

Review



Cite this article: Sallis H, Davey Smith G, Munafò MR. 2018 Genetics of biologically based psychological differences. *Phil. Trans. R. Soc. B* **373**: 20170162. <http://dx.doi.org/10.1098/rstb.2017.0162>

Accepted: 22 November 2017

One contribution of 20 to a theme issue 'Diverse perspectives on diversity: multi-disciplinary approaches to taxonomies of individual differences'.

Subject Areas:

behaviour, genetics

Keywords:

heritability, GWAS, behavioural traits, personality, temperament, Mendelian randomization

Author for correspondence:

Hannah Sallis
e-mail: hannah.sallis@bristol.ac.uk

Genetics of biologically based psychological differences

Hannah Sallis^{1,2,3}, George Davey Smith¹ and Marcus R. Munafò^{1,2}

¹MRC Integrative Epidemiology Unit, University of Bristol, Bristol, UK

²UK Centre for Tobacco and Alcohol Studies, School of Experimental Psychology, University of Bristol, Bristol, UK

³Centre for Academic Mental Health, Population Health Sciences, Bristol Medical School, University of Bristol, Bristol, UK

id HS, 0000-0002-4793-6290; MRM, 0000-0002-4049-993X

In recent years, substantial effort has gone into disentangling the genetic contribution to individual differences in behaviour (such as personality and temperament traits). Heritability estimates from twin and family studies, and more recently using whole genome approaches, suggest a substantial genetic component to these traits. However, efforts to identify the genes that influence these traits have had relatively little success. Here, we review current work investigating the heritability of individual differences in behavioural traits and provide an overview of the results from genome-wide association analyses of these traits to date. In addition, we discuss the implications of these findings for the potential applications of Mendelian randomization.

This article is part of the theme issue 'Diverse perspectives on diversity: multi-disciplinary approaches to taxonomies of individual differences'.

1. Introduction

Despite the early promise of behavioural genetic research, efforts to disentangle the genetic contribution to individual differences in behaviour (e.g. personality and temperament traits) have been slow. Early studies relied on a candidate gene approach to identify genes influencing these traits; however, many of these failed to replicate, despite having a plausible biological mechanism. More recent studies have used whole genome approaches to investigate the genetic architecture of behavioural traits. However, unlike for many other complex traits such as height [1,2] and schizophrenia [3], relatively few genetic variants have been identified that are robustly associated with temperament and individual differences in personality.

It has been argued that a small number of factors can be used to account for individual differences in personality [4–8]. Although there is no universally accepted framework of personality, the proposed factors (such as those suggested in Eysenck's five-factor model (FFM)) provide a good starting point when investigating the genetic architecture of these individual differences [7]. It is likely that individual molecular genetic effects associated with these traits will be small and large sample sizes will therefore be required to detect any association. By measuring specific traits, studies are able to harmonize phenotypes and pool analyses across cohorts. This will increase the power to detect genetic effects relevant to individual differences.

With recent advances in computing and availability of increasingly rich data sources, there has been a surge in the number of behavioural genetic studies investigating personality and temperament traits. Behavioural genetics has the potential both to quantify genetic influences on these traits and to indirectly quantify environmental influences [9]. Here, we provide a synthesis of these behavioural genetic studies to date. We begin by reviewing current work investigating the heritability of individual differences in personality and

Box 1. Heritability.

Heritability is the proportion of variation in a phenotype (V_P) that can be attributed to genetic differences for the particular context and timepoint at which it is estimated. This can include the proportion of variance due to additive genetic effects (V_A), known as narrow-sense heritability (h^2), or the proportion due to all genetic effects (V_G), known as broad-sense heritability (H^2) [12]. In addition to additive genetic effects, H^2 includes both between-loci interactions (epistasis) and within-loci interactions (dominance) effects.

Broad-sense heritability: $H^2 = V_G/V_P$

and

Narrow-sense heritability: $h^2 = V_A/V_P$.

Heritability can be estimated using a number of methods. These include twin, family and adoption studies, in addition to more recent methods that use data captured by genome-wide arrays. Estimates calculated according to these models can be thought of as ‘true’ heritability. These estimates will incorporate heritability due to variants across the entire genome, including rare variants and those not captured by SNPs included on genotyping platforms. However, twin models, by design, use closely related individuals. These individuals are thus likely to share a great deal of their environment, which could lead to false inflation of heritability estimates.

Additional methods estimate heritability using SNP-based approaches. This includes heritability due to variants robustly associated with the phenotype of interest that are identified through genome-wide association study (GWASs) (h_{GWAS}^2), and heritability attributed to all SNPs captured on a GWAS array (h_{SNP}^2). h_{SNP}^2 is typically larger than h_{GWAS}^2 , although h_{GWAS}^2 will increase with sample size as more variants are identified through GWASs, and is expected to approach h_{SNP}^2 [13,14]. These approaches generally estimate heritability using unrelated individuals, which should minimize any false inflation due to shared environment, but do not give an estimate of true narrow-sense heritability because they do not include all additive genetic variance.

While the classic twin design can be used to estimate the proportion of phenotypic variation that can be attributed to genetic differences within a population, SNP-based methods can provide an estimate of the amount of variation that could potentially be explained using information from genome-wide SNP chips. These chips tend to encompass common variation and work by genotyping a fraction of the genome, in the hope that these SNPs will ‘tag’ the majority of the remaining SNPs. This h_{SNP}^2 is an estimate of how much of the phenotypic variation could potentially be explained using the variants contained on these chips, rather than an estimate of ‘true’ h^2 . Any discrepancies between h_{SNP}^2 and the narrow-sense heritability estimated from twin studies may be attributed to variation explained by rare variants, small effect sizes not detectable using current sample sizes or common variation not tagged by SNPs included on the current chip.

temperament traits; for the purposes of this review, we will refer to individual differences in these traits as simply ‘individual differences’. We also look at the results from genome-wide association analyses of these traits to date and discuss the implications of these findings for the potential for applying Mendelian randomization (MR) to individual differences in behaviour [10,11].

2. Estimating heritability

To investigate the genetic contribution to a phenotype, it is useful to first estimate its heritability (box 1). Heritability is the proportion of variation in a phenotype that can be attributed to genetic differences; these estimates are specific to the particular context and the timepoint at which they are estimated [13,15–17]. For example, if a trait has a heritability of 30%, then 30% of the variation in this trait is assumed to be due to genetic variation. However, although these estimates provide an idea of the size of the genetic component for a particular trait, they do not give us any information about which genes are likely to be responsible for it [18].

(a) Using twin studies to estimate heritability

Twin studies have been extensively used to disentangle the role of genetics and environment on human traits [19]. These models assume there are three distinct influences on phenotypic variation (V_P) and these comprise additive genetic effects (V_A),

shared environmental effects (V_C) and non-shared environmental effects (V_E) [20]. Heritability estimates resulting from twin models are valid under certain assumptions. These assumptions include: (i) the twins are representative of the general population in terms of the trait, (ii) environmental effects are shared to the same extent by identical (MZ) and non-identical (DZ) twins, (iii) gene–environment interactions for the trait are minimal, and (iv) there is no assortative mating in the population [20]. Violations of these assumptions can result in biased (i.e. either increased or decreased) heritability estimates, and it seems reasonable that these assumptions could impact differently depending on the trait of interest. It should also be recognized that while one of the variance components is referred to as ‘non-shared environment’, this does not necessarily relate to environmental factors as they are usually understood, as the effects of somatic mutations, stochastic epigenetic changes and other random processes, as well as measurement error, are subsumed under this heading [13].

A recent meta-analysis by Polderman *et al.* [19] focusing on twin studies of human behavioural traits found that of all the phenotypes studied (more than 500 distinct traits), temperament and personality traits were among the top 10 most investigated. This study investigated the relative contribution of genetics and environment to a comprehensive list of traits studied over the previous 50 years, as well as assessing the presence of non-additive genetic effects. The authors suggest that of all the studies ($N = 568$) investigating temperament and personality traits, the majority (84%) of published

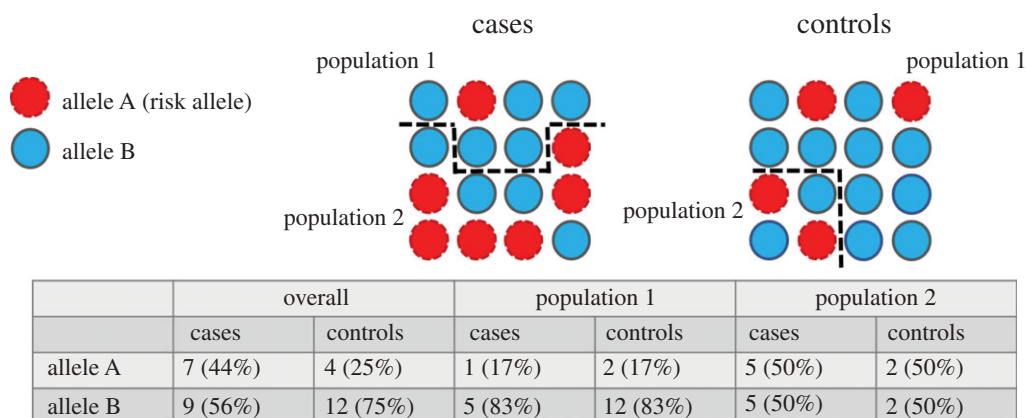


Figure 1. Population stratification. When looking at the sample overall, there is a higher frequency of risk allele A in cases than controls. However, there are two ‘hidden populations’ within this sample and the risk allele has a higher frequency in one population than the other. Cases and controls are sampled disproportionately, resulting in a false positive association for this variant. (Online version in colour.)

Box 2. Population stratification.

Allele frequencies can differ across populations, which can cause difficulties in association studies (figure 1). These differences in ancestry, or underlying population substructure, need to be taken into account to prevent finding associations that are due to this population substructure rather than to phenotype-associated variants. Methods of accounting for this include restricting samples to populations of common ancestry, adjusting results for a genomic inflation factor (λ) [27] or performing principal component analysis to account for variation in the data due to population differences, and adjusting the regression model for these components [28].

twin correlations were consistent with a simple model containing just additive genetic effects [16,19]. However, many of studies in the meta-analysis did not report non-additive variance components, because these are generally (and arguably erroneously) constrained to zero when there is no ‘significant’ effect. This outcome is linked to sample size; thus, it is possible that this lack of non-additive effect is a result of the sample sizes analysed [16,21].

A recent meta-analysis of personality and temperament traits has combined evidence from twin, family and adoption studies [22]. The author, Vukasović & Bratko, find evidence for a heritable component to individual differences in personality, with heritability estimated at approximately 40%. The authors investigated the contribution of the different study designs and found that heritability estimates were consistently higher among twin studies than family and adoption studies. The authors also looked at potential moderating effects of gender and personality model on heritability. Although there is some suggestion in the literature that heritability differs according to gender [23–26], there was no strong evidence of a moderating effect when combining across all studies. Twin studies and Eysenck’s theory of personality were both over-represented in the data; however, the authors only found evidence for a moderating effect of study design.

Despite the variation across models of personality structure considered in these studies, there is clear evidence that a proportion of variation in these individual difference traits can be attributed to genetic variation. Vukasović & Bratko [22] found no difference in heritability regardless of the model of personality used across studies. This is somewhat reassuring given the different models considered across cohorts and the ongoing development of personality theory.

An additional consideration when undertaking genetic analysis is the population under investigation. In this

meta-analysis, there is an underrepresentation of Asian populations and no representation of South American or African populations. Performing genetic analyses across different samples can be difficult owing to the possibility of population stratification (box 2 and figure 1). Risk alleles can occur with different frequencies across populations, and this can induce spurious associations. The population from which a sample has been drawn, therefore, needs to be taken into consideration to ensure that any observed association is not simply due to differences in allele frequency (see *ALDH2* Mendelian randomization example in section 5).

(b) Single-nucleotide polymorphism-based approaches to estimating heritability

Although the twin studies discussed previously provide estimates of true narrow-sense heritability (h^2), it can be problematic to obtain large enough samples to precisely estimate these effects. With recent advances in computing, and the increasing willingness of many cohorts to combine efforts through large-scale consortia, it has been possible to create increasingly large genotyped samples with data on a wide range of phenotypic measures. Unlike twin studies which rely on assumed levels of genetic correlation between twin pairs, we are now able to measure the degree of genetic similarity between individuals in a dataset and look at the extent to which genetic and phenotypic similarity correlates. This means that it is now possible to calculate heritability estimates using non-related individuals [14,29–33]. Methods have also been developed that use summary statistics from genome-wide association studies (GWASs) to calculate heritability estimates even when individual-level data are not available. These techniques use information on the correlation or linkage disequilibrium (LD) structure across the genome to estimate

Table 1. Neuroticism heritability estimates from recent GWASs.

study	sample	h^2_{SNP}	<i>N</i>
Bae <i>et al.</i> [39]	LLFS	0.252, $p = 1.7 \times 10^{-15}$	4595
de Moor <i>et al.</i> [40]	NTR	0.147, $p = 0.02$	3599
	QIMR	0.157, $p = 0.18$	3369
Smith <i>et al.</i> [41]	UK Biobank	0.136, s.e. = 0.015	91 370
Okbay <i>et al.</i> [42]	UK Biobank	0.091, s.e. = 0.007	170 910
Lo <i>et al.</i> [38]	23andMe	0.119, s.e. = 0.016	59 206
Luciano <i>et al.</i> [43]	UK Biobank	0.108, s.e. = 0.005	329 821
Nagel <i>et al.</i> [44]	UK Biobank	0.100, s.e. = 0.003	449 484

Box 3. Genome-wide association studies.

GWASs use a hypothesis-free method of identifying specific common variants associated with a phenotype. Analyses are performed for each SNP in the dataset, which can be either genotyped or imputed. This assumes that any causal SNPs will be either captured within the analysis or tagged by those that are included. GWAS performs a regression for each variant, with the appropriate method determined by the format of the phenotype of interest. The regression model includes the genotype at that SNP plus any other relevant confounding variables, such as principal components, in order to adjust for population stratification, or other covariates that could account for variation in the phenotype not acting through the SNP of interest.

To account for the extensive multiple testing performed by running these analyses across the genome, a Bonferroni corrected p -value has been calculated which accounts for the likely number of functional variants being tested. This is the widely used $p < 0.05$ level of 'statistical significance' divided by 1 million, an approximation of the number of independent tests carried out across the genome. The accepted level of genome-wide significance is therefore $p < 5 \times 10^{-8}$ [46]. This is then typically followed by replication in an independent, comparable sample.

the heritability of a trait from GWAS summary statistics [34]. Although these approaches incorporate the LD structure among SNPs, estimates could be biased if the distribution of causal variants is non-random with respect to LD. A potential solution is to stratify SNPs jointly by minor allele frequency and LD. This approach appears to provide unbiased estimates of h^2_{SNP} [14,33].

A recent study by Power & Pluess [35] used the FFM of personality to assess heritability of individual differences in behaviour in a sample of 5011 participants from the 1958 National Child Development Study (NCDS). They applied genome-wide approaches to estimate the heritability of each factor included in this model and found negligible estimates, with the exception of neuroticism ($h^2 = 0.15$, s.e. = 0.08) and openness ($h^2 = 0.21$, s.e. = 0.08). They also identified substantial genetic correlation (r_G) between the neuroticism and extraversion factors ($r_G = 0.82$, s.e. = 0.39) and between neuroticism and openness ($r_G = 1.00$, s.e. = 0.50).

Although previous studies have estimated SNP heritability estimates for various facets of personality [36,37], this was the first study that attempted to estimate heritability for each of the FFM domains simultaneously. The estimates from this study are much smaller than the heritability estimated by twin studies (approximately 40%), leading the authors to suggest that common variants account for only around a quarter of causal genetic variation [35].

A recent meta-analysis by Lo *et al.* [38] estimated heritability based on GWAS summary statistics and found estimates that were consistent for both neuroticism and openness, and an increased heritability estimate for extraversion. Power to detect smaller h^2 estimates was increased in this study, which involved a sample size of approximately 59 000 participants.

Table 1 shows the h^2_{SNP} from neuroticism with increasing sample size. Heritability remains approximately 10%, with increasing precision as the sample size improves.

3. Identifying genetic associations

Despite emerging evidence that individual difference traits are heritable, so far relatively few individual genetic variants associated with these traits have been identified. This is likely due to the polygenic nature of these traits. Under the evolutionary neutral model, most variants with large effects will be rare; the majority of phenotypic variation is likely to be due to common variants of small effect [45]. Initial studies investigating the genetic basis to individual difference traits relied on investigating candidate genes. With advancements in computing power, it has become possible to increase the scope of analyses to take into account the entire genome rather than focusing on a single candidate locus. While candidate gene studies were purely hypothesis-driven and required knowledge of the underlying biology, GWASs are instead hypothesis-generating and test for associations across the whole genome (box 3). Associations found here can inform future studies and provide valuable information on underlying mechanisms.

(a) Candidate gene studies

In a candidate gene study, the association between allele frequency at a particular variant and phenotype of an individual is investigated [47]. These variants are selected *a priori* based on some underlying knowledge of a plausible biological mechanism associated with a particular locus. In recent

years, the associations between the serotonergic and dopaminergic neurotransmitter systems have gained attention [48–50]; however, the evidence for these associations has been somewhat inconsistent.

Although candidate gene studies were once the norm for behavioural genetic studies, many findings failed to replicate leaving the literature awash with false positive results. A recent study investigated findings from the largest schizophrenia GWAS to date and found no evidence that variants in the most-studied candidate genes were more associated with the disorder than non-candidate genes [51]. Candidate gene studies suffered from issues such as small sample sizes and lack of power, confounding due to population stratification which was frequently unaccounted for, and selective reporting of ‘positive’ results [52]. As a result, there has been a move away from these analyses in favour of the more agnostic GWAS approach.

(b) Genome-wide association studies

(i) Analyses using the five-factor model of personality

GWASs have been successful in identifying loci associated with many complex traits, for example a recent GWAS of schizophrenia identified 108 loci robustly associated with the disorder [3], while over 200 susceptibility loci have been identified for inflammatory bowel disease [53]. Progress with regard to personality and temperament traits, however, has been slower.

A number of early GWASs focused on the FFM [39,54–57]; however, these were generally small and the majority of reported associations failed to replicate. Recently, a number of larger studies have used data from the UK Biobank to increase sample sizes [38,41–44]. The largest study to date, which focuses on five broad personality domains, is a recent meta-analysis of several GWASs which included data from 23andMe, Genetics of Personality Consortium, deCODE and UK Biobank ($N = 123\,132$ to $260\,861$) [38]. This study identified six genetic loci, five of which were novel.

The majority of genome-wide analyses of individual differences have been performed in samples of European ancestry. However, studies have also investigated the genetic architecture of these factors within a Korean sample [54,55]. These studies also focused on an FFM of individual differences, but used a version of the NEO-PI revised for use in the Korean population. With the exception of rs2146180, an intergenic variant on chromosome 9, which associated with the openness domain among a sample of young Korean females, there were no robustly associated genetic variants identified among these samples.

(ii) Analyses of specific traits

In addition to studies looking at each of these five domains, other studies have focused on a particular trait. Several studies have focused on neuroticism, with the largest study to date including data from UK Biobank, 23andMe and the Genetics of Personality Consortium ($N = 449\,484$). This study reports finding 136 independent genetic variants associated with neuroticism levels [44]. This suggests that for complex traits, for which there are likely to be many genetic effects of small size, increases in sample size are likely key to identifying genetic effects.

Other studies have focused on extraversion, although the samples for these have been smaller and to date have yielded

few robust findings [58,59]. However, as shown in the heritability literature, there appears to be substantial genetic correlation between the traits of neuroticism and extraversion [35]. Extraversion also appears to be one of the more heritable traits, so it is possible that these analyses are currently underpowered. As of yet, there are no phenotypic measures available in the larger cohorts (such as UK Biobank) that would allow a GWAS of similar magnitude to be carried out.

Excitement-seeking and temperament were also investigated using genome-wide association analyses; however, these also comprised relatively small samples and did not find robust evidence of variants associated with these traits [60–62].

With the availability of large samples such as UK Biobank, there have been recent GWAS successes for other individual difference traits such as depressive symptoms. Progress in identifying genetic variants associated with depression had previously been slow, but in recent years, three large GWASs of major depressive disorder have been published [42,63,64]. The latest of these reports 44 independent risk loci for depression [64]. Reported heritability estimates for depression are similar to those reported for the FFM. It seems feasible that with similar increases in sample size, analyses of behavioural traits will be powered to detect genetic variants of small effect, and we may yet see comparable success in detecting variants that influence individual differences in personality and temperament.

4. Interpreting genome-wide association study findings

With recent advances in computing and the large number of cohorts willing to pool resources, there has been a rapid growth in the number of published GWASs and publicly available summary statistics. These studies have led to the identification of a vast number of genetic variants associated with disease outcomes in addition to health and lifestyle behaviours. Genome-wide association analyses have made a vast contribution to our understanding of the biological mechanisms unpinning many complex traits [1–3,65–67]. However, although GWAS findings often represent direct genetic effects, the possibility that they reflect the effects of modifiable risk factors should not be overlooked [68]. An illustrative example is the nicotinic receptor gene cluster *CHRNA5-A3-B4* on chromosome 15, which is robustly associated with heaviness of smoking. This association was first reported in candidate gene studies [69] and later identified through a GWAS of smoking quantity [70] and studies of nicotine metabolite levels [71–73]. Subsequent functional work suggests that this gene cluster plays a pivotal role in nicotine dependence by reducing the aversive effects of nicotine [74]. While this variant explains some of the biological underpinnings of this trait, the *CHRNA5-A3-B4* locus has also been identified in GWASs of lung cancer [75] and chronic obstructive pulmonary disease [76]. In the case of these traits, it is likely that the observed association with the *CHRNA5-A3-B4* locus is moderated through smoking status and mediated through the resulting tobacco exposure among smokers [68,71].

For individual personality and temperament traits, it may also be possible to identify the effects of modifiable risk factors by dissecting the results of newly published genome-wide analyses. It is likely that any effects, acting either through

biological mechanisms or via modifiable exposures, are small, and large sample sizes will be required to detect these effects. With the recent release of the full UK Biobank sample, in addition to the increasingly collaborative direction in which genetic research appears to be heading, it seems likely that future analyses of these individual difference traits will be undertaken on greater samples, thus increasing power. Despite the fact that samples may not currently be powered to detect variants associated with modifiable risk factors at a genome-wide level, it is possible to interrogate the summary statistics relating to these analyses and investigate for enrichment of variants and networks that are known to relate to exposures of interest.

GWASs of other individual difference traits such as reproductive outcomes (measured by age at first birth and number of children ever born) [77] and educational attainment [78] have been successful in providing insights into the genetic architecture of these other behavioural phenotypes. Meta-analyses of GWASs of both age at first birth and number ever born have been conducted, and several independent loci have been identified that are robustly associated with either one or both of these traits [77]. The related genes have been identified as those that play a role in human reproduction and infertility. In addition, a large GWAS of educational attainment has identified 74 variants associated with years of schooling. These variants are disproportionately found in regions that regulate gene expression in the fetal brain [78]. These studies illustrate that well-powered studies can identify potentially relevant biological pathways, even for proximal phenotypes or behavioural phenotypes that at first appear to be largely environmentally driven.

(a) Genetic overlap: what does this tell us?

Once we have estimated the heritability of a trait, we know that there is a genetic component to this phenotype, although further work needs to be carried out to determine which genes influence this. If heritability has been calculated for multiple traits, then we know that there is a genetic component to each of these; however, we do not know if these are influenced by the same genes. Genetic correlation is a method to quantify how much genetic overlap there is between traits. It provides an estimate of the additive genetic effect shared between pairs of traits. It is possible to estimate genetic correlation using either publicly available summary statistics or individual-level data. The vast number of published GWASs means that effect sizes are available for a wide range of traits on specific variants, and we can use the correlation between these to assess the genetic overlap between traits. Although these do not provide information on the causality of any potential relationship, they shed light on the amount of shared genetic architecture across traits. Any overlap here could be due to pleiotropy (genetic effects on multiple traits), shared biological mechanisms between traits or a causal relationship from one trait to another, but the direction of this cannot be ascertained from these approaches.

There is a well-documented association between personality domains and a range of health behaviours and physical and mental health outcomes [67,79–85]. Recent studies have investigated the extent to which there is shared genetic architecture between personality and temperament traits and both mental and physical health [43,44,83,85]. Nagel *et al.* investigated genetic overlap between neuroticism and several mental health outcomes, anthropometric and health-related traits using previously published summary statistics. Strong

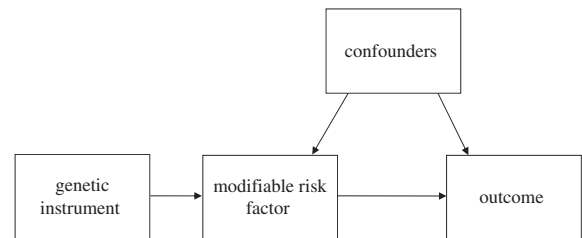


Figure 2. Directed acyclic graph illustrating Mendelian randomization.

evidence of genetic correlation was found for several of these outcomes, with the greatest correlations observed with anxiety, depression and subjective well-being. There was also evidence of moderate–low genetic correlation between neuroticism and a number of other outcomes including schizophrenia, attention deficit hyperactivity disorder (ADHD), anorexia nervosa, educational attainment and height [44].

5. Mendelian randomization

Once genetic correlation between traits has been established, it may be possible to disentangle the nature of the relationship using techniques such as Mendelian randomization (MR). MR is a method of investigating causality using observational data. Genetic variants are used as proxies, or instrumental variables, for a modifiable risk factor of interest. The principles and assumptions underlying MR have been described in detail elsewhere [10,11,86,87]. Assuming these assumptions hold, there should be no association with potential confounders and analyses should not be subject to reverse causation (figure 2). MR can therefore be used to investigate the causality and direction of observational associations where there is a strong genetic instrument for the modifiable risk factor.

For cases where there are strong genetic instruments for both the exposure and outcome, bidirectional MR methods can be used to provide stronger evidence of the direction of effect. Assuming that the instruments are equally as strong for the exposure and outcome, if we were to observe strong evidence of an association in one direction, but not in the other, this would strengthen our confidence that this is the true direction of effect. However, MR alone is not enough to provide definitive evidence of causality, and evidence from different approaches should be triangulated. MR estimates should be compared with those from other methods to investigate whether results are consistent across different approaches and to build a more complete picture of the true effect.

To date, there have been relatively few MR studies carried out using individual differences in personality, in part due to the lack of strong instruments for these traits. However, with the publication of increasingly large GWASs for these traits, we expect instrument strength to increase and MR analyses to become more feasible. MR has provided insights into the causal effects of other individual difference traits, such as educational attainment, depression and anxiety [88,89]. A recent MR suggests that additional education is protective against the risk of coronary heart disease, with similar causal associations with reduced smoking, body mass index and improved blood lipid profiles [88]. MR analyses investigating the effect of smoking behaviours on depression and anxiety have found no strong evidence of an effect in this direction [89].

As mentioned previously, allele frequencies can differ across populations, and this information can be useful when

designing MR analyses. A recent study investigating the role of alcohol use in adolescent internalizing and externalizing problems was performed within a sample of Chinese adolescents [90]. The *ALDH2*2* genotype is found almost exclusively in Asian populations and can be used as a strong instrument for alcohol use in these samples. This study found evidence of a causal role for alcohol in adolescent aggression and attention problems.

In addition, MR can provide insights into evolutionary developmental hypotheses as it allows us to look at varying levels of genetic vulnerability to a trait, rather than focusing on an individual's phenotype. It seems likely that particular traits would be susceptible to selection; for example, individuals with a diagnosis of schizophrenia tend to have lower reproductive success [91,92]. However, the disorder remains at a constant prevalence in the population. This is likely due to favourable characteristics that are linked to some genetic vulnerability for the disorder that, although elevated, is not high enough for the phenotype to manifest.

6. Further research directions

An alternative to increasing sample size is to refine the phenotype being investigated. Reducing the noise in a phenotype measure may lead to a clearer genetic signal, and thus, lower sample numbers would be required to detect this. Attempts to refine the phenotype are useful in terms of improving statistical power, but could limit the utility of any findings if results are not generalizable to the whole population. Therefore, there is a trade-off between statistical power and clinical utility. Ideally, studies combining the large sample sizes of resources, such as UK Biobank, with the detailed phenotyping of smaller cohort studies would provide the greatest power to detect genetic effects, if such effects exist. Longitudinal cohorts have the potential to contribute to this; measurement error will be reduced for longitudinally defined phenotypes, resulting in increased power.

Although there is ongoing discussion about the best model of individual differences in personality, the factor structure suggested by Eysenck and others is a useful starting point when investigating the genetic architecture of these traits [4–7]. Work has been done to investigate the impact of using different personality models when attempting to unpick the genetic structure of these traits, and little evidence of a difference across models has been found [22]. There is

some suggestion that a general factor of personality should be constructed [93,94]. Although there is little work that has been done to specifically investigate the genetic architecture of such a factor [95], the genetic correlation between various factors in the FFM provides some evidence for this theory [35]. Further work should be done to investigate the influence of using different models of these individual differences.

To date, the majority of studies investigating underlying genetics of behavioural traits have focused on samples of European ancestry. Individual molecular genetic effects are likely to be of small size, and large samples will be required to identify them. There is a finite number of samples available with relevant phenotypes and genetic data available; therefore, in order to maximize power and sample size, combining samples across populations is likely to play an important role in future genetic studies. Analytical approaches that can account for underlying population stratification will be required.

7. Conclusion

Despite ongoing debate about the structure of individual differences in personality, there is clear evidence of a genetic component to these traits. Consistent heritability estimates have been calculated for a number of domains using both twin studies and whole genome approaches. Increasingly large GWASs have been published which identify a number of genetic variants for individual differences in behavioural traits. To date, the largest study of a personality trait focuses on neuroticism, which has identified 136 variants of small effect. As the number of samples with both genetic and phenotypic data increases, it is likely that more such evidence will emerge. This has the potential both to provide insights into potential biological mechanisms and to generate instruments for use in future MR analyses investigating causal consequences of these individual differences in personality.

Data accessibility. This article has no additional data.

Author contributions. H.S. drafted the manuscript. G.D.S. and M.R.M. critically revised the manuscript. All authors gave final approval for publication.

Competing interests. We have no competing interests.

Funding. H.S., G.D.S. and M.R.M. work in the Medical Research Council Integrative Epidemiology Unit at the University of Bristol, which is supported by the Medical Research Council and the University of Bristol (MC_UU_12013/1 and MC_UU_12013/6).

References

- Marouli E *et al.* 2017 Rare and low-frequency coding variants alter human adult height. *Nature* **542**, 186–190. (doi:10.1038/nature21039)
- Wood AR *et al.* 2014 Defining the role of common variation in the genomic and biological architecture of adult human height. *Nat. Genet.* **46**, 1173–1186. (doi:10.1038/ng.3097)
- Schizophrenia Working Group of the Psychiatric Genomics Consortium. 2014 Biological insights from 108 schizophrenia-associated genetic loci. *Nature* **511**, 421–427. (doi:10.1038/nature13595)
- McCrae RR, Costa PT. 1985 Comparison of EPI and psychoticism scales with measures of the five-factor model of personality. *Pers. Individ. Differ.* **6**, 587–597. (doi:10.1016/0191-8869(85)90008-X)
- Tupes EC, Christal RE. 1992 Recurrent personality factors based on trait ratings. *J. Pers.* **60**, 225–251. (doi:10.1111/j.1467-6494.1992.tb00973.x)
- Zuckerman M. 1992 What is a basic factor and which factors are basic? Turtles all the way down. *Pers. Individ. Differ.* **13**, 675–681. (doi:10.1016/0191-8869(92)90238-K)
- Eysenck HJ. 1991 Dimensions of personality: 16, 5 or 3? Criteria for a taxonomic paradigm. *Pers. Individ. Differ.* **12**, 773–790. (doi:10.1016/0191-8869(91)90144-Z)
- Cloninger CR. 1994 Temperament and personality. *Curr. Opin. Neurobiol.* **4**, 266–273. (doi:10.1016/0959-4388(94)90083-3)
- Munafò MR. 2009 Behavioural genetics: from variance to DNA. In *The Cambridge Handbook of Personality Psychology* (eds PJ Corr, G Matthews), pp. 287–304. Cambridge University Press. (doi:10.1017/CB09780511596544.021)
- Davey Smith G, Ebrahim S. 2003 'Mendelian randomization': can genetic epidemiology contribute to understanding environmental determinants of disease? *Int. J. Epidemiol.* **32**, 1–22. (doi:10.1093/ije/dyg070)

11. Zheng J, Baird D, Borges M-C, Bowden J, Hemani G, Haycock P, Evans D, Davey Smith G. 2017 Recent developments in Mendelian randomization studies. *Curr. Epidemiol. Rep.* **4**, 330–345. (doi:10.1007/s40471-017-0128-6)
12. Wray NR, Visscher PM. 2008 Estimating trait heritability. *Nat. Educ.* **1**, 29.
13. Tropf FC *et al.* 2017 Hidden heritability due to heterogeneity across seven populations. *Nat. Hum. Behav.* **1**, 757–765. (doi:10.1038/s41562-017-0195-1)
14. Yang J, Zeng J, Goddard ME, Wray NR, Visscher PM. 2017 Concepts, estimation and interpretation of SNP-based heritability. *Nat. Genet.* **49**, 1304–1310. (doi:10.1038/ng.3941)
15. Matthews LJ, Turkheimer E. 2017 Flynn, the ageable method, and a metatheory of intelligence. *Stud. Hist. Philos. Biol. Biomed. Sci.* **65**, 35–40. (doi:10.1016/j.shpsc.2017.07.003)
16. Davey Smith G. 2011 Epidemiology, epigenetics and the 'gloomy prospect': embracing randomness in population health research and practice. *Int. J. Epidemiol.* **40**, 537–562. (doi:10.1093/ije/dyr117)
17. Beam CR, Turkheimer E. 2017 Gene–environment correlation as a source of stability and diversity in development. In *Gene–environment transactions in developmental psychopathology: the role in intervention research* (eds PH Tolan, BL Leventhal), pp. 111–130. Cham, Switzerland: Springer International Publishing.
18. Cesarini D, Visscher PM. 2017 Genetics and educational attainment. *npj Sci. Learn.* **2**, 4. (doi:10.1038/s41539-017-0005-6)
19. Polderman TJC, Benyamin B, de Leeuw CA, Sullivan PF, van Bochoven A, Visscher PM, Posthuma D. 2015 Meta-analysis of the heritability of human traits based on fifty years of twin studies. *Nat. Genet.* **47**, 702–709. (doi:10.1038/ng.3285)
20. Rijdsdijk FV, Sham PC. 2002 Analytic approaches to twin data using structural equation models. *Brief. Bioinform.* **3**, 119–133. (doi:10.1093/bib/3.2.119)
21. Davey Smith G, Relton CL, Brennan P. 2016 Chance, choice and cause in cancer aetiology: individual and population perspectives. *Int. J. Epidemiol.* **45**, 605–613. (doi:10.1093/ije/dyw224)
22. Vukasović T, Bratko D. 2015 Heritability of personality: a meta-analysis of behavior genetic studies. *Psychol. Bull.* **141**, 769–785. (doi:10.1037/bul0000017)
23. Eaves LJ, Heath AC, Neale MC, Hewitt JK, Martin NG. 1998 Sex differences and non-additivity in the effects of genes on personality. *Twin Res.* **1**, 131–137.
24. Fanous A, Gardner CO, Prescott CA, Cancro R, Kendler KS. 2002 Neuroticism, major depression and gender: a population-based twin study. *Psychol. Med.* **32**, 719–728. (doi:10.1017/S003329170200541X)
25. Finkel D, McGue M. 1997 Sex differences and nonadditivity in heritability of the multidimensional personality questionnaire scales. *J. Pers. Soc. Psychol.* **72**, 929–938. (doi:10.1037/0022-3514.72.4.929)
26. Rettew DC, Vink JM, Willemsen G, Doyle A, Hudziak JJ, Boomsma DI. 2006 The genetic architecture of neuroticism in 3301 Dutch adolescent twins as a function of age and sex: a study from the Dutch twin register. *Twin Res. Hum. Genet.* **9**, 24–29. (doi:10.1375/183242706776403028)
27. Lee SH *et al.* 2013 Genetic relationship between five psychiatric disorders estimated from genome-wide SNPs. *Nat. Genet.* **45**, 984–994. (doi:10.1038/ng.2711)
28. Price AL, Patterson NJ, Plenge RM, Weinblatt ME, Shadick NA, Reich D. 2006 Principal components analysis corrects for stratification in genome-wide association studies. *Nat. Genet.* **38**, 904–909. (doi:10.1038/ng1847)
29. Yang J, Lee SH, Goddard ME, Visscher PM. 2013 Genome-wide complex trait analysis (GCTA): methods, data analyses, and interpretations. *Methods Mol. Biol.* **1019**, 215–236. (doi:10.1007/978-1-62703-447-0_9)
30. Yang J, Lee SH, Goddard ME, Visscher PM. 2011 GCTA: a tool for genome-wide complex trait analysis. *Am. J. Hum. Genet.* **88**, 76–82. (doi:10.1016/j.ajhg.2010.11.011)
31. Speed D, Hemani G, Johnson MR, Balding DJ. 2012 Improved heritability estimation from genome-wide SNPs. *Am. J. Hum. Genet.* **91**, 1011–1021. (doi:10.1016/j.ajhg.2012.10.010)
32. Palla L, Dudbridge F. 2015 A fast method that uses polygenic scores to estimate the variance explained by genome-wide marker panels and the proportion of variants affecting a trait. *Am. J. Hum. Genet.* **97**, 250–259. (doi:10.1016/j.ajhg.2015.06.005)
33. Speed D, Cai N, Johnson MR, Nejentsev S, Balding DJ, Balding DJ. 2017 Reevaluation of SNP heritability in complex human traits. *Nat. Genet.* **49**, 986–992. (doi:10.1038/ng.3865)
34. Bulik-Sullivan B *et al.* 2015 An atlas of genetic correlations across human diseases and traits. *Nat. Genet.* **47**, 1236–1241. (doi:10.1038/ng.3406)
35. Power RA, Pluess M. 2015 Heritability estimates of the big five personality traits based on common genetic variants. *Transl. Psychiatry* **5**, e604. (doi:10.1038/tp.2015.96)
36. Vinkhuyzen AAE *et al.* 2012 Common SNPs explain some of the variation in the personality dimensions of neuroticism and extraversion. *Transl. Psychiatry* **2**, e102. (doi:10.1038/tp.2012.27)
37. Verweij KJH *et al.* 2012 Maintenance of genetic variation in human personality: testing evolutionary models by estimating heritability due to common causal variants and investigating the effect of distant inbreeding. *Evolution* **66**, 3238–3251. (doi:10.1111/j.1558-5646.2012.01679.x)
38. Lo M-T *et al.* 2017 Genome-wide analyses for personality traits identify six genomic loci and show correlations with psychiatric disorders. *Nat. Genet.* **49**, 152–156. (doi:10.1038/ng.3736)
39. Bae HT, Sebastiani P, Sun JX, Andersen SL, Daw EW, Terracciano A, Ferrucci L, Perls TT. 2013 Genome-wide association study of personality traits in the long life family study. *Front. Genet.* **4**, 65. (doi:10.3389/fgene.2013.00065)
40. de Moor *et al.* 2015 Meta-analysis of genome-wide association studies for neuroticism, and the polygenic association with major depressive disorder. *JAMA Psychiatry* **72**, 642–650. (doi:10.1001/jamapsychiatry.2015.0554)
41. Smith DJ *et al.* 2016 Genome-wide analysis of over 106 000 individuals identifies 9 neuroticism-associated loci. *Mol. Psychiatry* **21**, 749–757. (doi:10.1038/mp.2016.49)
42. Okbay A *et al.* 2016 Genetic variants associated with subjective well-being, depressive symptoms, and neuroticism identified through genome-wide analyses. *Nat. Genet.* **48**, 624–633. (doi:10.1038/ng.3552)
43. Luciano M *et al.* 2017 Association analysis in over 329,000 individuals identifies 116 independent variants influencing neuroticism. *Nat. Genet.* **50**, 6–11. (doi:10.1038/s41588-017-0013-8)
44. Nagel M *et al.* 2017 GWAS meta-analysis of neuroticism ($N = 449,484$) identifies novel genetic loci and pathways. *bioRxiv* (doi:10.1101/184820)
45. Munafò MR, Flint J. 2014 Common or rare variants for complex traits? *Biol. Psychiatry* **75**, 752–753. (doi:10.1016/j.biopsych.2014.03.010)
46. Clarke GM, Anderson CA, Pettersson FH, Cardon LR, Morris AP, Zondervan KT. 2011 Basic statistical analysis in genetic case-control studies. *Nat. Protoc.* **6**, 121–133. (doi:10.1038/nprot.2010.182)
47. Lohoff FW. 2010 Overview of the genetics of major depressive disorder. *Curr. Psychiatry Rep.* **12**, 539–546. (doi:10.1007/s11920-010-0150-6)
48. Malhotra AK, Virkkunen M, Rooney W, Eggert M, Linnoila M, Goldman D. 1996 The association between the dopamine D4 receptor (D4DR) 16 amino acid repeat polymorphism and novelty seeking. *Mol. Psychiatry* **1**, 388–391.
49. Lesch KP *et al.* 1996 Association of anxiety-related traits with a polymorphism in the serotonin transporter gene regulatory region. *Science* **274**, 1527–1531. (doi:10.1126/science.274.5292.1527)
50. Benjamin J, Li L, Patterson C, Greenberg BD, Murphy DL, Hamer DH. 1996 Population and familial association between the D4 dopamine receptor gene and measures of novelty seeking. *Nat. Genet.* **12**, 81–84. (doi:10.1038/ng0196-81)
51. Johnson EC, Border R, Melroy-Greif WE, de Leeuw CA, Ehringer MA, Keller MC. 2017 No evidence that schizophrenia candidate genes are more associated with schizophrenia than noncandidate genes. *Biol. Psychiatry* **82**, 702–708. (doi:10.1016/j.biopsych.2017.06.033)
52. Colhoun HM, McKeigue PM, Davey Smith G. 2003 Problems of reporting genetic associations with complex outcomes. *Lancet* **361**, 865–872. (doi:10.1016/S0140-6736(03)12715-8)
53. Liu JZ *et al.* 2015 Association analyses identify 38 susceptibility loci for inflammatory bowel disease and highlight shared genetic risk across populations. *Nat. Genet.* **47**, 979–986. (doi:10.1038/ng.3359)
54. Kim H-N, Roh S-J, Sung YA, Chung HW, Lee J-Y, Cho J, Shin H, Kim H-L. 2013 Genome-wide association

- study of the five-factor model of personality in young Korean women. *J. Hum. Genet.* **58**, 667–674. (doi:10.1038/jhg.2013.75)
55. Kim B-H *et al.* 2015 GWA meta-analysis of personality in Korean cohorts. *J. Hum. Genet.* **60**, 455–460. (doi:10.1038/jhg.2015.52)
 56. de Moor MHM *et al.* 2012 Meta-analysis of genome-wide association studies for personality. *Mol. Psychiatry* **17**, 337–349. (doi:10.1038/mp.2010.128)
 57. Terracciano A *et al.* 2010 Genome-wide association scan for five major dimensions of personality. *Mol. Psychiatry* **15**, 647–656. (doi:10.1038/mp.2008.113)
 58. van den Berg SM *et al.* 2016 Meta-analysis of genome-wide association studies for extraversion: findings from the Genetics of Personality Consortium. *Behav. Genet.* **46**, 170–182. (doi:10.1007/s10519-015-9735-5)
 59. Luciano M *et al.* 2012 Genome-wide association uncovers shared genetic effects among personality traits and mood states. *Am. J. Med. Genet. B Neuropsychiatr. Genet.* **159**, 684–695. (doi:10.1002/ajmg.b.32072)
 60. Terracciano A *et al.* 2011 Meta-analysis of genome-wide association studies identifies common variants in *CTNNA2* associated with excitement-seeking. *Transl. Psychiatry* **1**, e49. (doi:10.1038/tp.2011.42)
 61. Service SK *et al.* 2012 A genome-wide meta-analysis of association studies of Cloninger's temperament scales. *Transl. Psychiatry* **2**, e116. (doi:10.1038/tp.2012.37)
 62. Verweij KJH *et al.* 2010 A genome-wide association study of Cloninger's temperament scales: implications for the evolutionary genetics of personality. *Biol. Psychol.* **85**, 306–317. (doi:10.1016/j.biopsycho.2010.07.018)
 63. Hyde CL *et al.* 2016 Identification of 15 genetic loci associated with risk of major depression in individuals of European descent. *Nat. Genet.* **48**, 1031–1036. (doi:10.1038/ng.3623)
 64. Major Depressive Disorder Working Group of the PGC, Wray NR, Sullivan PF. 2017 Genome-wide association analyses identify 44 risk variants and refine the genetic architecture of major depression. *bioRxiv* (doi:10.1101/167577)
 65. Jostins L *et al.* 2012 Host–microbe interactions have shaped the genetic architecture of inflammatory bowel disease. *Nature* **491**, 119–124. (doi:10.1038/nature11582)
 66. Lee JC *et al.* 2017 Genome-wide association study identifies distinct genetic contributions to prognosis and susceptibility in Crohn's disease. *Nat. Genet.* **49**, 262–268. (doi:10.1038/ng.3755)
 67. Sullivan PF *et al.* 2017 Psychiatric genomics: an update and an agenda. *Am. J. Psychiatry* appi.ajp.2017.1. (doi:10.1176/appi.ajp.2017.17030283)
 68. Gage SH, Davey Smith G, Ware JJ, Flint J, Munafò MR, Koifman R. 2016 G = E: what GWAS can tell us about the environment. *PLoS Genet.* **12**, e1005765. (doi:10.1371/journal.pgen.1005765)
 69. Saccone SF *et al.* 2006 Cholinergic nicotinic receptor genes implicated in a nicotine dependence association study targeting 348 candidate genes with 3713 SNPs. *Hum. Mol. Genet.* **16**, 36–49. (doi:10.1093/hmg/ddl438)
 70. Thorgeirsson TE *et al.* 2008 A variant associated with nicotine dependence, lung cancer and peripheral arterial disease. *Nature* **452**, 638–642. (doi:10.1038/nature06846)
 71. Munafò MR *et al.* 2012 Association between genetic variants on chromosome 15q25 locus and objective measures of tobacco exposure. *J. Natl. Cancer Inst.* **104**, 740–748. (doi:10.1093/jnci/djs191)
 72. Le Marchand L, Derby KS, Murphy SE, Hecht SS, Hatsukami D, Carmella SG, Tiirikainen M, Wang H. 2008 Smokers with the *CHRNA* lung cancer-associated variants are exposed to higher levels of nicotine equivalents and a carcinogenic tobacco-specific nitrosamine. *Cancer Res.* **68**, 9137–9140. (doi:10.1158/0008-5472.CAN-08-2271)
 73. Keskitalo K *et al.* 2009 Association of serum cotinine level with a cluster of three nicotinic acetylcholine receptor genes (*CHRNA3/CHRNA5/CHRNA4*) on chromosome 15. *Hum. Mol. Genet.* **18**, 4007–4012. (doi:10.1093/hmg/ddp322)
 74. Fowler CD, Lu Q, Johnson PM, Marks MJ, Kenny PJ. 2011 Habenular $\alpha 5$ nicotinic receptor subunit signalling controls nicotine intake. *Nature* **471**, 597–601. (doi:10.1038/nature09797)
 75. Amos CI *et al.* 2008 Genome-wide association scan of tag SNPs identifies a susceptibility locus for lung cancer at 15q25.1. *Nat. Genet.* **40**, 616–622. (doi:10.1038/ng.109)
 76. Pillai SG *et al.* 2009 A genome-wide association study in chronic obstructive pulmonary disease (COPD): identification of two major susceptibility loci. *PLoS Genet.* **5**, e1000421. (doi:10.1371/journal.pgen.1000421)
 77. Barban N *et al.* 2016 Genome-wide analysis identifies 12 loci influencing human reproductive behavior. *Nat. Genet.* **48**, 1462–1472. (doi:10.1038/ng.3698)
 78. Okbay A *et al.* 2016 Genome-wide association study identifies 74 loci associated with educational attainment. *Nature* **533**, 539–542. (doi:10.1038/nature17671)
 79. Ligthart L, Boomsma DI. 2012 Causes of comorbidity: pleiotropy or causality? Shared genetic and environmental influences on migraine and neuroticism. *Twin Res. Hum. Genet.* **15**, 158–165. (doi:10.1375/twin.15.2.158)
 80. Ormel J, Jeronimus BF, Kotov R, Riese H, Bos EH, Hankin B, Rosmalen JGM, Oldehinkel AJ. 2013 Neuroticism and common mental disorders: meaning and utility of a complex relationship. *Clin. Psychol. Rev.* **33**, 686–697. (doi:10.1016/j.cpr.2013.04.003)
 81. Hakulinen C, Hintsanen M, Munafò MR, Virtanen M, Kivimäki M, Batty GD, Jokela M. 2015 Personality and smoking: individual-participant meta-analysis of nine cohort studies. *Addiction* **110**, 1844–1852. (doi:10.1111/add.13079)
 82. Malouff JM, Thorsteinsson EB, Rooke SE, Schutte NS. 2007 Alcohol involvement and the five-factor model of personality: a meta-analysis. *J. Drug Educ.* **37**, 277–294. (doi:10.2190/DE.37.3.d)
 83. Gale CR *et al.* 2016 Pleiotropy between neuroticism and physical and mental health: findings from 108 038 men and women in UK Biobank. *Transl. Psychiatry* **6**, e791. (doi:10.1038/tp.2016.56)
 84. Lahey BB. 2009 Public health significance of neuroticism. *Am. Psychol.* **64**, 241–256. (doi:10.1037/a0015309)
 85. Purves KL, Coleman JRI, Rayner C, Hettema JM, Deckert J, McIntosh AM, Nicodemus KK, Breen G, Eley TC. 2017 The common genetic architecture of anxiety disorders, *bioRxiv*. (doi:10.1101/203844)
 86. Davey Smith G, Ebrahim S. 2005 What can Mendelian randomisation tell us about modifiable behavioural and environmental exposures? *BMJ* **330**, 1076–1079. (doi:10.1136/bmj.330.7499.1076)
 87. Davey Smith G, Hemani G. 2014 Mendelian randomization: genetic anchors for causal inference in epidemiological studies. *Hum. Mol. Genet.* **23**, R89–R98. (doi:10.1093/hmg/ddu328)
 88. Tillmann T *et al.* 2017 Education and coronary heart disease: a Mendelian randomization study. *BMJ* **358**, j3542. (doi:10.1136/bmj.j3542)
 89. Taylor AE *et al.* 2014 Investigating the possible causal association of smoking with depression and anxiety using Mendelian randomisation meta-analysis: the CARTA consortium. *BMJ Open* **4**, e006141. (doi:10.1136/bmjopen-2014-006141)
 90. Chao M, Li X, McGue M. 2017 The causal role of alcohol use in adolescent externalizing and internalizing problems: a Mendelian randomization study. *Alcohol. Clin. Exp. Res.* **41**, 1953–1960. (doi:10.1111/acer.13493)
 91. Nettle D, Clegg H. 2006 Schizotypy, creativity and mating success in humans. *Proc. R. Soc. B* **273**, 611–615. (doi:10.1098/rspb.2005.3349)
 92. Bundy H, Stahl D, MacCabe JH. 2011 A systematic review and meta-analysis of the fertility of patients with schizophrenia and their unaffected relatives. *Acta Psychiatr. Scand.* **123**, 98–106. (doi:10.1111/j.1600-0447.2010.01623.x)
 93. Rushton JP, Bons TA, Ando J, Hur Y-M, Irwing P, Vernon PA, Petrides KV, Barbaranelli C. 2009 A general factor of personality from multitrait–multimethod data and cross-national twins. *Twin Res. Hum. Genet.* **12**, 356–365. (doi:10.1375/twin.12.4.356)
 94. Hopwood CJ, Wright AGC, Donnellan MB. 2011 Evaluating the evidence for the general factor of personality across multiple inventories. *J. Res. Pers.* **45**, 468–478. (doi:10.1016/j.jrp.2011.06.002)
 95. Veselka L, Schermer JA, Petrides KV, Vernon PA. 2009 Evidence for a heritable general factor of personality in two studies. *Twin Res. Hum. Genet.* **12**, 254–260. (doi:10.1375/twin.12.3.254)