	0.1	
					18	Child	3289	0.03	0.1	
					19	Child	2433	0.05	0.1	
					10	Parent	5571	0.04	0.1	
					12	Parent	5119	0.12	0.1	
					13	Parent	5031	0.07	0.1	
					17	Parent	3990	0.04	0.1	

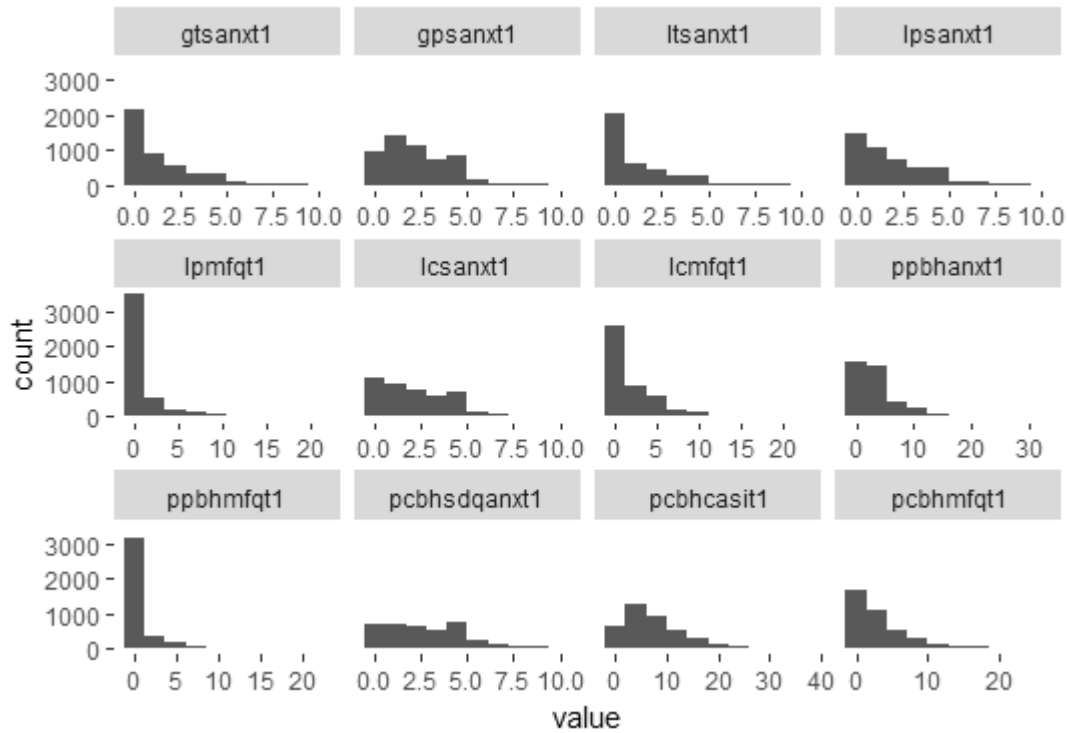
Note: Gen R= Generation R; NTR= Netherlands Twin Register; ALSPAC= Avon Longitudinal Study of Parents and Children; CBCL = Child Behaviour Checklist.

Three studies in samples other than TEDS have estimated SNP heritabilities for quantitative measures of childhood anxiety and depression, known as emotional problems, or internalising'. The average SNP heritability estimate from these studies is 10%. Although this is twice as high as the average in this study from the TEDS sample (5%), 10% SNP heritability is still well below half of the twin estimates of these phenotypes in the literature (40-60% in TEDS). Larger samples are needed to clarify this inconsistent and underpowered literature. For example, the study of pre-school internalising symptoms (Benke et al., 2014) had maximum 35% power with a sample of 2000 if true SNP heritability was half of the twin estimate they cite (50%), and a minimum of 10% power with a sample of 1000 and heritability of 20%. Moreover, the large ranges within the studies makes the estimates difficult to interpret.

Supplementary Table 2: Descriptive statistics for the 12 TEDS anxiety and depression scales

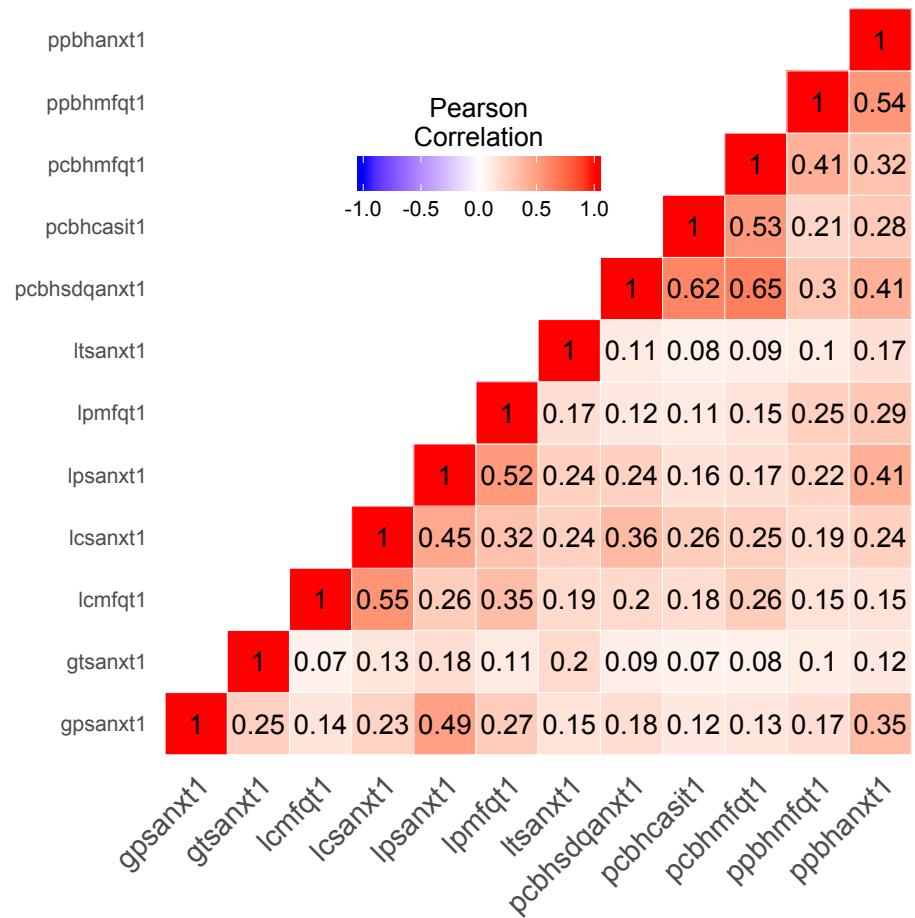
	Measure	Sample size (twin pairs)	No. items	Mean	Standard deviation
1	Parent-rated SDQ at 7	5367	5	2.13	1.80
2	Teacher-rated SDQ at 7	4419	5	1.30	1.77
3	Parent-rated MFQ at 12	4486	11	1.13	2.24
4	Parent-rated SDQ at 12	4487	5	1.77	1.91
5	Self-rated MFQ at 12	4493	11	2.30	3.28
6	Self-rated SDQ at 12	4475	5	2.17	2.05
7	Teacher-rated SDQ at 12	3713	5	1.16	1.77
8	Parent-rated ARBQ at 16	3845	19	3.43	4.08
9	Parent-rated MFQ at 16	3844	11	0.96	2.28
10	Self-rated CASI at 16	3831	18	7.85	5.86
11	Self-rated MFQ at 16	3832	13	3.59	4.37
12	Self-rated SDQ at 16	3830	5	2.76	2.27

Supplementary Figure 1: Histograms of TEDS anxiety and depression scales



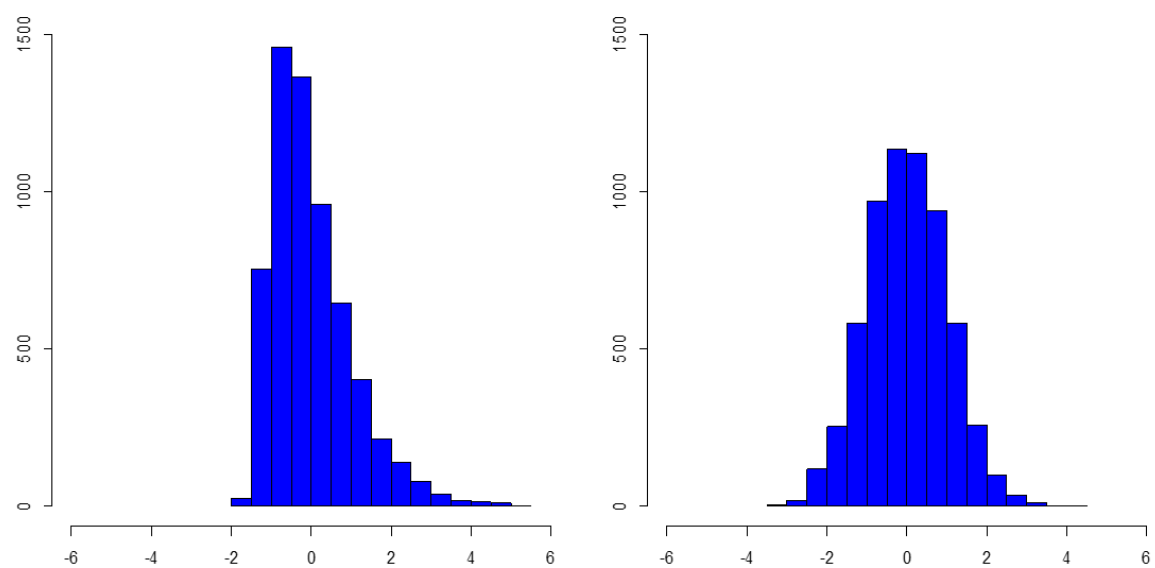
Note: 'gtsanxt1' = Teacher-rated SDQ at 7, 'gpsanxt1' = Parent-rated SDQ at 7, 'ltsanxt1' = Teacher-rated SDQ at 12, 'lpsanxt1' = Parent-rated SDQ at 12, 'lpmfqt1' = Parent-rated MFQ at 12, 'lcsanxt1' = Self-rated SDQ at 12, 'lcmfqt1' = Self-rated MFQ at 12, 'ppbhanxt1' = Parent-rated ARBQ at 16, 'ppbhmft1' = Parent-rated MFQ at 16, 'pcbhsdqanxt1' = Self-rated SDQ at 16, 'pcbhcasit1' = Self-rated CASI at 16, 'pcbhmft1' = Self-rated MFQ at 16.

Supplementary Figure 2: Phenotypic correlations of the 12 anxiety and depression scales at ages 7, 12 and 16



Note: correlations were calculated in R, before any data cleaning, and including only the people with genotype data and complete data for these 12 scales. See Note of Supplementary Figure 1 for variable names.

Supplementary Figure 3: histograms of CFA- (left) and IRT-derived (right) stable emotional problems scores



Note: the correlation between CFA and IRT factor scores is 0.92.

CFA model fit and factor loadings

Supplementary Table 3: Twin correlations for emotional problems stable latent factor from CFA

Twin correlations

MZ	0.85
DZ	0.489

Supplementary Table 4: Model Fit

The single-factor model did not provide a good fit to the data. CFI and TLI were less than 0.9, suggesting worse fit than a restricted baseline model, and RMSEA was significantly greater than 0.05, suggesting poor fit. However, poor fit might be expected: it could reflect the lack of stability of anxiety across childhood (as indicated by the low correlations across our measures (heat map above), and the low stability in the literature), or it could reflect the higher than modelled associations within raters and within time-points.

Fit indices for the CFA model of childhood emotional problems

CFI	0.68
TLI	0.69
RMSEA	0.08

Note: RMSEA = Root-Mean-Square Error of Approximation; CFI = Comparative Fit Index; TLI = Tucker-Lewis Index.

Supplementary Table 5: Standardised Factor Loadings

With regard to the CFA model, all measures load significantly onto the stable emotional problems factor. All measures loaded ≥ 0.5 , apart from the two teacher-rated measures, and self-rated MFQ at 12.

Measure	Factor Loading	SE
Parent-rated SDQ at 7	0.51	0.00
Teacher-rated SDQ at 7	0.42	0.01
Self-rated MFQ at 12	0.33	0.01
Self-rated SDQ at 12	0.61	0.01
Parent-rated SDQ at 12	0.50	0.01
Parent-rated MFQ at 12	0.62	0.01
Teacher-rated SDQ at 12	0.49	0.01
Self-rated SDQ at 16	0.61	0.01
Self-rated CASI at 16	0.56	0.01
Self-rated MFQ at 16	0.73	0.01
Parent-rated MFQ at 16	0.59	0.01
Parent-rated ARBQ at age 16	0.68	0.01

Supplementary Table 6: Residual variances

Measure	Residual variance	SE
Parent-rated SDQ at 7	0.74	0.00
Teacher-rated SDQ at 7	0.82	0.01
Self-rated MFQ at 12	0.89	0.01
Self-rated SDQ at 12	0.63	0.01
Parent-rated SDQ at 12	0.75	0.01
Parent-rated MFQ at 12	0.61	0.01
Teacher-rated SDQ at 12	0.76	0.01
Self-rated SDQ at 16	0.63	0.01
Self-rated CASI at 16	0.69	0.01
Self-rated MFQ at 16	0.47	0.01
Parent-rated MFQ at 16	0.65	0.01
Parent-rated ARBQ at age 16	0.54	0.01

Full genotyping information

DNA from 3665 individuals extracted from buccal cheek swabs was hybridized to AffymetrixGeneChip 6.0 SNP genotyping arrays using standard WTCCC2 experimental protocols. 4649 other individuals of European ancestry were genotyped on HumanOmniExpressExome-8v1.2 arrays at the Molecular Genetics Laboratories of the Medical Research Council Social, Genetic Developmental Psychiatry Centre, based on DNA extracted from saliva samples.

After initial quality control and genotype calling, the same quality control was conducted on the samples genotyped on the Affymetrix and Illumina platforms separately using PLINK (Purcell et al. 2007; Chang et al. 2015), R (GBIF.ORG n.d.), and vcftools (Danecek et al. 2011). Samples were removed based on call rate (<0.99), suspected non-European ancestry, heterozygosity, array signal intensity, and relatedness. SNPs were excluded if minor allele frequency was $<0.5\%$, if more than 1% of genotype data were missing, or if the Hardy Weinberg p-value was lower than 10^{-5} . Non-autosomal markers and indels were removed. Associations between the SNP and the platform, batch, or plate on which samples were genotyped were calculated; SNPs with p-values less than 10^{-3} were excluded. A total sample of 6710 individuals remained after quality control – 3093 individuals and 525 859 SNPs genotyped on the Affymetrix platform and 3617 individuals and 600 034 SNPs genotyped on the Illumina platform.

Genotypes from the two platforms were separately imputed using the Haplotype Reference Consortium 5 and Minimac3 1.0.13 (Howie et al. 2012; Fuchsberger et al. 2015) before merging genotype data obtained from both arrays.

We performed principal component analysis on a subset of 42 859 common ($MAF > 5\%$) autosomal HapMap3 SNPs (International HapMap 3 Consortium et al. 2010), after stringent pruning to remove markers in linkage disequilibrium ($r^2 > 0.1$) and excluding high linkage disequilibrium genomic regions so as to ensure that only genome-wide effects were detected.

SNP heritability: supplementary methods

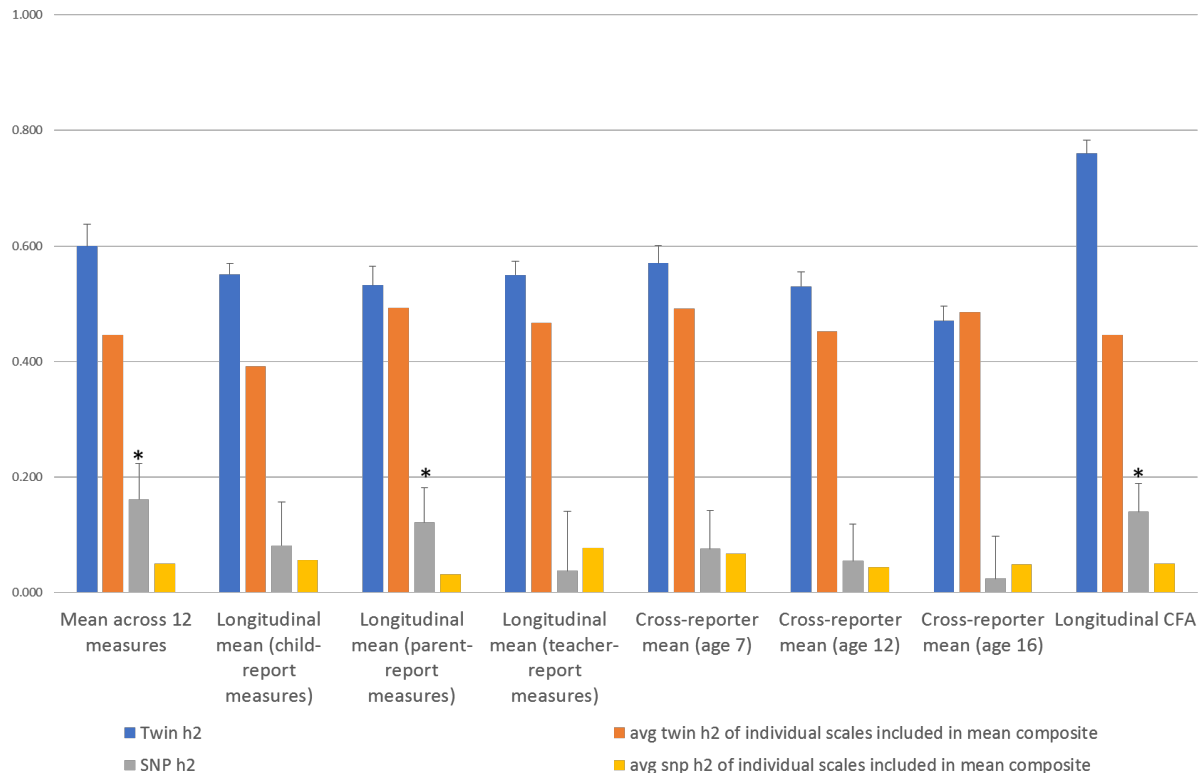
SNP heritabilities were estimated using genomic relatedness matrix restricted maximum likelihood (GREML), implemented in the Genome-wide Complex Trait Analysis (GCTA) program (Yang et al. 2011). This estimates genetic influence directly using genome-wide genotypes in large samples of unrelated individuals. First, genetic similarity for each pair of unrelated individuals across all genotyped or imputed genetic markers is calculated. Each pair's genetic similarity is then used to predict their phenotypic similarity. In the present study, one from each pair of individuals with pairwise identity-by-descent (IBD) of >0.025 (third degree relatives) were removed, so that chance genetic similarity could be used as a random effect in a linear mixed model. Comparing a matrix of pairwise genomic similarity to a matrix of pairwise phenotypic similarity using a random-effects mixed linear model, the variance of a trait can be decomposed into genetic and residual components.

Supplementary Table 7: Full twin and SNP heritability results:

Measure	Twin h ²	SNP h ²	Twin SE	Twin 95% CI	SNP SE	SNP 95% CI	Twin N (pairs)	SNP N (indiv)	p (SNP h ²)
Stable CFA score (simultaneous)	0.701	n/a	0.047	[0.608- 0.792]	n/a	n/a	6110	n/a	n/a
Stable CFA score (extracted FS)	0.760	0.14	0.023	[0.715- 0.805]	0.049	[0.044- 0.236]	6110	6002	0.002*
Stable IRT score (simultaneous)	0.611	n/a	0.043	[0.527- 0.695]	n/a	n/a	6110	n/a	n/a
Stable IRT score (extracted FS)	0.620	0.095	0.023	[0.575- 0.665]	0.048	[0.001- 0.189]	6110	6001	0.025*
Parent-rated SDQ anxiety at 7	0.475	0.009	0.037	[0.402- 0.548]	0.053	[-0.095- 0.113]	5359	5272	0.438
Teacher-rated SDQ anxiety at 7	0.509	0.126	0.018	[0.474- 0.544]	0.065	[-0.001- 0.253]	4419	4336	0.023*
Self-rated MFQ at 12	0.283	0.039	0.051	[0.183- 0.383]	0.064	[-0.086- 0.164]	4490	4414	0.27
Self-rated SDQ anxiety at 12	0.399	0.082	0.026	[0.348- 0.450]	0.065	[-0.045- 0.209]	4475	4396	0.106
Parent-rated SDQ anxiety at 12	0.531	0.022	0.041	[0.451- 0.611]	0.062	[-0.100- 0.144]	4482	4408	0.364
Parent-rated MFQ at 12	0.478	0.029	0.038	[0.404- 0.552]	0.063	[-0.094- 0.152]	4482	4407	0.326
Teacher-rated SDQ anxiety at 12	0.426	0.029	0.023	[0.381- 0.471]	0.077	[-0.122- 0.180]	3695	3652	0.359
Self-rated SDQ anxiety at 16	0.428	0.094	0.018	[0.393- 0.463]	0.075	[-0.053- 0.241]	3822	3759	0.108
Self-rated CASI at 16	0.445	0.069	0.019	[0.408- 0.482]	0.076	[-0.080- 0.218]	3824	3760	0.187
Self-rated MFQ at 16	0.403	0	0.024	[0.356- 0.450]	0.074	[-0.145- 0.145]	3824	3761	0.5
Parent-rated MFQ at 16	0.567	0.099	0.031	[0.506- 0.628]	0.073	[-0.044- 0.242]	3838	3774	0.08
Parent-rated ARBQ at age 16	0.413	0	0.039	[0.337- 0.489]	0.066	[-0.129- 0.129]	3840	3775	0.499

Note: 'h²' = heritability; 'SE' = standard error; '95% CI' = 95% confidence intervals; 'p' = the p-value associated with the SNP heritability estimate, and whether p<0.05 (indicated with '*'); 'FS' = stable emotional problems factor scores extracted from latent modelling; 'simultaneous' = simultaneous latent trait-twin modelling.

Supplementary Figure 4: Twin and SNP heritability results for a simple mean approach to aggregating emotional problems measures, with sensitivity analyses to investigate rater effects



Note: means were calculated only for individuals with data present on at least half of the variables. Twin and SNP heritability results for the means (blue and grey) are compared to average heritabilities of the constituent individual measures (orange and yellow). “*” indicates a significant SNP heritability estimate.

Supplementary Figure 4 shows that the simpler approach of creating a mean of the 12 anxiety and depression measures yields similar results to a latent modelling approach (comparing the first and eighth sections of the plot). As such, the finding of higher twin and SNP heritabilities holds for even a crude composite measure that adds up common *and* measure-specific error (rather than latent modelling to remove error). This bolsters the argument that aggregating across time and across raters taps into a more heritable, core ‘trait’ aspect of emotional development.

To explore the whether the effect of extracting longitudinal stability increasing heritability remained when using single-rater composites, we created longitudinal means for child-, parent- and teacher-reports separately (sections 2-4 of the plot). Results suggest that this is the case. Twin estimates increased from ~45% to ~55%. SNP heritability point estimates increased for child- and parent-report longitudinal means (with longitudinal aggregation boosting SNP heritability of parent-report measures into statistical significance), but not for the teacher-report composite. Given the similar point estimates and large standard error intervals around SNP heritability estimates, it would be inappropriate to compare SNP heritability estimates.

To investigate the contribution of trans-situational variance (i.e. rater agreement on childhood emotional problems), we created means of anxiety and depression measures across reporters but at individual ages (sections 5-7 of the plot). Results suggest that combining cross-sectional measures across reporters increases twin and SNP heritability estimates at ages 7 and 12, but not at 16. Again, standard errors for SNP heritability estimates are large, such that we cannot compare estimates for different composite measures, and we cannot say that aggregation across raters *significantly* increases SNP heritability. The heritability increase resulting from aggregating measures using a simple mean approach is weaker for SNP heritability than for twin estimates, but this is partly because the sample sizes for SNP heritability estimation were ~4000 (rather than over 6000 for latent modelling with maximum likelihood). The restriction of the mean composite approach to individuals with complete data is a disadvantage in comparison to the latent modelling approach. The finding that latent modelling is more powerful than simple approaches replicates evidence from the Netherlands Twin Register: a latent anxious depression phenotype was substantially more heritable than the scores

observed at any age, but measure created by simply summing data from the different ages was considerably less heritable (Lubke et al. 2016).

Overall, results suggest that both the mean composite and latent factor across time and across reporters shows increased heritability as a result of tapping into longitudinally stable aspects of behaviour that are also agreed upon by different raters. Future research could more thoroughly and explicitly test the structure of age-, scale-, and rater-specific influences on emotional problems. For example, in the Trait-State-Occasion model, significant covariances between observed variable error terms indicates that trait stability is inflated by shared method variance cross waves.

Investigation of the contribution of individuals with persistent emotional problems to the increased heritability of our stable phenotype.

In studies of adults, recurrent depression is more severe and heritable than single episode depression. It follows that the heritability for stable childhood emotional problems could be higher because we are capturing the higher heritability of severe recurrent symptoms, rather than stability at any level of emotional symptoms. This is especially worth investigating given the high genetic correlation of our stable childhood emotional problems phenotype with adult depression.

To test this idea, we calculated how many of the 12 measures each individual was an outlier for. Outliers were defined as individuals with z-scores ≥ 1.96 (i.e. cutting off the top 1% of scores). We removed twin pairs for whom twin 1 or 2 was an outlier for >6 of the 12 scales (i.e. removed 21 pairs). Then, we re-calculated the twin heritability of a single common factor (from CFA).

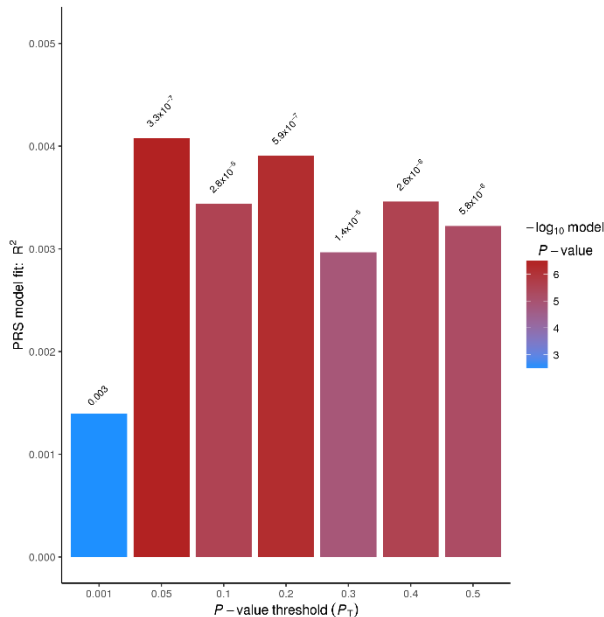
We found a lower twin heritability for stable emotional problems when excluding individuals with persistent problems (0.50 (se=0.05) vs 0.70 (se=0.05) for the CFA factor scores using all 6110 pairs). This suggests that a small number of young people with enduring severe emotional symptoms could be disproportionately contributing to the high heritability of stable emotional problems. However, 95% CIs for the two estimates just overlap (.40-.60 and .60-.80), so the difference is not statistically significant.

Adjusting the definition of outliers to $z \geq 1.64$ (the top 5% of scores) led to the removal of 43 twin pairs for whom at least one twin was an outlier on more than half of the measures. The heritability estimate hardly decreased below 0.50 (0.49 (se=0.05)), but the upper 95% CI decreased to .57, such that the CIs no longer overlapped with the estimates from the full sample. Overall, there is no strong evidence that individuals with severe persistent symptoms are inflating heritability estimates.

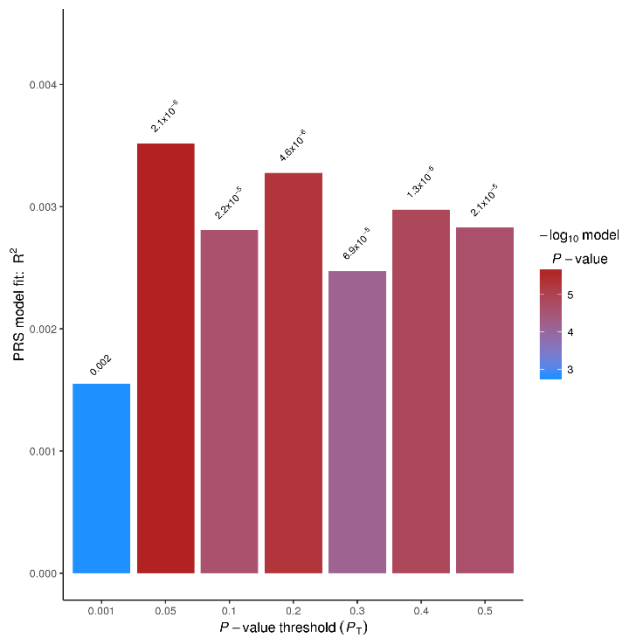
Supplementary Figures 5 and 6: Polygenic score results

The PRSice 2 plots below demonstrate that the R^2 is not dramatically better at the best p-value threshold than at the others. This is indicative of reliable prediction for a polygenic trait.

Supplementary Figure 5: Prediction of CFA scores using UK Biobank anxiety polygenic score



Supplementary Figure 6: Prediction of IRT scores using UK Biobank anxiety polygenic score



Supplementary Table 8: variance explained by polygenic scores for MDD, derived from summary statistics for a subset of 10,000 variants in the PGC sample including 23&Me.

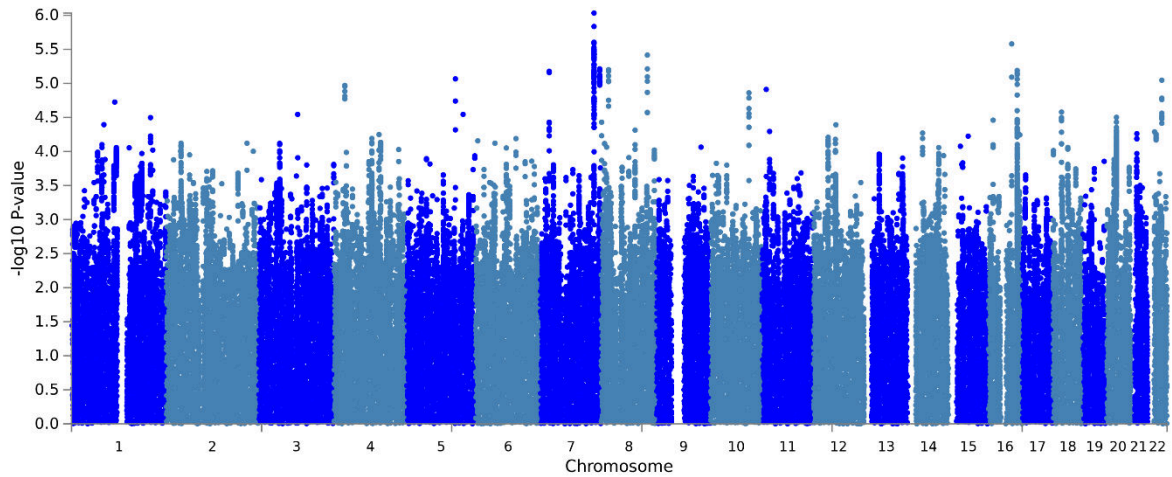
	R ²	P-value	Empirical P-value
Stable CFA score	0.0041	0.0000	0.0001*
Stable IRT score	0.0031	0.0000	0.0002*
Parent-rated SDQ at 7	0.0006	0.0754	0.1268
Teacher-rated SDQ at 7	0.0023	0.0015	0.0026*
Self-rated MFQ at 12	0.0051	0.0000	0.0001*
Self-rated SDQ at 12	0.0024	0.0010	0.0020*
Parent-rated SDQ at 12	0.0009	0.0411	0.0743
Parent-rated MFQ at 12	0.0014	0.0113	0.0213*
Teacher-rated SDQ at 12	0.0005	0.1633	0.2582
Self-rated SDQ at 16	0.0022	0.0021	0.0038*
Self-rated CASI at 16	0.0034	0.0001	0.0005*
Self-rated MFQ at 16	0.0020	0.0042	0.0091*
Parent-rated MFQ at 16	0.0014	0.0202	0.0366*
Parent-rated ARBQ at 16	0.0021	0.0036	0.0067*

Note: '*' indicates target TEDS phenotypes that were significantly predicted by polygenic scores at the more stringent empirical p-value thresholds.

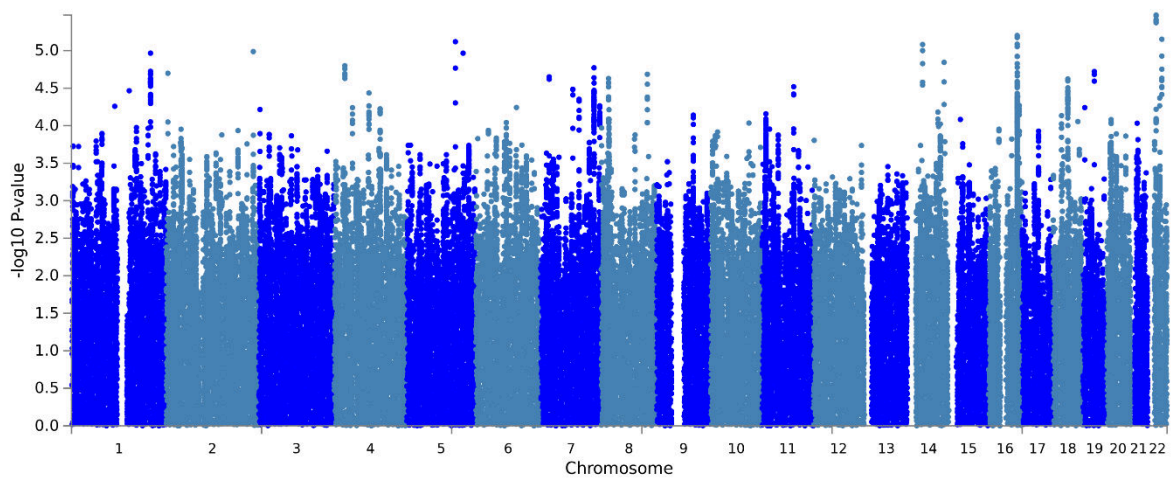
Supplementary Figures 7 and 8: Genome-wide association results

No significant SNPs were identified. There are no SNPs that exceed a $-\log_{10}$ p-value of more than 6. This was expected, given the relatively small sample (6110), and the lack of associations in previous, similarly powered studies of emotional problems (e.g. Benke et al., 2014 meta-analysis of preschool emotional problems, N=4596).

IRT score GWAS: FUMA plot:



CFA score GWAS: FUMA plot:



Supplementary Table 9: Genetic correlations between stable emotional problems (from CFA) and psychiatric traits, obtained using LD Score regression.

Trait	PMID	rg	se	z	p	h2_obs	h2_obs_se
Depressive symptoms	27089181	0.64	0.19	3.37	0.0008**	0.05	0
Adult depression (PGC) †		0.48	0.14	3.35	0.0008**	0	0
Adult anxiety (UK Biobank) †		0.45	0.15	2.92	0.0035**	0.11	0.01
Subjective well being	27089181	-0.36	0.15	-2.36	0.0182*	0.02	0
Major depressive disorder	22472876	0.37	0.2	1.9	0.0574	0.18	0.03
Schizophrenia	25056061	0.16	0.09	1.8	0.0718	0.45	0.02
PGC cross-disorder analysis	23453885	0.21	0.14	1.49	0.1368	0.18	0.01
Attention deficit hyperactivity disorder (GC)	27663945	0.34	0.3	1.14	0.2528	0.08	0.03
Attention deficit hyperactivity disorder (No GC)	27663945	0.34	0.3	1.14	0.2538	0.08	0.03
Attention deficit hyperactivity disorder	20732625	-0.27	0.3	-0.9	0.3677	0.22	0.1
Autism spectrum disorder	0	-0.06	0.15	-0.39	0.6972	0.46	0.05
Bipolar disorder	21926972	0.02	0.14	0.14	0.8857	0.45	0.04
Anorexia Nervosa	24514567	-0.01	0.13	-0.06	0.9528	0.4	0.04

Note: Pairs of traits are ordered according to the p-value of their genetic correlation. * = significant at $p < 0.05$, ** = significant at $p < 0.00385$ i.e. $(0.05/13)$ to correct for multiple tests; † = genetic correlation analysis performed separately (using LDSR) because phenotype was *not* in LD Hub.

The genetic correlations between our CFA-derived stable scores, and case-control depression, depressive symptoms, case-control generalised anxiety, and wellbeing, are significant at $p < 0.05$. The negative genetic correlation with wellbeing did not remain significant at the more stringent threshold ($p < 0.00385$). Genetic correlations have many possible interpretations other than true biological pleiotropy (Martin et al. 2017). For example, the relationship between adult depression and childhood stable emotional problems could be mediated by shared genetic risk with low levels of well-being. It has been suggested that the genetic correlation between MDD and depressive symptoms in the population could be accounted for by shared genetic risk with low levels of subjective well-being (Direk et al. 2017). Another issue is that nosological issues involving heterogeneity, comorbidity and misclassification might inflate phenotypic and genetic correlations.

Full authors list: Major Depressive Disorder Working Group of the Psychiatric Genomics Consortium

Naomi R Wray* 1, 2	Nese Direk 37, 38	Jesper Krogh 67
Stephan Ripke* 3, 4, 5	Conor V Dolan 10	Zoltán Kutalik 68, 69
Manuel Mattheisen* 6, 7, 8, 9	Erin C Dunn 39, 40, 41	Yihan Li 66
Maciej Trzaskowski* 1	Thalia C Eley 27	Penelope A Lind 28
Enda M Byrne 1	Valentina Escott-Price 42	Donald J MacIntyre 70, 71
Abdel Abdellaoui 10	Farnush Farhadi	Dean F MacKinnon 50
Mark J Adams 11	Hassan Kiadeh 43	Robert M Maier 2
Esben Agerbo 9, 12, 13	Hilary K Finucane 44, 45	Wolfgang Maier 72
Tracy M Air 14	Andreas J Forstner 35, 36, 46, 47	Jonathan Marchini 73
Till F M Andlauer 15, 16	Josef Frank 48	Hamdi Mbarek 10
Silviu-Alin Bacanu 17	Hélène A Gaspar 27	Patrick McGrath 74
Marie Bækvad-Hansen 9, 18	Michael Gill 49	Peter McGuffin 27
Aartjan T F Beekman 19	Fernando S Goes 50	Sarah E Medland 28
Tim B Bigdeli 17, 20	Scott D Gordon 51	Divya Mehta 2, 75
Elisabeth B Binder 15, 21	Jakob Grove 7, 8, 9, 52	Christel M Middeldorp 10, 76, 77
Douglas H R Blackwood 11	Lynsey S Hall 11, 53	Evelin Mihailov 78
Julien Bryois 22	Christine Søholm Hansen 9, 18	Yuri Milaneschi 19, 19
Henriette N Buttenschøn 8, 9, 23	Thomas F Hansen 54, 55, 56	Lili Milani 78
Jonas Bybjerg-Grauholm 9, 18	Stefan Herms 35, 36, 47	Francis M Mondimore 50
Na Cai 24, 25	Ian B Hickie 57	Grant W Montgomery 1
Enrique Castelao 26	Per Hoffmann 35, 36, 47	Sara Mostafavi 79, 80
Jane Hvarregaard Christensen 7, 8, 9	Georg Homuth 58	Niamh Mullins 27
Toni-Kim Clarke 11	Carsten Horn 59	Matthias Nauck 81, 82
Jonathan R I Coleman 27	Jouke-Jan Hottenga 10	Bernard Ng 80
Lucía Colodro-Conde 28	David M Hougaard 9, 18	Michel G Nivard 10
Baptiste Couvy-Duchesne 29, 30	Marcus Ising 60	Dale R Nyholt 83
Nick Craddock 31	Rick Jansen 19, 19	Paul F O'Reilly 27
Gregory E Crawford 32, 33	Eric Jorgenson 61	Hogni Oskarsson 84
Gail Davies 34	James A Knowles 62	Michael J Owen 85
Ian J Deary 34	Isaac S Kohane 63, 64, 65	Jodie N Painter 28
Franziska Degenhardt 35, 36	Julia Kraft 4	Carsten Bøcker Pedersen 9, 12, 13
Eske M Derks 28	Warren W. Kretzschmar 66	Marianne Giørtz Pedersen 9, 12, 13
		Roseann E. Peterson 17, 86
		Erik Pettersson 22
		Wouter J Peyrot 19
		Giorgio Pistis 26
		Danielle Posthuma 87, 88
		Jorge A Quiroz 89
		Per Qvist 7, 8, 9
		John P Rice 90
		Brien P. Riley 17
		Margarita Rivera 27, 91
		Saira Saeed Mirza 37
		Robert Schoevers 92
		Eva C Schulte 93, 94
		Ling Shen 61
		Jianxin Shi 95
		Stanley I Shyn 96
		Engilbert Sigurdsson 97
		Grant C B Sinnamon 98
		Johannes H Smit 19
		Daniel J Smith 99

Hreinn Stefansson
 100
 Stacy Steinberg
 100
 Fabian Streit 48
 Jana Strohmaier
 48
 Katherine E
 Tansey 101
 Henning
 Teismann 102
 Alexander Teumer
 103
 Wesley Thompson
 9, 55, 104, 105
 Pippa A Thomson
 106
 Thorgeir E
 Thorgeirsson 100
 Matthew Traylor
 107
 Jens Treutlein 48
 Vassily
 Trubetskoy 4
 André G
 Uitterlinden 108
 Daniel Umbricht
 109
 Sandra Van der
 Auwera 110
 Albert M van
 Hemert 111
 Alexander Viktorin
 22
 Peter M Visscher
 1, 2
 Yunpeng Wang 9,
 55, 105
 Bradley T. Webb
 112
 Shantel Marie
 Weinsheimer 9, 55
 Jürgen Wellmann
 102
 Gonneke
 Willemsen 10
 Stephanie H Witt
 48
 Yang Wu 1
 Hualin S Xi 113
 Jian Yang 2, 114
 Futao Zhang 1
 Volker Arolt 115
 Bernhard T Baune
 14
 Klaus Berger 102
 Dorret I Boomsma
 10
 Sven Cichon 35,
 47, 116, 117

Udo Dannlowski
 115
 EJC de Geus 10,
 118
 J Raymond
 DePaulo 50
 Enrico Domenici
 119
 Katharina
 Domschke 120
 Tõnu Esko 5, 78
 Hans J Grabe
 110
 Steven P
 Hamilton 121
 Caroline
 Hayward 122
 Andrew C Heath
 90
 Kenneth S
 Kendler 17
 Stefan Kloiber
 60, 123, 124
 Glyn Lewis 125
 Qingqin S Li 126
 Susanne Lucae
 60
 Pamela AF
 Madden 90
 Patrik K
 Magnusson 22
 Nicholas G
 Martin 51
 Andrew M
 McIntosh 11, 34
 Andres Metspalu
 78, 127
 Ole Mors 9, 128
 Preben Bo
 Mortensen 8, 9,
 12, 13
 Bertram Müller-
 Myhsok 15, 16,
 129
 Merete
 Nordentoft 9, 130
 Markus M Nöthen
 35, 36
 Michael C
 O'Donovan 85
 Sara A Paciga
 131
 Nancy L
 Pedersen 22
 Brenda WJH
 Penninx 19
 Roy H Perlis 39,
 132
 David J Porteous
 106

James B Potash 133
 Martin Preisig 26
 Marcella Rietschel 48
 Catherine Schaefer 61
 Thomas G Schulze 48,
 94, 134, 135, 136
 Jordan W Smoller 39,
 40, 41
 Kari Stefansson 100,
 137
 Henning Tiemeier 37,
 138, 139
 Rudolf Uher 140
 Henry Völzke 103
 Myrna M Weissman 74,
 141
 Thomas Werge 9, 55,
 142
 Cathryn M Lewis* 27,
 143
 Douglas F Levinson* 144
 Gerome Breen* 27, 145
 Anders D Børglum* 7, 8,
 9
 Patrick F Sullivan* 22,
 146, 147,

- 1, Institute for Molecular Bioscience, The University of Queensland, Brisbane, QLD, AU
- 2, Queensland Brain Institute, The University of Queensland, Brisbane, QLD, AU
- 3, Analytic and Translational Genetics Unit, Massachusetts General Hospital, Boston, MA, US
- 4, Department of Psychiatry and Psychotherapy, Universitätsmedizin Berlin Campus Charité Mitte, Berlin, DE
- 5, Medical and Population Genetics, Broad Institute, Cambridge, MA, US
- 6, Centre for Psychiatry Research, Department of Clinical Neuroscience, Karolinska Institutet, Stockholm, SE
- 7, Department of Biomedicine, Aarhus University, Aarhus, DK
- 8, iSEQ, Centre for Integrative Sequencing, Aarhus University, Aarhus, DK
- 9, iPSYCH, The Lundbeck Foundation Initiative for Integrative Psychiatric Research, DK
- 10, Dept of Biological Psychology & EMGO+ Institute for Health and Care Research, Vrije Universiteit Amsterdam, Amsterdam, NL
- 11, Division of Psychiatry, University of Edinburgh, Edinburgh, GB
- 12, Centre for Integrated Register-based Research, Aarhus University, Aarhus, DK
- 13, National Centre for Register-Based Research, Aarhus University, Aarhus, DK
- 14, Discipline of Psychiatry, University of Adelaide, Adelaide, SA, AU
- 15, Department of Translational Research in Psychiatry, Max Planck Institute of Psychiatry, Munich, DE
- 16, Munich Cluster for Systems Neurology (SyNergy), Munich, DE
- 17, Department of Psychiatry, Virginia Commonwealth University, Richmond, VA, US
- 18, Center for Neonatal Screening, Department for Congenital Disorders, Statens Serum Institut, Copenhagen, DK
- 19, Department of Psychiatry, Vrije Universiteit Medical Center and GGZ inGeest, Amsterdam, NL
- 20, Virginia Institute for Psychiatric and Behavior Genetics, Richmond, VA, US
- 21, Department of Psychiatry and Behavioral Sciences, Emory University School of Medicine, Atlanta, GA, US
- 22, Department of Medical Epidemiology and Biostatistics, Karolinska Institutet, Stockholm, SE
- 23, Department of Clinical Medicine, Translational Neuropsychiatry Unit, Aarhus University, Aarhus, DK
- 24, Human Genetics, Wellcome Trust Sanger Institute, Cambridge, GB
- 25, Statistical genomics and systems genetics, European Bioinformatics Institute (EMBL-EBI), Cambridge, GB
- 26, Department of Psychiatry, University Hospital of Lausanne, Prilly, Vaud, CH
- 27, MRC Social Genetic and Developmental Psychiatry Centre, King's College London, London, GB
- 28, Genetics and Computational Biology, QIMR Berghofer Medical Research Institute, Herston, QLD, AU
- 29, Centre for Advanced Imaging, The University of Queensland, Saint Lucia, QLD, AU
- 30, Queensland Brain Institute, The University of Queensland, Saint Lucia, QLD, AU
- 31, Psychological Medicine, Cardiff University, Cardiff, GB
- 32, Center for Genomic and Computational Biology, Duke University, Durham, NC, US
- 33, Department of Pediatrics, Division of Medical Genetics, Duke University, Durham, NC, US
- 34, Centre for Cognitive Ageing and Cognitive Epidemiology, University of Edinburgh, Edinburgh, GB
- 35, Institute of Human Genetics, University of Bonn, Bonn, DE
- 36, Life&Brain Center, Department of Genomics, University of Bonn, Bonn, DE
- 37, Epidemiology, Erasmus MC, Rotterdam, Zuid-Holland, NL
- 38, Psychiatry, Dokuz Eylül University School Of Medicine, Izmir, TR
- 39, Department of Psychiatry, Massachusetts General Hospital, Boston, MA, US
- 40, Psychiatric and Neurodevelopmental Genetics Unit (PNGU), Massachusetts General Hospital, Boston, MA, US
- 41, Stanley Center for Psychiatric Research, Broad Institute, Cambridge, MA, US
- 42, Neuroscience and Mental Health, Cardiff University, Cardiff, GB
- 43, Bioinformatics, University of British Columbia, Vancouver, BC, CA
- 44, Department of Epidemiology, Harvard T.H. Chan School of Public Health, Boston, MA, US
- 45, Department of Mathematics, Massachusetts Institute of Technology, Cambridge, MA, US
- 46, Department of Psychiatry (UPK), University of Basel, Basel, CH
- 47, Human Genomics Research Group, Department of Biomedicine, University of Basel, Basel, CH
- 48, Department of Genetic Epidemiology in Psychiatry, Central Institute of Mental Health, Medical Faculty Mannheim, Heidelberg University, Mannheim, Baden-Württemberg, DE
- 49, Department of Psychiatry, Trinity College Dublin, Dublin, IE

50, Psychiatry & Behavioral Sciences, Johns Hopkins University, Baltimore, MD, US

51, Genetics and Computational Biology, QIMR Berghofer Medical Research Institute, Brisbane, QLD, AU

52, Bioinformatics Research Centre, Aarhus University, Aarhus, DK

53, Institute of Genetic Medicine, Newcastle University, Newcastle upon Tyne, GB

54, Danish Headache Centre, Department of Neurology, Rigshospitalet, Glostrup, DK

55, Institute of Biological Psychiatry, Mental Health Center Sct. Hans, Mental Health Services Capital Region of Denmark, Copenhagen, DK

56, iPSYCH, The Lundbeck Foundation Initiative for Psychiatric Research, Copenhagen, DK

57, Brain and Mind Centre, University of Sydney, Sydney, NSW, AU

58, Interfaculty Institute for Genetics and Functional Genomics, Department of Functional Genomics, University Medicine and Ernst Moritz Arndt University Greifswald, Greifswald, Mecklenburg-Vorpommern, DE

59, Roche Pharmaceutical Research and Early Development, Pharmaceutical Sciences, Roche Innovation Center Basel, F. Hoffmann-La Roche Ltd, Basel, CH

60, Max Planck Institute of Psychiatry, Munich, DE

61, Division of Research, Kaiser Permanente Northern California, Oakland, CA, US

62, Psychiatry & The Behavioral Sciences, University of Southern California, Los Angeles, CA, US

63, Department of Biomedical Informatics, Harvard Medical School, Boston, MA, US

64, Department of Medicine, Brigham and Women's Hospital, Boston, MA, US

65, Informatics Program, Boston Children's Hospital, Boston, MA, US

66, Wellcome Trust Centre for Human Genetics, University of Oxford, Oxford, GB

67, Department of Endocrinology at Herlev University Hospital, University of Copenhagen, Copenhagen, DK

68, Institute of Social and Preventive Medicine (IUMSP), University Hospital of Lausanne, Lausanne, VD, CH

69, Swiss Institute of Bioinformatics, Lausanne, VD, CH

70, Division of Psychiatry, Centre for Clinical Brain Sciences, University of Edinburgh, Edinburgh, GB

71, Mental Health, NHS 24, Glasgow, GB

72, Department of Psychiatry and Psychotherapy, University of Bonn, Bonn, DE

73, Statistics, University of Oxford, Oxford, GB

74, Psychiatry, Columbia University College of Physicians and Surgeons, New York, NY, US

75, School of Psychology and Counseling, Queensland University of Technology, Brisbane, QLD, AU

76, Child and Youth Mental Health Service, Children's Health Queensland Hospital and Health Service, South Brisbane, QLD, AU

77, Child Health Research Centre, University of Queensland, Brisbane, QLD, AU

78, Estonian Genome Center, University of Tartu, Tartu, EE

79, Medical Genetics, University of British Columbia, Vancouver, BC, CA

80, Statistics, University of British Columbia, Vancouver, BC, CA

81, DZHK (German Centre for Cardiovascular Research), Partner Site Greifswald, University Medicine, University Medicine Greifswald, Greifswald, Mecklenburg-Vorpommern, DE

82, Institute of Clinical Chemistry and Laboratory Medicine, University Medicine Greifswald, Greifswald, Mecklenburg-Vorpommern, DE

83, Institute of Health and Biomedical Innovation, Queensland University of Technology, Brisbane, QLD, AU

84, Humus, Reykjavik, IS

85, MRC Centre for Neuropsychiatric Genetics and Genomics, Cardiff University, Cardiff, GB

86, Virginia Institute for Psychiatric & Behavioral Genetics, Virginia Commonwealth University, Richmond, VA, US

87, Clinical Genetics, Vrije Universiteit Medical Center, Amsterdam, NL

88, Complex Trait Genetics, Vrije Universiteit Amsterdam, Amsterdam, NL

89, Solid Biosciences, Boston, MA, US

90, Department of Psychiatry, Washington University in Saint Louis School of Medicine, Saint Louis, MO, US

91, Department of Biochemistry and Molecular Biology II, Institute of Neurosciences, Center for Biomedical Research, University of Granada, Granada, ES

92, Department of Psychiatry, University of Groningen, University Medical Center Groningen, Groningen, NL

93, Department of Psychiatry and Psychotherapy, Medical Center of the University of Munich, Campus Innenstadt, Munich, DE

- 94, Institute of Psychiatric Phenomics and Genomics (IPPG), Medical Center of the University of Munich, Campus Innenstadt, Munich, DE
- 95, Division of Cancer Epidemiology and Genetics, National Cancer Institute, Bethesda, MD, US
- 96, Behavioral Health Services, Kaiser Permanente Washington, Seattle, WA, US
- 97, Faculty of Medicine, Department of Psychiatry, University of Iceland, Reykjavik, IS
- 98, School of Medicine and Dentistry, James Cook University, Townsville, QLD, AU
- 99, Institute of Health and Wellbeing, University of Glasgow, Glasgow, GB
- 100, deCODE Genetics / Amgen, Reykjavik, IS
- 101, College of Biomedical and Life Sciences, Cardiff University, Cardiff, GB
- 102, Institute of Epidemiology and Social Medicine, University of Münster, Münster, Nordrhein-Westfalen, DE
- 103, Institute for Community Medicine, University Medicine Greifswald, Greifswald, Mecklenburg-Vorpommern, DE
- 104, Department of Psychiatry, University of California, San Diego, San Diego, CA, US
- 105, KG Jebsen Centre for Psychosis Research, Norway Division of Mental Health and Addiction, Oslo University Hospital, Oslo, NO
- 106, Medical Genetics Section, CGEM, IGMM, University of Edinburgh, Edinburgh, GB
- 107, Clinical Neurosciences, University of Cambridge, Cambridge, GB
- 108, Internal Medicine, Erasmus MC, Rotterdam, Zuid-Holland, NL
- 109, Roche Pharmaceutical Research and Early Development, Neuroscience, Ophthalmology and Rare Diseases Discovery & Translational Medicine Area, Roche Innovation Center Basel, F. Hoffmann-La Roche Ltd, Basel, CH
- 110, Department of Psychiatry and Psychotherapy, University Medicine Greifswald, Greifswald, Mecklenburg-Vorpommern, DE
- 111, Department of Psychiatry, Leiden University Medical Center, Leiden, NL
- 112, Virginia Institute of Psychiatric & Behavioral Genetics, Virginia Commonwealth University, Richmond, VA, US
- 113, Computational Sciences Center of Emphasis, Pfizer Global Research and Development, Cambridge, MA, US
- 114, Institute for Molecular Bioscience; Queensland Brain Institute, The University of Queensland, Brisbane, QLD, AU
- 115, Department of Psychiatry, University of Münster, Münster, Nordrhein-Westfalen, DE
- 116, Institute of Medical Genetics and Pathology, University Hospital Basel, University of Basel, Basel, CH
- 117, Institute of Neuroscience and Medicine (INM-1), Research Center Juelich, Juelich, DE
- 118, Amsterdam Public Health Institute, Vrije Universiteit Medical Center, Amsterdam, NL
- 119, Centre for Integrative Biology, Università degli Studi di Trento, Trento, Trentino-Alto Adige, IT
- 120, Department of Psychiatry and Psychotherapy, Medical Center, University of Freiburg, Faculty of Medicine, University of Freiburg, Freiburg, DE
- 121, Psychiatry, Kaiser Permanente Northern California, San Francisco, CA, US
- 122, Medical Research Council Human Genetics Unit, Institute of Genetics and Molecular Medicine, University of Edinburgh, Edinburgh, GB
- 123, Department of Psychiatry, University of Toronto, Toronto, ON, CA
- 124, Centre for Addiction and Mental Health, Toronto, ON, CA
- 125, Division of Psychiatry, University College London, London, GB
- 126, Neuroscience Therapeutic Area, Janssen Research and Development, LLC, Titusville, NJ, US
- 127, Institute of Molecular and Cell Biology, University of Tartu, Tartu, EE
- 128, Psychosis Research Unit, Aarhus University Hospital, Risskov, Aarhus, DK
- 129, University of Liverpool, Liverpool, GB
- 130, Mental Health Center Copenhagen, Copenhagen University Hospital, Copenhagen, DK
- 131, Human Genetics and Computational Biomedicine, Pfizer Global Research and Development, Groton, CT, US
- 132, Psychiatry, Harvard Medical School, Boston, MA, US
- 133, Psychiatry, University of Iowa, Iowa City, IA, US
- 134, Department of Psychiatry and Behavioral Sciences, Johns Hopkins University, Baltimore, MD, US
- 135, Department of Psychiatry and Psychotherapy, University Medical Center Göttingen, Goettingen, Niedersachsen, DE
- 136, Human Genetics Branch, NIMH Division of Intramural Research Programs, Bethesda, MD, US
- 137, Faculty of Medicine, University of Iceland, Reykjavik, IS

- 138, Child and Adolescent Psychiatry, Erasmus MC, Rotterdam, Zuid-Holland, NL
- 139, Psychiatry, Erasmus MC, Rotterdam, Zuid-Holland, NL
- 140, Psychiatry, Dalhousie University, Halifax, NS, CA
- 141, Division of Epidemiology, New York State Psychiatric Institute, New York, NY, US
- 142, Department of Clinical Medicine, University of Copenhagen, Copenhagen, DK
- 143, Department of Medical & Molecular Genetics, King's College London, London, GB
- 144, Psychiatry & Behavioral Sciences, Stanford University, Stanford, CA, US
- 145, NIHR BRC for Mental Health, King's College London, London, GB
- 146, Genetics, University of North Carolina at Chapel Hill, Chapel Hill, NC, US
- 147, Psychiatry, University of North Carolina at Chapel Hill, Chapel Hill, NC, US

LD Hub acknowledgements

We gratefully acknowledge all the studies and databases that made GWAS summary data available:

ADIPOGen (Adiponectin genetics consortium), **C4D** (Coronary Artery Disease Genetics Consortium), **CARDIoGRAM** (Coronary ARtery Disease Genome wide Replication and Meta-analysis), **CKDGen** (Chronic Kidney Disease Genetics consortium), **dbGAP** (database of Genotypes and Phenotypes), **DIAGRAM** (DIABetes Genetics Replication And Meta-analysis), **ENIGMA** (Enhancing Neuro Imaging Genetics through Meta Analysis), **EAGLE** (EARly Genetics & Lifecourse Epidemiology Eczema Consortium, excluding 23andMe), **EGG** (Early Growth Genetics Consortium), **GABRIEL** (A Multidisciplinary Study to Identify the Genetic and Environmental Causes of Asthma in the European Community), **GCAN** (Genetic Consortium for Anorexia Nervosa), **GEFOS** (GEnetic Factors for OSteoporosis Consortium), **GIANT** (Genetic Investigation of ANthropometric Traits), **GIS** (Genetics of Iron Status consortium

), **GLGC** (Global Lipids Genetics Consortium), **GPC** (Genetics of Personality Consortium), **GUGC** (Global Urate and Gout consortium), **HaemGen** (haematological and platelet traits genetics consortium), **HRgene** (Heart Rate consortium), **IIBDGC** (International Inflammatory Bowel Disease Genetics Consortium), **ILCCO** (International Lung Cancer Consortium), **IMSGC** (International Multiple Sclerosis Genetic Consortium), **MAGIC** (Meta-Analyses of Glucose and Insulin-related traits Consortium), **MESA** (Multi-Ethnic Study of Atherosclerosis), **PGC** (Psychiatric Genomics Consortium), **Project MinE** consortium, **ReproGen** (Reproductive Genetics Consortium), **SSGAC** (Social Science Genetics Association Consortium) and **TAG** (Tobacco and Genetics Consortium), **TRICL** (Transdisciplinary Research in Cancer of the Lung consortium), **UK Biobank**.

We gratefully acknowledge the contributions of Alkes Price (the systemic lupus erythematosus GWAS and primary biliary cirrhosis GWAS) and Johannes Kettunen (lipids metabolites GWAS).

References

- Chang, C.C., Chow, C.C., Tellier, L.C., Vattikuti, S., Purcell, S.M. and Lee, J.J. 2015. Second-generation PLINK: rising to the challenge of larger and richer datasets. *GigaScience* 4, p. 7.
- Danecek, P., Auton, A., Abecasis, G., Albers, C.A., Banks, E., DePristo, M.A., Handsaker, R.E., Lunter, G., Marth, G.T., Sherry, S.T., McVean, G., Durbin, R. and 1000 Genomes Project Analysis Group 2011. The variant call format and VCFtools. *Bioinformatics* 27(15), pp. 2156–2158.
- Direk, N., Williams, S., Smith, J.A., Ripke, S., Sullivan, P.F., et al. 2017. An Analysis of Two Genome-wide Association Meta-analyses Identifies a New Locus for Broad Depression Phenotype. *Biological Psychiatry* 82(5), pp. 322–329.
- Fuchsberger, C., Abecasis, G.R. and Hinds, D.A. 2015. minimac2: faster genotype imputation. *Bioinformatics* 31(5), pp. 782–784.
- GBIF.ORG R: a language and environment for statistical computing [Online]. Available at: <http://www.gbif.org/resource/81287> [Accessed: 8 May 2017].
- Howie, B., Fuchsberger, C., Stephens, M., Marchini, J. and Abecasis, G.R. 2012. Fast and accurate genotype imputation in genome-wide association studies through pre-phasing. *Nature Genetics* 44(8), pp. 955–959.
- International HapMap 3 Consortium, Altshuler, D.M., Gibbs, R.A., Peltonen, L., McEwen, J.E., et al. 2010. Integrating common and rare genetic variation in diverse human populations. *Nature* 467(7311), pp. 52–58.
- Lubke, G.H., Miller, P.J., Verhulst, B., Bartels, M., van Beijsterveldt, T., Willemsen, G., Boomsma, D.I. and Middeldorp, C.M. 2016. A powerful phenotype for gene-finding studies derived from trajectory analyses of symptoms of anxiety and depression between age seven and 18. *American Journal of Medical Genetics. Part B, Neuropsychiatric Genetics* 171(7), pp. 948–957.
- Martin, J., Taylor, M.J. and Lichtenstein, P. 2017. Assessing the evidence for shared genetic risks across psychiatric disorders and traits. *Psychological Medicine*, pp. 1–16.
- Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M.A.R., Bender, D., Maller, J., Sklar, P., de Bakker, P.I.W., Daly, M.J. and Sham, P.C. 2007. PLINK: a tool set for whole-genome association and population-based linkage analyses. *American Journal of Human Genetics* 81(3), pp. 559–575.

Appendix 2. Supplementary Information for Chapter 3 (Familial Influences on Neuroticism and Education in the UK Biobank).

Supplementary Table 1: Full GREML-KIN model-fitting results for Neuroticism

Note: G (unrelated) refers to results from a standard GREML model using a relatedness cutoff of <0.025; selected model = GK (bold, underlined); GKFSC = full model; LL=log likelihood.

Model	h2g (se)	h2kin (se)	ef2 (se)	es2 (se)	ec2 (se)	LL	N
G (unrelated)	0.10 (0.01)					-73638.936	44686
GKFSC	0.11 (0.01)	0.20 (0.09)	0.01 (0.05)	0.00 (0.02)	0.04 (0.05)	-107829.455	65361
GKFC	0.11 (0.01)	0.20 (0.09)	0.02 (0.05)	0.04 (0.05)		-107829.458	65361
GKSC	0.11 (0.01)	0.17 (0.03)		0.00 (0.02)	0.03 (0.01)	-107829.524	65361
GKC	0.11 (0.01)	0.16 (0.02)			0.03 (0.01)	-107829.541	65361
GKFS	0.11 (0.01)	0.12 (0.03)	0.03 (0.01)	0.00 (0.02)		-107829.907	65361
GKF	0.11 (0.01)	0.11 (0.02)	0.03 (0.01)			-107829.966	65361
GFSC	0.11 (0.01)		0.08 (0.02)	0.00 (0.02)	(-)0.05 (0.02)	-107832.111	65361
GFC	0.11 (0.01)		0.08 (0.01)		(-)0.05 (0.01)	-107832.112	65361
GKS	0.11 (0.01)	0.17 (0.03)		0.00 (0.02)		-107835.138	65361
<u>GK</u>	<u>0.11 (0.01)</u>	<u>0.16 (0.02)</u>				<u>-107835.16</u>	<u>65361</u>
GFS	0.11 (0.01)		0.04 (0.01)	0.04 (0.01)		-107836.775	65361
GF	0.12 (0.01)		0.05 (0.01)			-107843.082	65361
GSC	0.12 (0.01)			0.08 (0.01)	0.03 (0.01)	-107845.177	65361
GS	0.12 (0.01)			0.08 (0.01)		-107850.821	65361
GC	0.15 (0.01)				0.03 (0.01)	-107880.81	65361
G	0.15 (0.008)					-107886.434	65361
KFSC		0.30 (0.09)	(-)0.02 (0.05)	0.00 (0.02)	0.04 (0.05)	-107909.95	65361
KFC		0.30 (0.09)	0.01 (0.05)		0.04 (0.05)	-107909.955	65361
KC		0.27 (0.01)			0.03 (0.01)	-107910.004	65361
KSC		0.27 (0.03)		0.00 (0.02)	0.03 (0.01)	-107910.004	65361
KFS		0.22 (0.03)	0.03 (0.01)	0.00 (0.02)		-107910.338	65361
KF		0.22 (0.02)	0.03 (0.01)			-107910.346	65361
K		0.27 (0.01)				-107915.393	65361
KS		0.27 (0.03)		0.00 (0.02)		-107915.393	65361
FSC			0.13 (0.02)	0.00 (0.02)	(-)0.10 (0.02)	-107915.441	65361
FC			0.13 (0.01)		(-)0.10 (0.01)	-107915.457	65361
SC				0.13 (0.01)	0.03 (0.01)	-107950.759	65361

S	0.13 (0.01)	-107956.204	65361
C	0.03 (0.01)	-108086.077	65361
FS	Error: the variance-covariance matrix V is not positive definite.		65361
F	Error: the variance-covariance matrix V is not positive definite.		65361

Supplementary Table 2: Full GREML-KIN model-fitting results for Education (I)

Note: G (unrel) refers to results from a standard GREML model using a relatedness cutoff of <0.025; selected model = GKC (bold, underlined); GKFSC = full model; LL=log likelihood.

	GRMg	GRMkin	ERMfam	ERMsib	ERMcouple	LL	n	Standard errors				
Model	h2g	h2kin	ef2	es2	ec2			h2g	h2kin	ef2	es2	ec2
G (unrel)	0.17					-21775.04	44520	0.01				
GKFSC	0.12	0.42	0.00	0.00	0.36	-29884.04	65072	0.01	0.09	0.05	0.02	0.05
KFSC		0.54	0.00	0.00	0.36	-29989.21	65072		0.09	0.05	0.02	0.05
GFSC	0.12	0.21	0.00	0.16		-29894.33	65072	0.01		0.01	0.02	0.02
GKSC	0.12	0.43		0.01	0.36	-29884.04	65072	0.01	0.03		0.02	0.01
GKFC	0.12	0.42	0.01		0.36	-29884.07	65072	0.01	0.09	0.05		0.05
GKFS	0.11	-0.11	0.33	-0.07		-29929.32	65072	0.01	0.01	0.01	0.01	
FSC			0.27	0.00	0.10	-29894.33	65072			0.01	0.02	0.02
KSC		0.54		0.00	0.37	-29989.21	65072		0.03		0.02	0.01
KFC		0.54	0.01		0.36	-29989.24	65072		0.09	0.05		0.05
KFS	Error: the variance-covariance matrix V is not positive definite.						65072					
GSC	0.15			0.20	0.38	-29980.86	65072	0.01			0.01	0.01
GFC	0.12		0.21		0.15	-29894.37	65072	0.01		0.01		0.01
GFS	Error: the variance-covariance matrix V is not positive definite.						65072					
<u>GKC</u>	<u>0.12</u>	<u>0.42</u>			<u>0.36</u>	<u>-29884.07</u>	<u>65072</u>	<u>0.01</u>	<u>0.02</u>			<u>0.01</u>
GKS	0.15	0.38		0.01		-31014.3	65072	0.01	0.03		0.02	
GKF	0.10	-0.13	0.32			-29951.07	65072					
SC				0.27	0.37	-30160.08	65072				0.01	0.01
FC			0.27		0.10	-30007.02	65072			0.01		0.01
FS	Error: the variance-covariance matrix V is not positive definite.						65072					
KC		0.55			0.37	-29989.24	65072		0.01			0.01
KS		0.55		0.01		-31168.61	65072		0.03		0.02	

KF	Error: the variance-covariance matrix V is not positive definite.					65072						
GC	0.23			0.35	-30226.87	65072	0.01				0.01	
GS	0.18		0.18		-31099.41	65072	0.01				0.01	
GF	Error: the variance-covariance matrix V is not positive definite.					65072						
GK	0.15	0.40			-31014.47	65072	0.01	0.02				
C				0.38	-135393.9	65072					0.01	
S			0.28		-31346.42	65072					0.01	
F	Error: the variance-covariance matrix V is not positive definite.					65072						
K		0.56			-31168.76	65072		0.01				
G	0.27				-31308.98	65072	0.01					

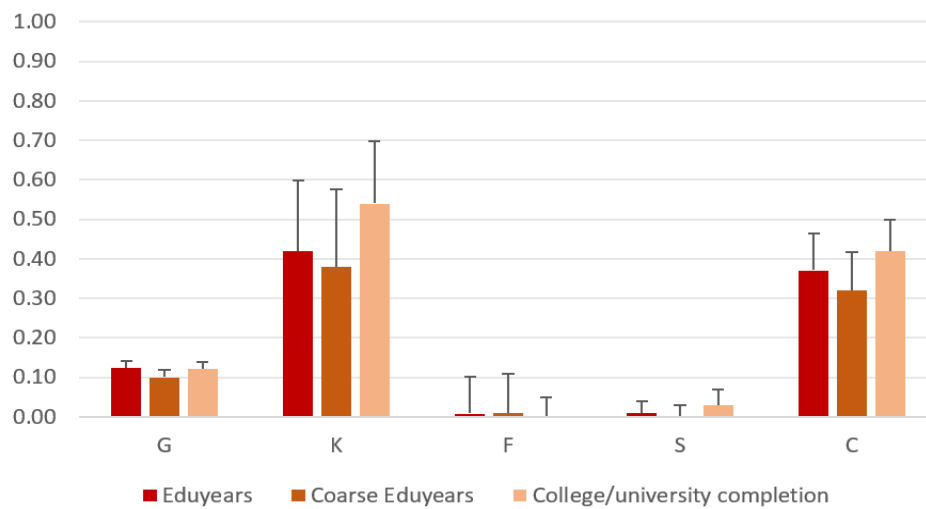
Supplementary Table 3: Full GREML-KIN model-fitting results for Education (II)

Note: G (unrel) refers to results from a standard GREML model using a relatedness cutoff of <0.025; selected model = GKC (bold, underlined); GKFSC = full model; LL=log likelihood.

	GRMg	GRMkin	ERMfam	ERMsis	ERMcouple	LL	n	Standard errors				
Model	h2g	h2kin	ef2	es2	ec2			h2g	h2kin	ef2	es2	ec2
G (unrel)	0.16					-44961.25	21202	0.03				
GKFSC	0.10	0.38	0.01	0.04	0.33	-14583.87	30793	0.02	0.13	0.07	0.02	0.07
KFSC		0.48	0.02	0.04	0.33	-14603.07	30793		0.02	0.12	0.06	0.06
GFSC	0.11		0.20	0.05	0.15	-14586.79	30793	0.02		0.02	0.02	0.02
GKSC	0.10	0.40		0.04	0.35	-14583.88	30793	0.02	0.04		0.02	0.01
GKFC	0.10	0.40	0.04		0.31	-14330.5	30793	0.02	0.12	0.06		0.06
GKFS	Error: the variance-covariance matrix V is not positive definite.						30793	0.02	0.03	0.01	0.02	
FSC			0.25	0.05	0.10	-14607.87	30793			0.02	0.02	0.02
KSC		0.51		0.04	0.35	-14603.1	30793		0.04		0.02	0.01
KFC		0.49	0.04		0.31	-14604.94	30793		0.12	0.06		0.06
KFS	Error: the variance-covariance matrix V is not positive definite.						30793		0.02	0.01	0.02	
GSC	0.16			0.22	0.34	-14622.91	30793	0.02			0.01	0.01
GFC	0.11		0.23		0.11	-14589.02	30793	0.02		0.01		0.02
GFS	Error: the variance-covariance matrix V is not positive definite.						30793	0.02		0.01	0.01	
<u>GKC</u>	<u>0.10</u>	<u>0.47</u>			<u>0.35</u>	<u>-14585.87</u>	<u>30793</u>	<u>0.02</u>	<u>0.03</u>			<u>0.01</u>

GKS	0.14	0.37		0.05		-14844.34	30793	0.02	0.04	0.02
GKF	0.10	-0.21	0.34			-14593.8	30793	0.02	0.03	0.01
SC				0.30	0.35	-14673.97	30793			0.01 0.01
FC			0.29		0.06	-14610.21	30793		0.01	0.01
FS	Error: the variance-covariance matrix V is not positive definite.						30793			
KC		0.58			0.35	-14605.11	30793		0.02	0.01
KS		0.52		0.05		-15077.12	30793		0.04	0.02
KF	Error: the variance-covariance matrix V is not positive definite.						30793		0.02	0.01
GC	0.31				0.33	-14747.85	30793	0.01		0.01
GS	0.20			0.20		-15082.07	30793	0.01		0.01
GF	Error: the variance-covariance matrix V is not positive definite.						30793	0.01	0.01	
GK	0.14	0.45				-15050.59	30793	0.02	0.03	
C					0.36	-15030.62	30793			0.01
S				0.31		-15152.45	30793			0.01
F	Error: the variance-covariance matrix V is not positive definite.						30793			
K		0.60				-15080.09	30793		0.02	
G	0.35					-15191.23	30793	0.01		

Supplementary Figure 1: GREML-KIN model-fitting results for alternative education variables with fewer response categories.



Note: full models with all five variance components are presented here. As indicated in the main text, G= population-level effects of common genotyped SNPs; K= kin-based genetic effects; F= nuclear family (siblings, parent-offspring, couple); S= sibling similarity; C= couple similarity. The coarse edueyears variable was constructed as in the Lee et al. (2018) paper, by collapsing the full edueyears variable from 6 to 3 categories (representing 10, 13 and 19 years of education, respectively). The college/university completion variable was constructed as a binary variable, with individuals scoring 1 if they report university as their highest level of qualification, and 0 if their highest qualification was below that.

Supplementary Figure 1 shows that estimates for the 3 phenotypes are not significantly different – they have overlapping 95% confidence intervals. However, it is interesting that that heritability is higher for the binary college completion variable (66%) than for more detailed and accurate phenotypes capturing more variation. Heritabilities were 54% and 48% for years of education measured with 6 and 3 categories, respectively. Supplementary analyses in Lee et al. (2018) indicated that heritability was significantly reduced in UK Biobank when response options were dropped. The inconsistency with our results could be explained by their use of the full sample, rather than a much smaller sample of relatives, or by their use of a different heritability method (LD score regression).

Supplementary Table 4: Number of SNPs dropped when applying INFO>0.8 cutoff, by MAF bin.

	All INFO	INFO >0.8	No. SNPs dropped
MAF>=0.001 & MAF<=0.01	5051730	4002558	1049172
MAF> 0.01 & MAF<=0.1	4012461	3521025	491436
MAF> 0.1 & MAF<=0.2	1798173	1465411	332762
MAF> 0.2 & MAF<=0.3	1367401	1039265	328136
MAF> 0.3 & MAF<=0.4	1212347	878672	333675
MAF> 0.4 & MAF<=0.5	1144244	801580	342664

Reference

Lee, J.J., Wedow, R., Okbay, A., et al. 2018. Gene discovery and polygenic prediction from a genome-wide association study of educational attainment in 1.1 million individuals. *Nature Genetics* 50(8), pp. 1112–1121.

Appendix 3. Supplementary Information for Chapter 4 (Comparison of adopted and non-adopted individuals reveals gene-environment interplay for education in the UK Biobank).

Table S1: results from training education polygenic scores in 50,000 non-adopted individuals in the UK Biobank. Note: this p-value (10^{-50}) held for all of the standard set-values tested in PRSice (0.001, 0.05, 0.1, 0.2, 0.3, 0.4, 0.5, 1).

Thresh old	PRS.R ²	Full.R ²	Null.R ²	Coefficient	SE	P	Num_SNP
1	0.071066	0.11043	0.039365	196340	3110.32	$p < 10^{-50}$	87372

Table S2: full polygenic prediction results. PRS.SE = standard error around the R² value, derived from bootstrapping with 1000 replications. Note that the samples comprised 6500 non-adopted individuals and 6311 adoptees.

Group	PRS.R ²	PRS.SE	Coefficient	SE	P	N-SNPs	Empirical P
Adopted	0.037	0.0044	27581	1759.93	2.47×10^{-54}	87372	9.99×10^{-05}
Non-adopted	0.074	0.0063	38359.5	1681.36	6.78×10^{-111}	87372	9.99×10^{-05}

Table S3: full results from GxE (education polygenic score x adoption status) model. The outcome variable was years of education. All quantitative covariates were standardised. The table shows results for a linear model testing for additive interaction. The interaction term of interest is highlighted in yellow.

	Estimate	SE	t	Pr(> t)
(Intercept)	13.42	0.4	33.91	<2x10 ⁻¹⁶
PRS_std	1.2	0.28	4.27	2x10 ⁻⁰⁵
adopted	0.12	0.53	0.23	0.82
Sex.0.0	0.63	0.12	5.24	1.66x10 ⁻⁰⁷
Age.at.recruitment.0.0_scaled	-0.8	0.06	-12.8	<2x10 ⁻¹⁶
PC1_scaled	0.19	0.08	2.52	0.01
PC2_scaled	-0.07	0.07	-0.95	0.34
PC3_scaled	0.11	0.07	1.74	0.08
PC4_scaled	4.1x10 ⁻⁰³	0.06	0.07	0.95
PC5_scaled	-0.03	0.07	-0.5	0.62
PC6_scaled	-0.13	0.07	-1.92	0.06
PC7_scaled	-0.07	0.07	-1.1	0.27
PC8_scaled	-0.05	0.06	-0.87	0.38
PC9_scaled	0.12	0.08	1.53	0.13
PC10_scaled	-0.22	0.07	-3.15	1.61x10 ⁻⁰³
Centre11002	0.69	0.52	1.31	0.19
Centre11003	0.11	0.52	0.21	0.83
Centre11004	-0.27	0.51	-0.53	0.59
Centre11005	1.15	0.52	2.2	0.03
Centre11006	0.03	0.5	0.07	0.95
Centre11007	1.21	0.47	2.6	0.01
Centre11008	-0.14	0.46	-0.3	0.76
Centre11009	-0.56	0.47	-1.2	0.23
Centre11010	0.17	0.44	0.39	0.7
Centre11011	0.2	0.45	0.44	0.66
Centre11012	1.54	0.57	2.69	0.01
Centre11013	0.36	0.46	0.78	0.43
Centre11014	-0.22	0.46	-0.49	0.63
Centre11016	-0.02	0.45	-0.05	0.96
Centre11017	-0.09	0.48	-0.18	0.86

Centre11018	1.07	0.48	2.22	0.03
Centre11020	0.64	0.48	1.33	0.18
Centre11021	0.48	0.47	1.01	0.31
Centre11022	-0.62	0.95	-0.65	0.52
Centre11023	-3.52	1.51	-2.33	0.02
PRS_std:adopted	-0.33	0.09	-3.65	2.66x10 ⁻⁰⁴
PRS_std:Sex.0.0	-0.06	0.09	-0.66	0.51
PRS_std:Age.at.recruitment.0.0_scaled	-0.03	0.04	-0.64	0.52
PRS_std:PC1_scaled	-0.04	0.05	-0.93	0.35
PRS_std:PC2_scaled	-0.03	0.05	-0.69	0.49
PRS_std:PC3_scaled	-0.1	0.04	-2.21	0.03
PRS_std:PC4_scaled	-0.05	0.04	-1.08	0.28
PRS_std:PC5_scaled	-0.01	0.05	-0.25	0.80
PRS_std:PC6_scaled	0.03	0.05	0.72	0.47
PRS_std:PC7_scaled	0.02	0.05	0.38	0.70
PRS_std:PC8_scaled	0.02	0.04	0.37	0.71
PRS_std:PC9_scaled	1.7x10 ⁻⁰⁴	0.05	4x10 ⁻⁰³	1.00
PRS_std:PC10_scaled	3.6x10 ⁻⁰³	0.05	0.08	0.94
PRS_std:Centre11002	0.37	0.36	1.03	0.30
PRS_std:Centre11003	0.46	0.36	1.28	0.20
PRS_std:Centre11004	-0.27	0.37	-0.74	0.46
PRS_std:Centre11005	0.76	0.37	2.09	0.04
PRS_std:Centre11006	0.25	0.35	0.73	0.47
PRS_std:Centre11007	0.25	0.33	0.75	0.45
PRS_std:Centre11008	0.01	0.33	0.04	0.97
PRS_std:Centre11009	0.14	0.33	0.43	0.67
PRS_std:Centre11010	-0.14	0.31	-0.45	0.65
PRS_std:Centre11011	0.08	0.31	0.26	0.8
PRS_std:Centre11012	0.45	0.39	1.16	0.25
PRS_std:Centre11013	0.03	0.32	0.11	0.92
PRS_std:Centre11014	0.19	0.33	0.59	0.56
PRS_std:Centre11016	0.26	0.32	0.8	0.43
PRS_std:Centre11017	0.11	0.35	0.33	0.74

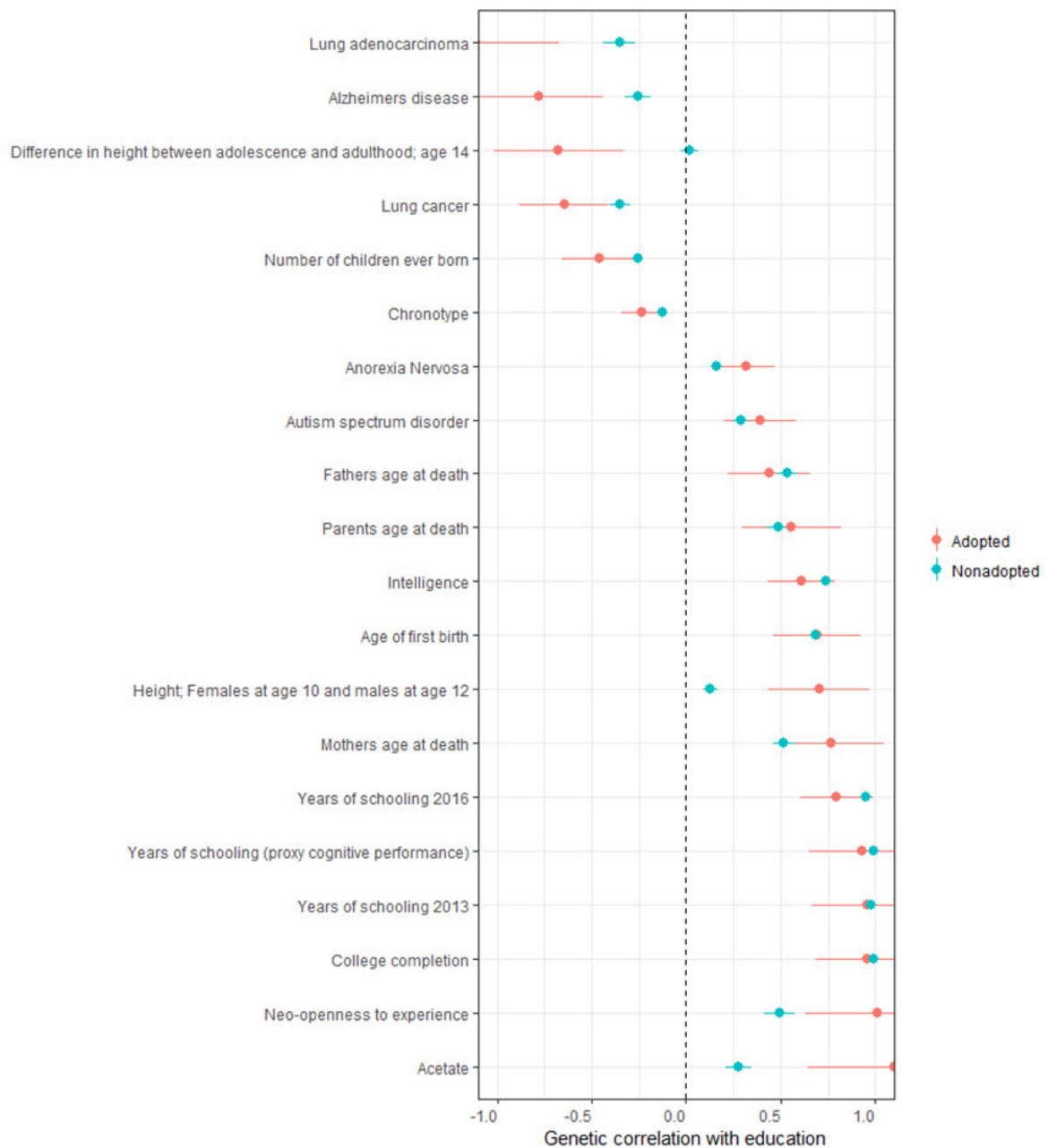
PRS_std:Centre11018	0.35	0.34	1.03	0.31
PRS_std:Centre11020	0.39	0.34	1.16	0.24
PRS_std:Centre11021	0.21	0.33	0.63	0.53
PRS_std:Centre11022	-0.21	0.62	-0.34	0.73
PRS_std:Centre11023	0.14	1.03	0.14	0.89
adopted:Sex.0.0	-0.27	0.17	-1.58	0.12
adopted:Age.at.recruitment.0.0_scaled	-0.23	0.09	-2.70	0.01
adopted:PC1_scaled	-0.03	0.09	-0.37	0.71
adopted:PC2_scaled	0.04	0.10	0.46	0.64
adopted:PC3_scaled	0.03	0.09	0.36	0.72
adopted:PC4_scaled	-0.07	0.09	-0.85	0.39
adopted:PC5_scaled	0.11	0.09	1.16	0.25
adopted:PC6_scaled	0.05	0.10	0.49	0.62
adopted:PC7_scaled	-0.02	0.10	-0.25	0.80
adopted:PC8_scaled	0.12	0.09	1.40	0.16
adopted:PC9_scaled	-0.15	0.09	-1.60	0.11
adopted:PC10_scaled	0.12	0.09	1.28	0.20
adopted:Centre11002	-0.44	0.72	-0.61	0.54
adopted:Centre11003	-0.61	0.70	-0.87	0.39
adopted:Centre11004	0.10	0.71	0.13	0.89
adopted:Centre11005	-1.12	0.73	-1.52	0.13
adopted:Centre11006	0.07	0.68	0.10	0.92
adopted:Centre11007	-1.42	0.64	-2.23	0.03
adopted:Centre11008	-0.03	0.63	-0.06	0.96
adopted:Centre11009	0.27	0.63	0.43	0.67
adopted:Centre11010	-0.68	0.59	-1.14	0.26
adopted:Centre11011	-0.42	0.60	-0.70	0.48
adopted:Centre11012	-0.73	0.77	-0.95	0.34
adopted:Centre11013	-0.88	0.62	-1.43	0.15
adopted:Centre11014	0.07	0.64	0.12	0.91
adopted:Centre11016	-0.29	0.62	-0.46	0.64
adopted:Centre11017	0.02	0.67	0.03	0.97
adopted:Centre11018	-0.23	0.66	-0.35	0.73

adopted:Centre11020	-0.59	0.66	-0.91	0.36
adopted:Centre11021	-0.53	0.65	-0.82	0.41
adopted:Centre11022	0.96	1.29	0.74	0.46
adopted:Centre11023	1.04	2.15	0.48	0.63

Table S4: Comparison of adopted and non-adopted individuals for mean years of education per decile of polygenic score for education.

Decile	Adopted			Non-adopted				
	Mean education years	SE	N	Mean education years	SE	N	zdiff	zdiffp
1	-0.26	0.04	723	-0.48	0.04	559	-3.97	7.07x10⁻⁰⁵
2	-0.18	0.04	710	-0.27	0.04	571	-1.74	0.081
3	-0.20	0.04	672	-0.27	0.04	609	-1.21	0.225
4	-0.09	0.04	640	-0.11	0.04	641	-0.50	0.619
5	-0.02	0.04	638	-0.12	0.04	643	-1.95	0.051
6	0.02	0.04	623	0.002	0.04	658	-0.37	0.708
7	0.12	0.04	625	0.09	0.04	656	-0.65	0.515
8	0.15	0.04	583	0.21	0.04	698	1.11	0.269
9	0.27	0.04	562	0.26	0.04	719	-0.05	0.962
10	0.36	0.04	535	0.46	0.03	746	1.93	0.053

Figure S1: genetic correlations between traits in the LD Hub database and years of education (plus standard errors) for adopted (red) and for non-adopted (blue) individuals. The figure only shows results for traits that showed significant genetic correlations with education in the sample of adoptees at $p < 0.05$.



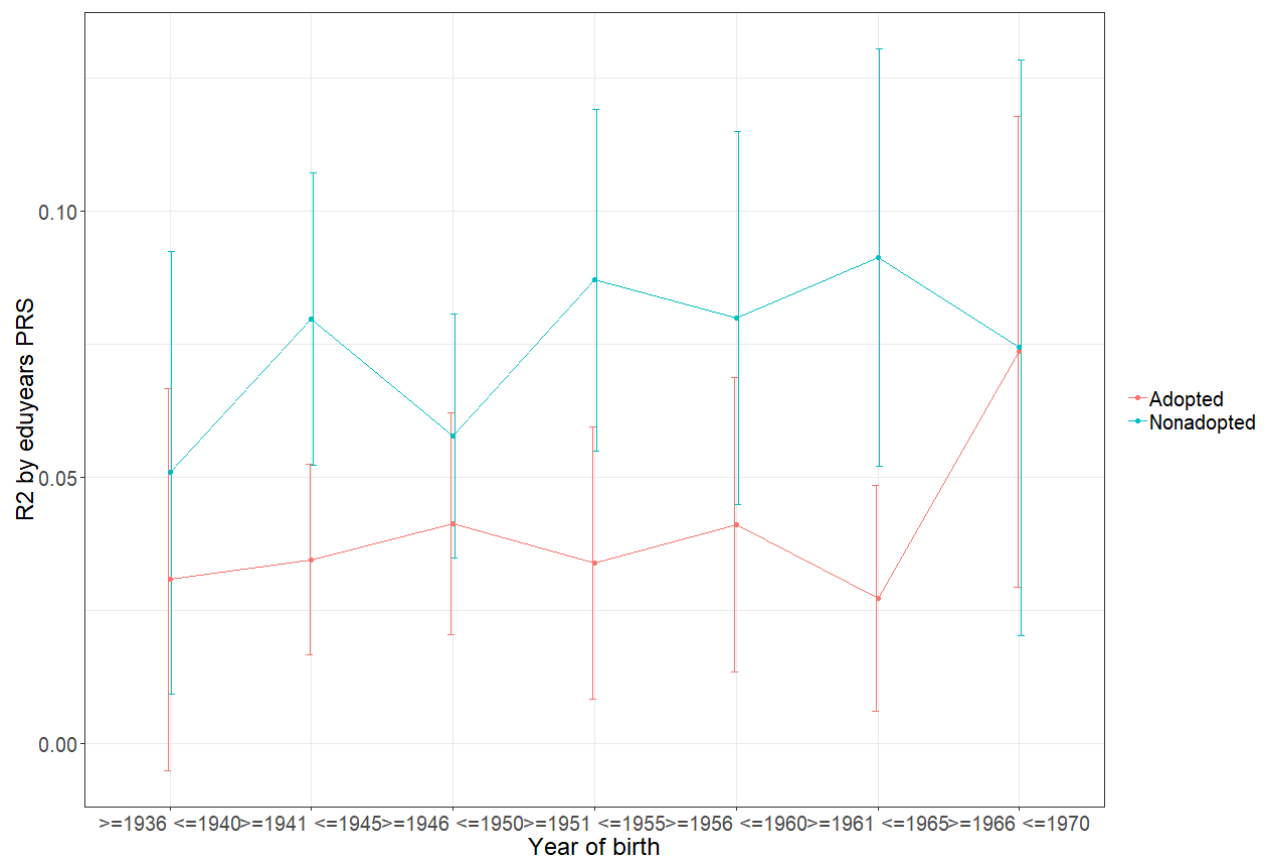
Due to the relatively small sample of adoptees, only 4 traits were significantly genetically correlated with education years in the adopted sample after correcting for multiple testing by using a stringent p-value of 0.000202 (0.05/247): years of schooling (2016), college completion, intelligence, and years of schooling (proxy). All traits in this figure also showed significant genetic correlations with education years in the non-adopted sample at $p < 0.000202$, except height of females age 10 and males at age 12 ($p = 0.0004$), Alzheimer's disease ($p = 0.0002$) and difference in height between adolescence and adulthood ($p = 0.7265$). None of the genetic correlations with education were significantly different between adoptees and non-adopted individuals at $p < 0.000202$.

Table S5: Genetic correlations with binary adopted/not adopted variable, from LD Hub, ranked by p-value. The table only shows correlations that remained significant after a Bonferroni correction for multiple testing ($p < 0.0002$ i.e. $0.05/247$).

Trait	Genetic correlation	se	p
Years of schooling 2016	-0.5216	0.0615	2.28×10^{-17}
Age of first birth	-0.6589	0.0836	3.23×10^{-15}
College completion	-0.5584	0.0882	2.41×10^{-10}
Years of schooling (proxy cognitive performance)	-0.522	0.0851	8.38×10^{-10}
Years of schooling 2013	-0.4739	0.0823	8.39×10^{-9}
Depressive symptoms	0.5176	0.1026	4.60×10^{-7}
Intelligence	-0.306	0.0679	6.66×10^{-6}
Obesity class 1	0.2796	0.0652	1.77×10^{-5}
Waist-to-hip ratio	0.2507	0.0626	6.23×10^{-5}
Obesity class 2	0.3368	0.0851	7.58×10^{-5}
Extreme bmi	0.2943	0.0748	8.32×10^{-5}
Mothers age at death	-0.6567	0.1679	9.14×10^{-5}
Body mass index	0.2486	0.0648	0.0001
Insomnia	0.4258	0.1098	0.0001

Note: These significant negative genetic correlations with educational attainment, age of motherhood, and poor physical health agree with evidence that adoptees are more likely than individuals from comparable origins to be born to young mothers, to have had a low birthweight, and to have received suboptimal obstetric care (Maughan et al. 1998), although the two groups were highly similar apart from these factors. The fact that adoptee status had a low heritability (6%; liability scale) means that interpretation of these genetic correlations should be cautious.

Figure S2: Comparing the variance explained in years of education by polygenic scores (PRS) for years of education, per year-of-birth grouping, for adoptees and non-adopted individuals. 95% CIs were obtained by bootstrapping with 1000 replications.



Any systematic differences between adoptive and biological families are likely to have changed over time. The variance explained by polygenic scores is more variable in the group of individuals reared by their biological relatives, but generally increases over time. The plot shows that prediction is stronger for adoptees in the youngest age group (although this increase was not significant). Due to increasing use of contraceptives across the 20th century, adoption became less related to young motherhood and more related to removal of children from high risk environments. Adoptions occurring later in the twentieth century were more likely to involve older children rather than babies (Maughan et al. 1998; Keating 2008). We speculate that these changes could mean that younger adoptees in our sample are less generalisable, more likely to have come from high risk environments, and more likely to have spent more time with their biological relatives, experiencing passive gene-environment correlation. However, other than the non-significant spike in the youngest age group of adoptees, polygenic prediction remains stable across cohorts (at ~4%). It could be argued that the variance explained by polygenic scores in adoptees provide a rough benchmark of ‘direct’ genetic influence remaining throughout periods of great social change.

Table S6: sample sizes by year-of-birth grouping.

	Adopted	Non-adopted group d (N=6500)
>=1936 <=1940	341	448
>=1941 <=1945	1687	1413
>=1946 <=1950	1312	1503
>=1951 <=1955	839	1124
>=1956 <=1960	781	867
>=1961 <=1965	869	796
>=1966 <=1970	482	349

The post-war spike in the number of adoptions (and number of births generally) is consistent with the 'baby scoop' and the 'baby boom'.

References

- Keating, J. 2008. *A Child for Keeps: The History of Adoption in England, 1918-45*. Springer.
- Maughan, B., Collishaw, S. and Pickles, A. 1998. School achievement and adult qualifications among adoptees: a longitudinal study. *Journal of Child Psychology and Psychiatry, and Allied Disciplines* 39(5), pp. 669–685.

Appendix 4. Supplementary Information for Chapter 5 (How important are parents in the development of child anxiety and depression? A genomic analysis of parent-offspring trios in the Norwegian Mother Father and Child Cohort Study).

Supplementary Table 1: MoBa genotyping and imputation information

Cohort	Genotyping chip	Genotyping center	Cohort design	N SNPs			N individuals		
				Genotyped	Used for imputation	Imputed	Sent	Genotyped	Imputed
Harvest	Illumina HumanCoreExome12v1.1	Genomics core facility, Trondheim, Norway		542,585	384,406	39,975,206	20,668	1,692	20,231
				542,585				18,972	
	Illumina HumanCoreExome24v1.0			547,644	395,736	39,963,876	12,874	12,874	12,757
Rotterdam	Illumina Global Screening Array MD v.1.0	ERASMUS MC, Rotterdam, Netherlands	Trios	692,367	567,903	39,791,709	27,000	17,949	17,742
	Illumina Global Screening Array MD v.1.0			692,338	571,759	39,787,853		9,041	8,959
Norment	Illumina Global Screening Array MD v.01.0 + 50k custom content (OmniExpress overlap)		Trios	693,143	588,807	39,770,805	9,841	9,632	9,628
	Illumina HumanOmniExpress-24v1.0		Unrelated parents	708,882	665,648	39,693,964	6,040	2,976	2,851
	Illumina HumanOmniExpress-24v1.0			710,146	667,978	39,691,634		2,983	2,983
	Illumina InfiniumOmniExpress-24v1.2		Trios	712,628	683,577	39,676,035	19,611	17,730	17,730
ADHD	Illumina InfiniumOmniExpress-24v1.2	deCODE Genetics, Reykjavik, Iceland	Case trios and singleton controls	713,599	670,860	39,688,752	5,790	5,410	5,229
MoBa Genetics Release 1							101,824	99,259	98,110

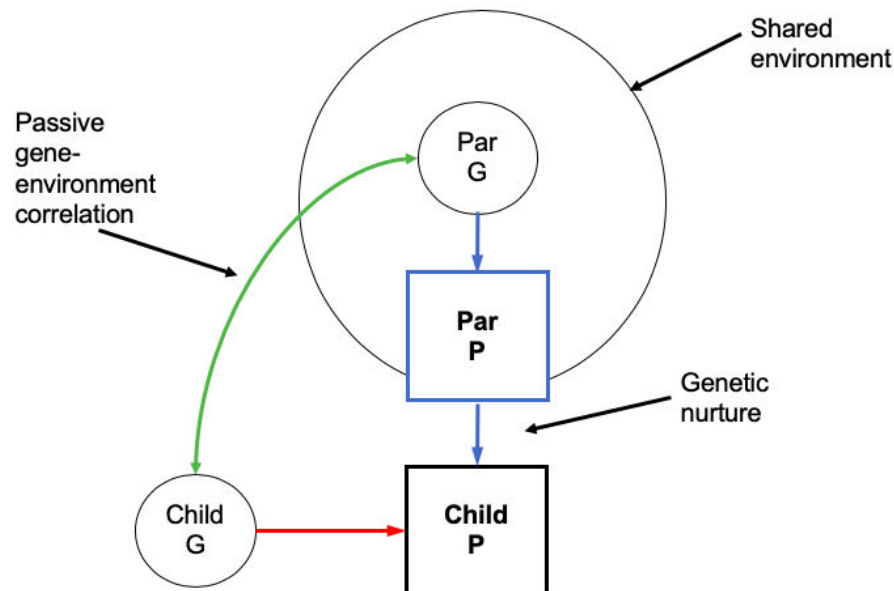
Supplementary Table 2: Full RDR results. Note that V_g = direct effect of child genetics; $V_{e \sim g}$ = genetic nurture i.e. variance in child's environment explained by parent genetics; $C_{g,e}$ = variance explained by covariance between direct and nurturing genetic effects; SE=standard error; W/M/Y=weeks months years of the child's life. The key results reported in the main text are shown in bold.

Phenotype	Covariates	V_g	$V_{e \sim g}$	$C_{g,e}$	SE_ V_g	SE_ $V_{e \sim g}$	SE_ $C_{g,e}$	N trios
Anxiety (sex resid)	NA	0.05	0.00	0.03	0.07	0.07	0.07	11598
Anxiety (sex resid)	10 PCs, batch	0.05	-0.01	0.01	0.07	0.07	0.07	11598
Anxiety (sex resid)	10 PCs, batch, longitudinal mat anx/dep	0.01	-0.02	0.05	0.07	0.07	0.07	11566
Anxiety (sex resid)	10 PCs, batch, mat anx/dep at 15W	0.03	0.02	0.02	0.07	0.07	0.07	11409
Anxiety (sex resid)	10 PCs, batch, mat anx/dep at 30W	0.04	-0.01	0.03	0.07	0.07	0.07	11350
Anxiety (sex resid)	10 PCs, batch, mat anx/dep at 6M	0.03	-0.03	0.05	0.07	0.07	0.07	11261
Anxiety (sex resid)	10 PCs, batch, mat anx/dep at 18M	0.04	-0.03	0.03	0.07	0.07	0.08	10571
Anxiety (sex resid)	10 PCs, batch, mat anx/dep at 3Y	-0.01	-0.08	0.07	0.08	0.09	0.09	9607
Anxiety (sex resid)	10 PCs, batch, mat anx/dep at 8Y	-0.01	-0.05	0.09	0.07	0.07	0.08	11409
Anxiety (sex resid) - only trios genotyped on the same batch	10 PCs, batch	0.07	-0.02	0.04	0.08	0.08	0.08	10378
Anxiety (sex resid), components constrained >0	10 PCs, batch	0.05	0	0.01	0.07	0.07	0.07	11598
Depression (sex resid)	NA	0.16	0.13	-0.10	0.07	0.07	0.07	11584
Depression (sex resid)	10 PCs, batch	0.19	0.14	-0.16	0.07	0.07	0.07	11584
Depression (sex resid)	10 PCs, batch, longitudinal mat anx/dep	0.17	0.05	-0.11	0.07	0.07	0.07	11552
Depression (sex resid)	10 PCs, batch, mat anx/dep at 15W	0.21	0.09	-0.13	0.07	0.07	0.07	11395
Depression (sex resid)	10 PCs, batch, mat anx/dep at 30W	0.15	0.06	-0.09	0.07	0.07	0.07	11335
Depression (sex resid)	10 PCs, batch, mat anx/dep at 6M	0.16	0.07	-0.11	0.07	0.07	0.07	11247
Depression (sex resid)	10 PCs, batch, mat anx/dep at 18M	0.20	0.13	-0.16	0.07	0.07	0.08	10559
Depression (sex resid)	10 PCs, batch, mat anx/dep at 3Y	0.17	0.01	-0.09	0.08	0.08	0.09	9595
Depression (sex resid)	10 PCs, batch, mat anx/dep at 8Y	0.15	0.05	-0.12	0.07	0.07	0.07	11390

Depression (sex resid) - only trios genotyped on the same batch	10 PCs, batch	0.19	0.12	-0.18	0.07	0.07	0.08	10364
Depression (sex resid), components constrained >0	10 PCs, batch	0.14	0.09	0.00	0.07	0.07	0.07	11584

Supplementary Figure 1: Comparison of related concepts: passive gene-environment correlation, genetic nurture, and the shared environment. Note: G=genetic, P=phenotype, Par=parent.

The term 'genetic nurture' was coined recently (Kong et al. 2018), but overlaps somewhat with existing concepts from quantitative genetics: passive gene-environment correlation and shared environment. We now discuss the similarities and differences between these concepts with reference to Supplementary Figure 1 (SF1). In SF1, the child phenotype is influenced directly by genes present in the child (red). Since children are genetically related to their parents, direct genetic effects become correlated (green) with genetic effects affecting children *indirectly* via parental behaviour (blue).



Genetic nurture and passive gene-environment correlation both refer to the intertwining of genetic and environmental effects generated because parental behaviour is heritable. 'Genetic nurture', estimated using intergenerational genomic data, describes the indirect effect of the parent genome **over and above the effects of genes shared with the child**. This can be done by estimating effects of non-transmitted alleles (Kong et al. 2018; Bates et al. 2018), or by estimating the independent effect of parental DNA in a multiple regression controlling for child DNA (i.e. the blue path in SF1 is estimated, adjusting for the green and red paths). In contrast, 'passive gene-environment correlation' is focussed on **shared genes** between parents and children (green double-headed arrow), which create associations between child genes and parent phenotypes. Here, we refer to passive gene-environment correlation specifically as the correlation between the direct genetic effect and the genetic nurture effect. This differs slightly from its usage in pedigree studies, which stresses shared genes independently affecting the child and parent phenotypes. There, the parent genetic effect is on the measured parent phenotype, and does not necessarily in turn affect the child.

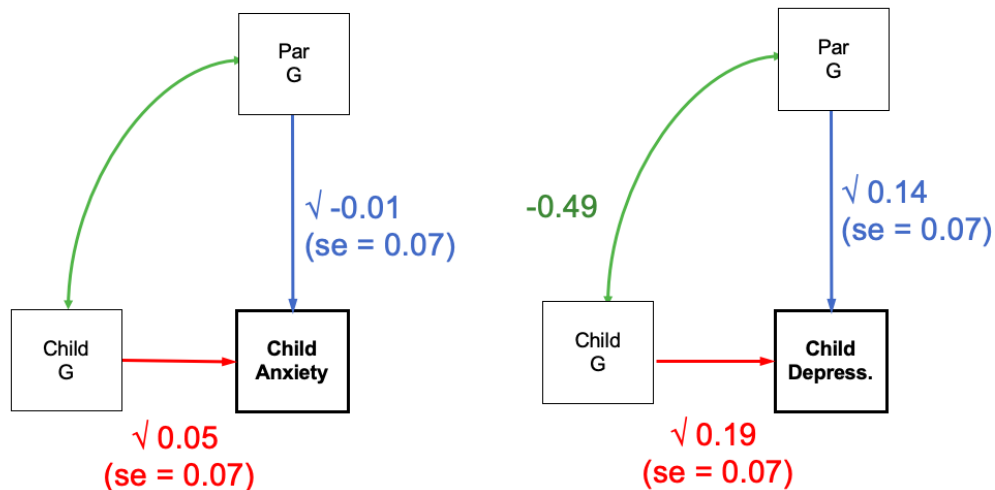
Since genetic nurture represents an *environmentally-mediated* genetic effect of parents, rather than an effect arising from the *correlation* between genes and environments, it may be conceptually closer to estimates of the shared environmental variance in the classical twin design than to passive gene-environment correlation (note that genetic nurture and passive gene-environment correlation are likely to both load onto the shared environment component in twin models). Supplementary Figure 1 depicts how the shared environment is highly overlapping with genetic nurture. Both constructs capture main effects of the family environment, which are uncorrelated with child-specific genetic effects, and therefore do not differ among siblings. Genetic nurture only includes parent effects, whereas the shared environment captures any non-genetic effect making siblings similar. However, it cannot be assumed that genetic nurture effects are necessarily a subset of the shared environment variance estimated in a pedigree design. Genetic nurture may also share properties with the genetic component in pedigree models, because it would also increase resemblance among cousins and half siblings when they are included in the model (the shared environment is typically defined as one that is shared by siblings living together so would not make e.g. cousins similar).

References

Bates, T.C., Maher, B.S., Medland, S.E., McAloney, K., Wright, M.J., Hansell, N.K., Kendler, K.S., Martin, N.G. and Gillespie, N.A. 2018. The Nature of Nurture: Using a Virtual-Parent Design to Test Parenting Effects on Children's Educational Attainment in Genotyped Families. *Twin Research and Human Genetics* 21(2), pp. 73–83.

Kong, A., Thorleifsson, G., Frigge, M. L., ... Stefansson, K. (2018). The nature of nurture: Effects of parental genotypes. *Science*, 359(6374), 424–428.

Supplementary Figure 2: Simplified path diagram for RDR results. Note: Bold=focal phenotype, G=genetic, Par=parent. Red and blue values show the direct child genetic and parental genetic nurture path estimates, respectively. The green value for depression represents the correlation between direct and nurturing genetic effects. This was estimated by dividing the covariance component [-0.16] by two times the square root of the product of the variance components [$2 \times \sqrt{(0.19 \times 0.14)}$]. The correlation could not be computed for anxiety, due to the negative genetic nurture path estimate. Not shown: residual effects on the phenotypes, including parent genetic effects not tagged by common SNPs, individual-specific environmental effects, chance factors and error.



Supplementary Figure 3: Simplified path diagram for pedigree results. Note: Bold=focal phenotype, square=observed, circle=latent, G=genetic, C= shared environment. Not shown: residual (includes non-shared environmental effects and error). The correlation between genetic and shared environmental influences is not estimated in this model.

