

# Socio-economic status is a social construct with heritable components and genetic consequences

Abdel Abdellaoui<sup>1,✉</sup>, Hilary C. Martin<sup>2</sup>, Martin Kolk<sup>3</sup>, Adam Rutherford<sup>4</sup>, Michael Muthukrishna<sup>5,6</sup>, Felix Tropf<sup>7,8</sup>, Melinda C. Mills<sup>9,10,11</sup>, Brendan Zietsch<sup>12</sup>, Karin J.H. Verweij<sup>1</sup>, Peter M. Visscher<sup>13,14</sup>

<sup>1</sup>Department of Psychiatry, Amsterdam UMC, University of Amsterdam, Amsterdam, the Netherlands

<sup>2</sup>Human Genetics Programme, Wellcome Sanger Institute, Wellcome Genome Campus, Hinxton, UK

<sup>3</sup>Demography Unit, Department of Sociology, Stockholm University, Stockholm, Sweden

<sup>4</sup>Department of Genetics, Evolution and Environment, University College London, London, UK

<sup>5</sup>Department of Psychological and Behavioural Science, London School of Economics and Political Science, London, UK

<sup>6</sup>The Database of Religious History, University of British Columbia, Vancouver, Canada

<sup>7</sup>Centre for Longitudinal Studies, University College London, London, UK

<sup>8</sup>Department of Sociology, Purdue University, West Lafayette, USA

<sup>9</sup>Leverhulme Centre for Demographic Science, Nuffield Department of Population Health and Nuffield College, University of Oxford, Oxford, UK

<sup>10</sup>Department of Economics, Econometrics and Finance, Faculty of Economics and Business, University of Groningen, Groningen, the Netherlands

<sup>11</sup>Department of Genetics, University Medical Centre Groningen, Groningen, the Netherlands

<sup>12</sup>Centre for Psychology and Evolution, School of Psychology, University of Queensland, Brisbane, Australia

<sup>13</sup>Big Data Institute, Li Ka Shing Centre for Health Information and Discovery, Nuffield Department of Population Health, University of Oxford, Oxford, UK

<sup>14</sup>Institute for Molecular Bioscience, University of Queensland, Brisbane, Australia

✉Correspondence: [a.abdellaoui@amsterdamumc.nl](mailto:a.abdellaoui@amsterdamumc.nl)

## ABSTRACT

**In civilizations, individuals are born into or sorted into different levels of socio-economic status (SES). SES clusters in families and geographically, and is robustly associated with genetic effects. Here, we first review the history of scientific research on the relationship between SES and heredity. We then discuss recent findings in genomics research in light of the hypothesis that SES is a dynamic social construct that involves genetically influenced traits that help in achieving or retaining a socio-economic position, and can affect the distribution of genes associated with such traits. Social stratification results in people with differing traits being sorted into strata with different environmental exposures, which can result in evolutionary selection pressures through differences in mortality, reproduction, and non-random mating. Genomics research is revealing previously concealed genetic consequences of the way society is organized, yielding insights that should be approached with caution in pursuit of a fair and functional society.**

Human societies throughout history have often been stratified by socio-economic status (SES), with different groups of people having access to different levels of power, prestige, wealth, health, freedom, and overall quality of life (see Box 1).<sup>1</sup> While some believe that, in a meritocratic system, inequality can serve as an incentive for individuals to be more productive, inequality – or an excess thereof – is also considered disruptive, with detrimental effects on social cohesion.<sup>2-4</sup> An increase in social inequality is generally accompanied by growing disparities in mental and physical health.<sup>4,5</sup> While in recent times, inequality between countries has broadly declined, inequality within many countries has been increasing, especially since the 1980s.<sup>6,7</sup> Social inequality is an inherently societal phenomenon driven by cultural, structural, economic, political, and technological forces, although, as we show here, it is also associated with genetic variation. While behavioral genetics research is actively examining the

relationship between genetics and SES,<sup>8</sup> most studies within the broader social sciences aiming to understand social inequalities tend to focus on societal factors.<sup>6,9</sup> By not including genetic effects, a significantly contributing force is omitted that may be increasing in importance due to recent societal changes. When acknowledging these genetic effects, however, it is important to tread with caution. Recent history has shown that attempts to control the genetic make-up of populations – in the form of eugenics – can result in serious violations of human rights, including limiting access to education and labor markets, involuntary sterilization, infanticide, and genocide.<sup>10-12</sup>

In the first half of this article, we review the history of social stratification and the scientific study of its relationship with DNA. We summarize recent developments in genomics research that have provided us with a wealth of data on the relationships between genetic effects and socio-economic outcomes, albeit overwhelmingly in populations of European ancestry and with a bias towards individuals with higher SES (Box 2). In the second half of this article, we discuss how these new data could be interpreted in the context of SES as a dynamic social construct that could exert natural selection pressures on genes associated with socially advantageous traits.

## **HISTORY OF SOCIAL STRATIFICATION AND HEREDITY**

### *Social Stratification*

Civilizations are generally defined as complex societies with urban development, some form of government, symbolic systems of communication (e.g., writing), and social stratification.<sup>13,14</sup> While there are no known civilizations without some form of social stratification, it is not a defining characteristic of human societies in general. Most of the hunter-gatherer societies known today are relatively egalitarian.<sup>1,15,16</sup> Social stratification became more pronounced with the rise of larger and more complex human societies that arose during developments of the Neolithic era starting about 12,000 years ago.<sup>17</sup> Broadly speaking, the gradual shift from hunter-gatherer to sedentary agricultural societies enabled surplus resource accumulation, which led to an increase in population size and division of labor.<sup>18</sup> This allowed for different levels of prestige to develop through job specialization and more unequal accumulation of possessions. In numerous societies, elite classes arose that gained control over food supplies, land, means of production, and the labor of much of their population. In many instances, legal and structural systems developed over time that codified these social hierarchies, reinforcing the power and privilege of the elites and cementing the stratification within societies.<sup>19</sup>

Human history has known relatively rigid social stratification systems with little movement between socio-economic levels, often maintained through religious beliefs that legitimized the divine mandate of rulers.<sup>20</sup> Social status was often ascribed, with children inheriting their parents' status. Many of the phenomena discussed in this article apply to systems that show at least some degree of merit-based social mobility. The earliest recorded example of a formal merit-based system emerged in 6<sup>th</sup> century BC China, where Confucian scholars advocated education for all, and introduced the notion that those who govern should do so based on individual merit rather than inherited status.<sup>21</sup> Such meritocratic principles were later applied by Genghis Khan in the 13<sup>th</sup> century, selecting leaders based on ability rather than family.<sup>22</sup> European scholars translated Confucian texts in the 17<sup>th</sup> century, exposing them to alternative perspectives on governance and social organization. These Confucian ideals probably contributed to the intellectual milieu of the Enlightenment movement, where merit-based social systems gained prominence.<sup>23</sup> The medieval European estate system, where noble or common status was largely determined by birth, made way for socio-economic orders that aimed for more equal opportunities. As

the Industrial Revolution unfolded, bringing increased production, economic growth, and social change, a modern, more merit-based socio-economic system began to emerge, marking the transition towards a new social order that could accommodate an ever-expanding population, while also increasing a visible underclass.

Compared to many pre-industrial socio-economic orders, merit-based hierarchies increase opportunities across the population, allocate talent more efficiently, and stimulate progress through competition between people and between firms. The term ‘meritocracy’, however, was originally coined in a negative light in the 1958 satire “The Rise of the Meritocracy” by Michael Young.<sup>24</sup> This book describes a dystopian future, in which meritocracy has led to a newly stratified society, replacing an aristocracy of birth by an aristocracy of talent, with a disenfranchised lower class of the less meritorious. If behaviors associated with merit (e.g., intelligence, persistence, creative talent) are partly heritable, variation in genetics within families could still facilitate social mobility. Enduring accumulation of resources within families however could limit this mobility, gradually reverting meritocracy back towards an aristocracy of birth.

### *Heredity*

Contemporary research shows SES to be central to social stratification, focusing on intergenerational transmission of education,<sup>25</sup> occupation,<sup>26</sup> class,<sup>27</sup> earnings, and wealth,<sup>28</sup> and variations across countries, history, gender, and ethnicity. Social science research on social stratification and intergenerational transmission of SES has largely ignored or actively resisted the study of its relationship with genetic factors, partly due to ethical concerns and historical misuse of genetics in social policy.<sup>29</sup> To better understand this oversight, it is important to consider the history of genetics research and its societal impact.

Scientific research exploring connections between genetics and socio-economic success has a turbulent and controversial history. During the 16<sup>th</sup> century, early ideas about biological heredity were influenced by legal concepts of cross-generational inheritance of property and wealth.<sup>30</sup> The concept of heritability began to be formalized in the 19th and early 20th centuries in the light of the work of Mendel and Darwin, whose work revealed the laws of inheritance and mechanisms of evolution. Charles Darwin’s half-cousin, Francis Galton, explored the heritability of traits linked to merit and socio-economic success in his book, *Hereditary Genius* (1869). In this period, a prelude to the emergence of the field of genetics, Galton and his followers put more emphasis on ‘nature’ than nurture. In his book, Galton applied statistics to show that offspring of “eminent” figures had a higher chance of succeeding in what were perceived to be high-profile professions.<sup>31</sup> Inspired by these findings, Galton became a proponent of improving what became known as the ‘genetic quality’ of a population through selective parenthood, thus initiating and spearheading the emerging eugenics movement.<sup>32</sup> This movement became widely supported in many countries across the world and across the political spectrum by established intellectuals and medical authorities.<sup>10</sup> Eugenics proponents intended to explore and enact policies that would increase the overall wellbeing of majority populations or dominant social groups, but inevitably at the expense of others who were deemed economically costly or socially undesirable, and suffered stigmatization and persecution as a result.<sup>11</sup> In many cases, eugenic ideas resulted in state-sponsored violence against marginalized groups, primarily via enforced or coerced sterilisation.<sup>33-35</sup> The destructive power of the eugenics movement reached genocidal levels in the Second World War, after which its public support declined. The legacy of involuntary sterilization is still detectable, with population register data revealing that individuals categorized with severe mental and physical disabilities (up to 1970 in Finland and 1976 in Sweden) often remained childless.<sup>36</sup> Enforced or coercive

sterilization continues in several countries to this day, including China and India, the two most populous countries on earth, often targeting lower socio-economic groups as a means of population control.<sup>11</sup> In the second half of the 20<sup>th</sup> century, the scientific field of heredity did become largely decoupled from social applications in most countries and made progress through decades of twin and family studies.<sup>37</sup>

When it comes to socio-economic success, merit in contemporary industrialized societies typically involves strong performance in the educational system and/or labor market, both of which have intelligence – defined and measured in various ways – as one of their strongest predictors, alongside non-cognitive predictors such as parent's socioeconomic status and individual-level traits such as conscientiousness.<sup>38-42</sup> Intelligence was the first trait studied using the classical twin method,<sup>43,44</sup> which estimates heritability by comparing the resemblance of identical and fraternal twins. The considerable heritability of intelligence, i.e., the extent to which genetic differences explain individual differences within a population, and its increase from childhood (~.43) to adulthood (~.65), have become among the most replicated findings in twin research.<sup>37,45,46</sup> A change in heritability can occur because genetic influences change over time or, more likely, because the variance in or influence of environments change.

Twin and family studies suggest that the heritability of traits affecting socio-economic outcomes varies with societal equality. Theoretically, equalizing opportunities could lead to reduced heritability by diminishing the impact of genetic predispositions through increased access to resources and support systems. Alternatively, heritability might increase if environmental influences become more uniform across the population, allowing genetic differences to explain a larger proportion of the remaining variation. In welfare states with more equal opportunities and high intergenerational mobility, the heritability of education seemed to increase, primarily due to reduced environmental influences and a relative constancy of genetic influences, increasing the relative contribution of genetics to the total variance.<sup>47</sup> In Europe, the heritability of educational attainment for women grew from near zero for those born around 1900 to almost 60% by the century's end (Figure 1),<sup>48</sup> possibly reflecting improving educational systems and more equal opportunities.

### *The Genomics Era*

Advances in genotyping technologies in the 21st century enabled genome-wide association studies (GWASs), first applied in 2005 after being proposed in 1996.<sup>49-52</sup> In a GWAS, millions of genetic variants capturing most common genetic variation in a population are measured, and the effect of each variant on a trait of interest is estimated. The effects of individual genetic variants on complex traits turned out to be hard to distinguish from noise, due to their extremely small effect sizes and the heavy multiple testing burden. By pooling data from many different cohorts, GWAS consortia reached large sample sizes, now on the order of millions of participants for some traits, allowing GWASs to successfully identify tens of thousands of variants associated with hundreds of traits related to physical and mental health as well as behavior.<sup>53</sup>

The first large-scale GWAS on educational attainment was published in 2013,<sup>54</sup> conducted in ~125 thousand individuals, identifying only three associated genetic variants explaining only ~0.02% of individual differences. A small harvest for such a large study, but enough to prompt further increases in sample size. Subsequent larger-scale GWASs reached up to 3 million individuals, identifying thousands of significantly associated variants.<sup>55-57</sup> GWASs on income and occupational status, each with sample sizes in the hundreds of thousands, found hundreds of associations, with very similar polygenic signals across these SES indicators.<sup>58-60</sup> The signal of income GWASs is consistent across Western countries, including the UK, Netherlands, Norway, Estonia, and the US (average  $r_g = 0.88$ ),<sup>59</sup> indicating minimal



confounding due to population stratification (Box 2). Many of the studies discussed below are based on polygenic scores constructed from the educational attainment GWASs, as these currently provide the most predictive scores due to their statistical power. Recently, the first GWAS on educational attainment was conducted in a large East Asian dataset – South Koreans and Taiwanese – which showed a remarkably similar signal to the European GWASs ( $r_g = 0.87$ ),<sup>61</sup> again suggesting limited confounding due to population stratification.

As with many behavioral traits, increasing sample sizes reveal progressively smaller effect sizes, in line with the infinitesimal model – a concept that builds upon the early foundations in quantitative genetics conceptualized by R.A. Fisher a century earlier.<sup>62</sup> Modern GWASs confirm that complex traits are influenced by a large number of genetic variants, each with a small effect,<sup>63,64</sup> but collectively explaining substantial variance. One way to harness the power of the aggregate genetic effects is through so-called polygenic scores, in which the alleles an individual carries are weighted by their estimated effects on a trait and then summed to produce genetic predictors. Polygenic scores for educational attainment from the latest GWAS of 3 million individuals explain 12-16% of educational differences in Europeans,<sup>57</sup> of which about half is expected to be due to the clustering of economic resources in families and assortative mating (Box 3).<sup>65</sup> Larger GWASs are expected to improve precision of genetic effect estimates and thereby the predictive power of polygenic scores, although this depends on similarities between the discovery GWAS and the target datasets.<sup>66</sup> These polygenic scores capture part of the heritability of socio-economic outcomes, reflecting biological, social, and demographic processes, and correlated environmental exposures (Box 3). Genetic scores computed from birth with predictive power on future socio-economic outcomes hold potential value for research as well as policy development. Currently, however, their predictive power arises from a largely elusive combination of underlying traits and environmental influences.<sup>67</sup> Analyses of GWAS signals for SES outcomes using enrichment tools applied to diseases show the strongest enrichment in brain and neuronal processes,<sup>56,59</sup> consistent with the role for cognitive and behavioral traits in SES. Rather than solely striving to expose underlying biology and perfect genetic effect size estimates, the field has great potential when striving to understand what causes genetic effects to vary across different environments and populations—a direction we explore in the following sections on socio-economic status as a selection pressure.

## **SOCIAL STRATIFICATION AS A SELECTION PRESSURE**

When Charles Darwin presented his theory of evolution by natural selection,<sup>68</sup> he wrote that nature selects for adaptations that give organisms an advantage in the three struggles of life: with the physical environment, with other species, and with members of one's own species. Over time, humans probably reduced the first two selection pressures while intensifying competition within populations, where winners are rewarded with the favorable social and environmental circumstances that come with higher SES.

GWASs on socio-economic outcomes produce polygenic signals that contain genetic effects on a mixture of traits and environmental effects.<sup>67</sup> Here, we describe social forces that bind these traits and environmental effects together. In human societies, people become sorted into more favorable or less favorable environments, depending on a combination of their inherited privileges and their performance in the socio-economic system. This sorting could create selection pressures through differential mortality, reproductive success, and mate choice. SES is a dynamic social construct which could target different associated traits and genes across time and space. In more merit-based systems, positive selection would probably act on genes associated with attributes and skills deemed beneficial by that

particular society, assuming these translate to greater reproductive success; however, as we explain further below in the section *Reproductive Success*, in practice, the direction of the selection effect may have varied over time, and there can be non-linear effects.<sup>69</sup> In particular, formal education and the introduction of money as a standardized medium of exchange could be expected to have sharpened this selection pressure. An educational system is a relatively efficient way to nurture talents and stratify the population based on those talents at a relatively early age, while money offers a more efficient way to keep score of a person's "earned" SES compared to earlier barter systems or reliance on resources like land.

Historically, varying environmental and social conditions impacted mortality and reproductive success, leading to complex patterns in the genetic architecture of complex traits. The highly polygenic and pleiotropic nature of complex traits, in addition to complex population structures, complicate detection of past selection pressures at the DNA level.<sup>70</sup> Selection pressures on complex traits are dispersed across many minor effects. The most pervasive selection pressures detectable in genetic data across most complex traits, including educational attainment, are negative selection pressures, indicated by genetic variants with larger effects being kept at lower frequencies.<sup>71,72</sup> This pattern reflects a consistent constraint on extreme phenotypes (stabilizing selection), flattening the distribution of effect sizes and resulting in high polygenicity.<sup>73</sup> Thousands of genes show a strong signature of historical negative selection against damaging variants,<sup>74-76</sup> and in contemporary populations, rare damaging variants in these genes have been associated with lower intelligence and educational attainment.<sup>77-79</sup> Modern selection pressures are expected to be detectable through associations between polygenic effects and patterns of mortality, reproduction, and non-random mating, all of which appear to be strongly driven by SES-related outcomes that vary across time and space, as discussed in more detail below.

#### *Genes, Geography, and Mortality*

After the advent of agriculture, societies became larger and more complex.<sup>80</sup> Over the millennia of the Neolithic era, population growth led to urbanized settlements that covered geographic regions orders of magnitude larger than their predecessors. Different populations covered different geographical areas, and within populations and cities, socio-economic strata covered different regions.<sup>81</sup> The largest patterns of genetic variation align with differences between ancestral populations, which correlate strongly with geography, because individuals tend to reproduce with people who live closer to them geographically (see *Population Stratification* in Box 2).<sup>82-84</sup> If socio-economic outcomes are based on heritable merit-based outcomes, such as the performance in an educational system, it would be expected that the genes associated with these outcomes would show regional differences within populations as well.

Analyses of the geographic distribution of polygenic scores for a variety of behavioral, cognitive, and health-related outcomes across Great Britain and across Estonia revealed that, once ancestry differences were minimized, the strongest regional differences were in educational attainment polygenic scores.<sup>85,86</sup> These geographic differences aligned with regional socio-economic differences; in Great Britain, lower polygenic scores clustered in economically disadvantaged coal mining areas,<sup>85</sup> and in Estonia, higher polygenic scores were more concentrated in the two prospering university towns.<sup>86</sup> Migration contributed to these differences, as individuals with higher polygenic scores tended to move to more prosperous regions, increasing regional differences throughout the 20th century.<sup>85,86</sup>

Polygenic scores for educational attainment and other SES outcomes are derived from GWASs estimating genetic effects on traits relevant to contemporary socio-economic systems. We can get a sense of the contribution of the underlying traits by estimating genetic correlations ( $r_g$ ) between traits,

which vary between -1 (100% shared variance due to the same genetic effects in the opposite direction) through 0 (no overlap in genetic effects) to 1 (100% of shared variance due to the same genetic effects in the same direction). The genetic correlation between educational attainment, income, and occupational status are ~0.9 in many developed countries, suggesting that mostly the same genetic signal is being picked up for all three traits, namely that of socio-economic status in those specific societies.<sup>58-</sup>

<sup>61</sup> Of all traits investigated so far, intelligence shows the highest genetic correlation with socio-economic status ( $r_g$  with educational attainment, income, and occupational status ~0.7); personality dimensions and mental and physical health outcomes also share a significant portion of their genetic effects with these SES outcomes.<sup>56,58-60</sup> Polygenic scores of traits that contribute more to socio-economic success tend to show stronger regional differences in Great Britain (Figure 2) and in Estonia than traits that are not.<sup>85,86</sup>

Genetic effects associated with educational attainment, income, or occupational status are a patchwork of many underlying heritable outcomes, but these effects, as estimated in a GWAS, are also intertwined with environmental effects. Environmental factors are diverse and include cultural, social, economic, and geographical contexts, ranging from societal structures like housing quality, dietary options, healthcare, and education systems to natural conditions such as air pollution. The increased efficiency with which populations in modern societies are stratified according to these heritable outcomes into different layers of environmental exposures leads to correlations between genes and environment that result in both ‘double advantages’ and ‘double disadvantages’. As a result, conducting a standard population-based GWAS on an SES-related outcome is partially equivalent to doing a GWAS on being born into a better environment and/or the ability to move to a better environment. Molecular genetic evidence shows that SES-related genetic effects and environmental influences cluster on both a family-level<sup>87,88</sup> and regional level<sup>85,89</sup>, resulting in systematic differences in these environmental exposures. These environmental factors that correlate with genetics can cause additional (regional) differences in mental and physical health outcomes, such as substance use and BMI,<sup>89</sup> but also more heritable traits, such as height (heritability ~80%; Figure 3).

People with genetic variants that make it easier for them to get a better education are more likely to move to better neighborhoods, while the people left behind are in worse living circumstances with higher mortality rates and greater risk for health problems like obesity, diabetes,<sup>85</sup> and infectious diseases. Regional differences in Covid-19 infection and mortality rates, for instance, also show significant genetic correlations with SES (Figure 4), as it was easier for people in certain occupations, smaller households, and better housing to avoid the disease by working from home and to be physically better prepared for infection through healthy diet and exercise.<sup>90</sup> Covid-19 was not likely to exert a strong selection pressure, because mortality was only high among those past the reproductive age. Previous pandemics, however, with higher mortality rates for younger people, such as the black death, smallpox, and the Spanish flu, had higher mortality in areas of lower SES as well.<sup>91</sup> Besides the typical health consequences associated with lower-quality living conditions, deadly pandemics – whose frequency has increased since the advent of agriculture and cities<sup>92,93</sup> – could potentially impact allele frequencies of genetic variants that affect traits that are more consistently associated with social stratification.

### *Reproductive Success*

Natural selection affects allele frequencies through differences in reproductive success, influenced by differential mortality rates, fertility rates, and mate choice. Like other large mammals, human populations have historically faced density-dependent checks, where resource availability affected

population density through mortality and fertility rates.<sup>94</sup> In humans, these dynamics are likely intertwined with socio-economic status (SES).

A collection of ~15,000 British wills from the 16<sup>th</sup> to the 20<sup>th</sup> century showed a positive relationship between income and net fertility in pre-Industrial centuries, with the wealthiest leaving nearly twice as many offspring as the poorest.<sup>95,96</sup> This was probably influenced by higher child mortality rate in lower SES groups,<sup>96,97</sup> and greater mating opportunities for higher SES males, as women tend to prefer men with more resources across cultures with different mating systems, different levels of gender equality, and different religious orientations.<sup>98</sup> Pre-Industrial data from the 13<sup>th</sup> to the 21<sup>st</sup> century across multiple countries confirm this positive income-fertility relationship.<sup>99</sup> As societies underwent significant transformations due to the Industrial Revolution, including changes in population density, urbanization, and industrialization, a general shift was observed across the world from a positive relationship to a negative relationship between income and reproductive success.<sup>99</sup> Several explanations posed for this reversal include changes in child mortality, birth control, and women's education and workforce participation.<sup>99-101</sup> Studies on industrialized societies have also shown sex differences in the association between wealth and fertility.<sup>102</sup>

In contemporary Western populations, including Great Britain, common genetic variants associated with higher SES show a negative correlation with offspring count.<sup>69,103-106</sup> This implies a recent decline of these variants, despite their relationship with decreased mortality described in the previous paragraph. In contrast, rare genetic variants with stronger deleterious effects on intelligence, educational attainment, and income, negatively impact reproduction rates in Great Britain, especially in men.<sup>77</sup> This may reflect a non-linear relationship between intelligence and reproductive success, with rare damaging variants being more predictive at the lower end of the intelligence spectrum.

More recently, the relationship between SES and reproductive success, which seems to be non-linear and sex-specific, seems to be reverting back to an overall positive one, both within and between a number of high-income countries; this is likely driven by multiple factors, such as an increased compatibility between women's careers and families due to more favorable family policies and social norms, cooperative fathers, and more flexible labor markets.<sup>107-109</sup>

### *Assortative Mating*

Besides determining who propagates their genes, mate choice can also impact genetic variation by inducing spouse similarities. If people systematically choose mates that resemble them for a certain trait, a phenomenon known as assortative mating, this affects genetic variation by widening the genetic and thereby phenotypic distribution in the offspring population and increasing resemblance between family members.<sup>62,110</sup> Humans tend to meet and choose partners similar to themselves in terms of ethnicity, religion, and socio-economic status.<sup>111,112</sup> The strongest DNA similarity in spouses is for ancestry-related variation,<sup>113</sup> followed by polygenic effects associated with educational attainment, which show the highest assortative mating levels among studied traits so far.<sup>114-117</sup> Interestingly, the educational attainment associated loci show a higher correlation between spouses than expected based on the phenotypic spousal correlations. Possible explanations include additional assortative mating on the basis of more heritable underlying traits (e.g., intelligence), matching based on characteristics of both the mate and their family members, or inaccurate genetic effect estimates (see Boxes 2 and 3).<sup>116,118,119</sup>

While assortative mating on education may have been strengthened more recently by the increased heritability of educational attainment in women (Figure 1) and more women joining the workforce,<sup>120</sup> studies into partnership markets around the Industrial revolution suggest that assortative mating on

socio-economic outcomes is not a recent phenomenon.<sup>121,122</sup> Data on 422,374 British inhabitants from 1600 to 2022 show substantial assortative mating and persistence of social status across generations despite significant social changes over time.<sup>123</sup> In Norway, assortative mating continues to increase genetic similarities within families for genetic variants associated with educational attainment, and thus has not yet reached equilibrium.<sup>117</sup> Equilibrium, where genetic similarities within families stabilize despite ongoing assortative mating, is expected to be achieved after many generations of stable assortative mating on this outcome. The ongoing increase in familial genetic similarity suggests a recent increase in assortative mating, potentially contributing to growing inequalities in contemporary Norwegian society.

Assortative mating probably increases with the geographic clustering of SES, since physical proximity increases the chance of finding a similar partner. SES has been clustering geographically since ancient times.<sup>81</sup> Distances traveled with SES-related migration increased in recent times, which may have increased geographic clustering of SES-associated alleles,<sup>85</sup> increasing assortative mating. If current rates of assortative mating on traits that influence SES persist or increase, this could further increase social inequalities on both an economic and genetic level over generations, making them harder to overcome. Assortative mating can also make economic factors become more intertwined with genetics over generations, as environmental advantages tied to SES influence mate selection. In both Great Britain and Norway, for example, it has been shown that earlier born siblings, who have a higher SES due to environmental factors, marry spouses with higher polygenic scores for educational attainment.<sup>124</sup> The increasing correlation within individuals between SES-associated alleles across the genome, and between these alleles and environmental factors that influence SES, could have societal and evolutionary consequences, but also complicate the task of accurately quantifying genetic effects associated with SES (see Box 3).

## CONCLUSIONS

Social inequality has long been inherent in the way human societies are organized, arising because certain outcomes are more valued and rewarded than others, and reinforced by the familial clustering and transmission of status, resources, and genetic predispositions. Populations become stratified into social environments with differing levels of health risks, safety, and opportunity, leading to disparities in mental and physical health. When combining significant life quality differences between social layers with a certain amount of social mobility, it becomes desirable to climb the social ladder, stimulating many to try, but allowing only those with the most advantageous talents to succeed. Over time, this could influence the genetic make-up of populations through differential mortality, fertility, and non-random mating.

The strength and nature of these selection pressures vary across time and cultural contexts. Certain cognitive abilities may have conferred a more consistent advantage throughout our recent evolutionary history, particularly in more merit-based societies, which could have contributed to the increasingly complex societies that make humans such a unique primate. Technological advances may have impacted the effects of SES on genetic variation in multiple ways, including through improved and more accessible educational systems. These developments may, as a side-effect, have made humans more efficient at stratifying the population according to genetic talent, further inducing geographic clustering and assortative mating on SES-related genetic variants, potentially increasing genetic differences associated with social inequality. The more direct genetic effects, however, are significantly smaller

than initially estimated in population-based genetic association studies, due to assortative mating and environmental influences that get entangled with genetic effects (see Box 3).

At the heart of these discussions is not a call for genetic intervention, but rather a call for a deeper understanding and awareness that our social structures are part of an evolving environment that, over time, shapes both social and genetic outcomes, potentially influencing human evolution.<sup>125,126</sup> The relationship between social stratification and genetic effects is complicated, consisting of a network of complex traits and environmental circumstances woven together by social and economic forces created by increasingly complex societies. While precise estimates of genetic effects are important, this field of research holds greater potential in uncovering how correlations between genes and social outcomes shift across societies and time.<sup>47,48,127</sup> Processes like migration, mate choice, reproductive success, and mortality shape how populations develop and could create feedback loops that reinforce or reshape inequalities. Understanding these dynamics could help us trace societal structures of the past through traces left in the genome by, for example, assortative mating,<sup>128</sup> reveal causes of present inequalities, and anticipate how inequalities might evolve with future social changes. While economic inequalities could be more readily mitigated through targeted policy changes, genetic influences (if causal) may persist across generations, presenting longer-term challenges. This line of research can provide researchers across disciplines with a framework for studying the dynamic interplay between genetics, complex traits, and social organization. While mistakes from the past should keep us vigilant about the potential for harmful effects of genetically informed social policies, advances in interdisciplinary genomics research can help us better understand the processes underlying the way societies are organized and their consequences.

### Box 1: Definition of Socio-Economic Concepts

- **Social Stratification:** The hierarchical organization of societies based on access to resources, power, and socio-economic status (SES), ranging from high to low. Group membership may or may not persist across generations.
- **Socio-Economic Status:** A social construct reflecting how society assigns value to certain outcomes, skills, traits, behaviors, achievements, and assets. SES captures and influences cultural norms within a specific society. In social science research, SES is typically described across four correlated dimensions:
  - *Income:* Total earnings from all sources (e.g., wages, bonuses, investments, rental income), typically averaged over multiple years and measured at the individual or household level.
  - *Education:* The formal acquisition of knowledge and skills through schooling, training, or self-directed learning. Education is both a cause and consequence of cognitive ability. It is often measured by years of attainment, earned credentials, and standardized test scores.
  - *Occupation:* Employment status, which can fluctuate over adulthood. Typically measured through occupational classes, continuous status scales, or ranking systems that combine education and occupational status (e.g., ISEI, SIOPS, CAMSIS)<sup>58</sup>.
  - *Wealth:* Net worth, defined as value of assets (e.g., property, financial holdings, possessions) minus liabilities. Wealth includes both self-generated and intergenerational components and is particularly difficult to measure at the top of the distribution.
- **(Intergenerational) Social Mobility:** The degree to which an individual's SES rank correlates with that of their parents, reflecting changes in SES across generations.
- **Social Inequality:** The unequal distribution of resources, opportunities, and quality of life across individuals or households within a society at a given time.

425

426

427

428

## Box 2: Factors Affecting Estimation of Genetic Effects on Socio-Economic Outcomes (Part I)

### Sources of bias in estimating genetic effects

Genome-wide association studies (GWASs) can help illuminate genetic architectures of complex traits, but can also present a somewhat distorted view of this genetic architecture, particularly for behavioral and socio-economic outcomes. The three main sources of bias are:

- **Population stratification:** The largest patterns of genetic variation within a population generally reflect ancestry differences, often from a more distant past. When cultural or environmental differences align with these genome-wide allele frequency differences, GWASs can produce false positive associations across the genome. While controlling for ancestry differences reduces bias, it is difficult to eliminate entirely.<sup>129</sup> Furthermore, controlling for ancestry could also obscure real genetic signals when the causal genetic variants themselves correlate with ancestry differences. A more effective approach to control for population stratification involves within-family analyses, where genetic differences between siblings are compared,<sup>88</sup> though these tend to, currently, have much smaller sample sizes than population-based GWASs.
- **Ancestry-related ascertainment bias:** GWAS datasets lack global diversity,<sup>130</sup> with 72% of GWAS discoveries (2005 – 2018) originating from the US, UK and Iceland.<sup>131</sup> To minimize false positives from population stratification, GWASs often exclude ethnic minorities and control for small ancestry differences. This narrows the scope of GWASs on SES-related outcomes. Besides missing how SES interacts with genetics in non-Western societies, they fail to capture influential social factors within Western societies such as systemic racism or discrimination experienced by minority groups.
- **SES-related ascertainment bias:** People who are willing and able to participate in genetics research tend to be more healthy and have higher SES, introducing a sampling bias.<sup>132</sup> UK Biobank, for example, shows smaller regional SES differences compared to census data.<sup>85</sup> This bias can distort our estimates of genetic and environmental contributions to complex traits across the full SES spectrum. Additionally, collider bias may arise when participation is influenced by both genetic factors and SES, potentially distorting associations between them. Using population-based weights to adjust for the overrepresentation of healthier, wealthier participants can help correct some of the biases.<sup>133</sup>

430

431

432



### Box 3: Factors Affecting Estimation of Genetic Effects on Socio-Economic Outcomes (Part II)

#### Inflation of genetic effect estimates due to socio-economic structures

A substantial part of socio-economic status (SES) is inherited through the family, household, and community one is born into, intertwining environmental and genetic influences. This can inflate genetic effect estimates in population-based GWASs. Polygenic scores for educational attainment based on population-based GWASs predict outcomes twice as strongly in children raised by biological families compared to adoptees,<sup>134</sup> and ~1.6 times more between families than within families<sup>135</sup>. Conducting GWASs within families<sup>88</sup> or geographic regions<sup>89</sup> significantly reduces the magnitude of genetic effects on SES-related outcomes. The three main sources of inflation of genetic effects on SES as estimated in population-based GWASs are:

- **Indirect genetic effects:** These occur when genetic influences arise not from an individual's own genes, but from the shared loci of family members. These effects can arise from parents (e.g., through parental investment or the prenatal environment), extended family members, or even multi-generational/dynastic effects (e.g., great-grandparents or beyond)<sup>136</sup>. Parents pass down half of their genes to their offspring. Polygenic scores for educational attainment constructed from untransmitted parental alleles predict socio-economic outcomes in children,<sup>87</sup> likely reflecting indirect effects and/or assortative mating. Since genetic effects cluster geographically,<sup>85</sup> indirect effects from individuals outside the direct family that are geographically and socially nearby could influence GWAS signals as well.<sup>89</sup>
- **Gene-environment correlations:** Genes cluster within families, communities, and neighborhoods that provide favorable or unfavorable environmental circumstances. Living circumstances differ greatly between poorer and richer households and neighborhoods. Genes that have an effect on traits that influence socio-economic outcomes therefore correlate with effects that come from the environments that people are born in (passive gene-environment correlations) or are able to move to throughout their lives (active gene-environment correlations). This can make genetic effects seem stronger than they are and can induce a correlation between genetic variants and other traits (e.g., BMI and height) that are influenced by the environmental factors that come with higher or lower socio-economic outcomes.
- **Assortative mating:** Socio-economic outcomes show strong positive assortative mating, where people select mates that resemble them. These outcomes are highly polygenic, i.e., influenced by many alleles with small effects. When trait-increasing alleles, which can be found throughout the whole genome, are transmitted together to the offspring due to assortative mating, they become increasingly correlated with each other in the population, leading to overestimation of their individual genetic effects. This is analogous to, for example, the effects of a good teacher on a student's education being overestimated because good teachers cluster together in schools in better neighborhoods, and thus children in the class of one good teacher tend to have classes from other good teachers as well.

433

434

435

## Figure Captions

**Figure 1: Changes in the heritability of educational attainment over time in Europe.** The Figures show the percentage of variation in educational attainment explained by genetic and environmental influences with 95% confidence intervals, as estimated in a meta-analysis of 28 European twin cohorts. Shared environmental influences refer to environmental factors that make siblings more similar to each other, while unique environmental influences are factors that do not, and also include measurement error. In twin studies, shared environmental influences can be overestimated at the expense of genetic influences when assortative mating occurs, as it increases the genetic similarity between siblings, mimicking shared environmental influences. The Figure is based on data from Table 2 from Silventoinen et al (2020).<sup>48</sup>

**Figure 2: Traits that show higher genetic correlations with educational attainment tend to show stronger regional differences.** Y-axis indicates the absolute genetic correlation ( $r_g$ ) of the trait with educational attainment (EA) (Lee et al, 2018), excluding all British. X-axis shows the Moran's I, a measure of geographic clustering, of 31 polygenic scores in ~320k individuals in Great Britain. The Moran's I of the educational attainment polygenic score is 0.6 (not shown). This Figure was adapted from Supplementary Figure 5 in Abdellaoui et al (2019).

**Figure 3: Polygenic prediction of average phenotypes per region likely captures environmental influences.** Polygenic scores for educational attainment (EA) capture environmental effects on a regional level that are not visible when examining individual-level data. Panel (a) shows this based on ~320k unrelated UK Biobank participants of European descent. The polygenic score for height explains 21% of individual differences in height, while the polygenic score for EA explains only 1% of individual differences in height. When we consider the average scores per region, however, the polygenic score for EA explains more regional differences in height (64%) than the polygenic score for height does (52%), presumably because the regional average of the EA polygenic score better captures regional differences in poor versus rich environments, and these affect height. Panel (b) displays the geographic distributions of regional averages of polygenic scores for height and EA, alongside the regional average for phenotypic height. Black lines denote coal mining regions, where environmental circumstances associated with socio-economic deprivation tend to cluster with polygenic scores for EA. Without such clustering, coal mining regions would be among the taller regions of the country, which phenotypically they are not. For statistical analyses regarding these regional differences and the migration herein, see Abdellaoui et al. (2019)<sup>85</sup>. The hypothesized causal diagram in panel (c) illustrates environmental influences on height (E) that cluster regionally with genetic influences associated with SES (G(SES)), making the regional average of those genes ( $\overline{G(SES)}_{\text{regional}}$ ) predictive of the regional average of phenotypic height ( $\overline{\text{Height}}_{\text{regional}}$ ). G(Height) denotes the genetic influences on height on an individual level (Height). For details on the data and QC for panels (a) and (b), see Abdellaoui et al 2022<sup>89</sup>; for details on polygenic score computation and geographic regions, see Abdellaoui et al, 2019<sup>85</sup>.

**Figure 4: Genetic correlations show that Covid-19 infections and deaths originate in higher SES regions and spread more widely in lower SES regions in England.** Figures a and b show data for

Covid-19 cases and deaths respectively. The Figures show results from a total of 2,924 regional GWASs (RGWASs; four per day - cases and deaths, cumulative and weekly - for 2 years, i.e., 731 days) performed on 1.2 million common single nucleotide polymorphisms (SNPs) from 396,042 individuals of European descent living in England. As opposed to a traditional GWAS, in an RGWAS the subjects are given the phenotype of the region they live in (315 regions), which often results in genetic signals associated with socio-economic outcomes due to their geographic clustering.<sup>85</sup> Covid-19 data on the 315 regions were obtained from Public Health England.<sup>76</sup> Each dot is one RGWAS (one day), for which either the weekly or the cumulative cases (**a**) or deaths (**b**) were analyzed as the phenotype. The upper panel shows the variation explained by all 1.2 million SNPs (the SNP-based heritability or SNP- $h^2$ ). The large middle panel shows the genetic correlation ( $r_g$ ) of the genetic signal with the educational attainment (EA) GWAS from Lee et al (2018)<sup>56</sup>. The gray shaded areas around the points indicate 95% confidence intervals for both the SNP- $h^2$  and genetic correlations. These genetic correlations shows several positive peaks, including at the start of the pandemic in March 2020 and around the start of the spread of the new and more contagious B1.1.7 variant in December 2020, both reflecting more infections in richer regions of the country (in or near London), after which the genetic correlation with education becomes negative again. The two bottom panels show the total number of cumulative and weekly Covid-19 cases (**a**) or deaths (**b**). For more information on quality control and statistical approaches used, see Abdellaoui (2020)<sup>137</sup>, where the same results were reported for cases only for the first ~6 weeks of the Covid-19 pandemic.

## REFERENCES

- 1 Flannery, K. & Marcus, J. *The creation of inequality: how our prehistoric ancestors set the stage for monarchy, slavery, and empire*. (Harvard University Press, 2012).
- 2 Mijs, J. J. The unfulfillable promise of meritocracy: Three lessons and their implications for justice in education. *Social Justice Research* **29**, 14-34 (2016).
- 3 Alesina, A. & La Ferrara, E. Participation in heterogeneous communities. *The quarterly journal of economics* **115**, 847-904 (2000).
- 4 Wilkinson, R. G. & Pickett, K. E. Income inequality and social dysfunction. *Annual review of sociology* **35**, 493-511 (2009).
- 5 Pickett, K. E. & Wilkinson, R. G. Income inequality and health: a causal review. *Social science & medicine* **128**, 316-326 (2015).
- 6 Dewar, M. *et al.* What makes a fair society? Insights and evidence. (Joint Research Centre (Seville site), 2017).
- 7 Piketty, T. & Saez, E. Inequality in the long run. *Science (New York, N.Y.)* **344**, 838-843 (2014).
- 8 Harden, K. P. & Koellinger, P. D. Using genetics for social science. *Nature Human Behaviour* **4**, 567-576, doi:10.1038/s41562-020-0862-5 (2020).
- 9 Markovits, D. *The Meritocracy Trap: Or, The Tyranny of Just Deserts*. (Penguin UK, 2019).
- 10 Dikötter, F. Race culture: Recent perspectives on the history of eugenics. *The American Historical Review* **103**, 467-478 (1998).
- 11 Rutherford, A. *Control: The Dark History and Troubling Present of Eugenics*. (WW Norton & Company, 2022).
- 12 Kevles, D. J. *In the name of eugenics: Genetics and the uses of human heredity*. (Harvard University Press, 1995).
- 13 Haviland, W. A., Prins, H. E., McBride, B. & Walrath, D. *Cultural anthropology: The human challenge*. (Cengage Learning, 2013).
- 14 Llobera, J. R. *An invitation to anthropology: the structure, evolution and cultural identity of human societies*. (Berghahn Books, 2003).
- 15 Widlok, T. & Tadesse, W. G. *Property and equality*. Vol. 2 (Berghahn Books, 2005).
- 16 Boehm, C. & Boehm, C. *Hierarchy in the forest: The evolution of egalitarian behavior*. (Harvard University Press, 2009).
- 17 Bocquet-Appel, J.-P. When the world's population took off: the springboard of the Neolithic Demographic Transition. *Science (New York, N.Y.)* **333**, 560-561 (2011).
- 18 Henrich, J. & Boyd, R. Division of labor, economic specialization, and the evolution of social stratification. *Current Anthropology* **49**, 715-724 (2008).
- 19 Weber, M. *Economy and society: An outline of interpretive sociology*. Vol. 1 (University of California press, 1978).
- 20 Bellah, R. N. *Religion in human evolution: From the Paleolithic to the Axial Age*. (Harvard University Press, 2011).
- 21 Liu, Y. *Higher education, meritocracy and inequality in China*. (Springer, 2016).
- 22 Weatherford, J. *Genghis Khan and the making of the modern world*. (Crown, 2005).
- 23 Schwarz, B. *The expansion of England: race, ethnicity and cultural history*. (Routledge, 2005).
- 24 Young, M. D. *The rise of the meritocracy*. (Transaction Publishers, 1994).
- 25 Black, S. E., Devereux, P. J. & Salvanes, K. G. Why the apple doesn't fall far: Understanding intergenerational transmission of human capital. *American economic review* **95**, 437-449 (2005).
- 26 Hauser, R. M. & Grusky, D. B. Cross-national variation in occupational distributions, relative mobility chances, and intergenerational shifts in occupational distributions. *American Sociological Review*, 723-741 (1988).
- 27 Erikson, R., Goldthorpe, J. H. & Portocarero, L. Intergenerational class mobility in three Western European societies: England, France and Sweden. *The British journal of sociology* **30**, 415-441 (1979).
- 28 Solon, G. Cross-country differences in intergenerational earnings mobility. *Journal of Economic Perspectives* **16**, 59-66 (2002).

553 29 Meyer, M. N. *et al.* Wrestling with social and behavioral genomics: risks, potential benefits, and  
554 ethical responsibility. *Hastings Center Report* **53**, S2-S49 (2023).

555 30 López-Beltrán, C. in *Heredity produced: At the crossroads of biology, politics, and culture, 1500–1870*  
556 105-132 (The MIT Press, 2007).

557 31 Galton, F. *Hereditary genius*. (D. Appleton, 1891).

558 32 Galton, F. *Inquiries into human faculty and its development*. (Macmillan, 1883).

559 33 Hyatt, S. A shared history of shame: Sweden's four-decade policy of forced sterilization and the  
560 eugenics movement in the United States. *Ind. Int'l & Comp. L. Rev.* **8**, 475 (1997).

561 34 Rutecki, G. W. Forced sterilization of native Americans: Later twentieth century physician  
562 cooperation with national eugenic policies? *Ethics & Medicine* **27**, 33 (2011).

563 35 Grekul, J., Krahn, A. & Odynak, D. Sterilizing the “Feeble-minded”: Eugenics in Alberta,  
564 Canada, 1929–1972. *Journal of Historical Sociology* **17**, 358-384 (2004).

565 36 Liu, A. *et al.* Evidence from Finland and Sweden on the relationship between early-life diseases  
566 and lifetime childlessness in men and women. *Nature Human Behaviour* **8**, 276-287 (2024).

567 37 Polderman, T. J. *et al.* Meta-analysis of the heritability of human traits based on fifty years of  
568 twin studies. *Nature Genetics*, **47**, 702-709, doi:10.1038/ng.3285 (2015).

569 38 Furnham, A. & Cheng, H. Factors influencing adult earnings: Findings from a nationally  
570 representative sample. *The Journal of Socio-Economics* **44**, 120-125 (2013).

571 39 Kappe, R. & Van Der Flier, H. Predicting academic success in higher education: what’s more  
572 important than being smart? *European Journal of Psychology of Education* **27**, 605-619 (2012).

573 40 von Stumm, S. *et al.* Predicting educational achievement from genomic measures and  
574 socioeconomic status. *Developmental science* **23**, e12925 (2020).

575 41 Schmidt, F. L. & Hunter, J. E. The validity and utility of selection methods in personnel  
576 psychology: Practical and theoretical implications of 85 years of research findings.  
577 *Psychological bulletin* **124**, 262 (1998).

578 42 Roth, B. *et al.* Intelligence and school grades: A meta-analysis. *Intelligence* **53**, 118-137 (2015).

579 43 Merriman, C. The intellectual resemblance of twins. *Psychological Monographs* **33**, i (1924).

580 44 Rende, R. D., Plomin, R. & Vandenberg, S. G. Who discovered the twin method? *Behavior*  
581 *genetics* **20**, 277-285 (1990).

582 45 Bouchard, T. J. & McGue, M. Familial studies of intelligence: A review. *Science (New York,*  
583 *N.Y.)* **212**, 1055-1059 (1981).

584 46 Plomin, R., DeFries, J. C., Knopik, V. S. & Neiderhiser, J. M. Top 10 replicated findings from  
585 behavioral genetics. *Perspectives on psychological science* **11**, 3-23 (2016).

586 47 Engzell, P. & Troup, F. C. Heritability of education rises with intergenerational mobility.  
587 *Proceedings of the National Academy of Sciences* **116**, 25386-25388 (2019).

588 48 Silventoinen, K. *et al.* Genetic and environmental variation in educational attainment: an  
589 individual-based analysis of 28 twin cohorts. *Scientific reports* **10**, 1-11 (2020).

590 49 DeWan, A. *et al.* HTRA1 promoter polymorphism in wet age-related macular degeneration.  
591 *Science (New York, N.Y.)* **314**, 989-992 (2006).

592 50 The Wellcome Trust Case Control Consortium. Genome-wide association study of 14,000 cases  
593 of seven common diseases and 3,000 shared controls. *Nature* **447**, 661-678 (2007).

594 51 Klein, R. J. *et al.* Complement factor H polymorphism in age-related macular degeneration.  
595 *Science (New York, N.Y.)* **308**, 385-389 (2005).

596 52 Risch, N. & Merikangas, K. The future of genetic studies of complex human diseases. *Science*  
597 *(New York, N.Y.)* **273**, 1516-1517 (1996).

598 53 Abdellaoui, A., Yengo, L., Verweij, K. J. H. & Visscher, P. M. 15 years of GWAS discovery:  
599 Realizing the promise. *The American Journal of Human Genetics* **110**, 179-194,  
600 doi:https://doi.org/10.1016/j.ajhg.2022.12.011 (2023).

601 54 Rietveld, C. A. *et al.* GWAS of 126,559 individuals identifies genetic variants associated with  
602 educational attainment. *Science (New York, N.Y.)* **340**, 1467-1471,  
603 doi:10.1126/science.1235488 (2013).

604 55 Okbay, A. *et al.* Genome-wide association study identifies 74 loci associated with educational  
605 attainment. *Nature* **533**, 539-542, doi:10.1038/nature17671 (2016).

- 56 Lee, J. J. *et al.* 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 (2018).
- 57 Okbay, A. *et al.* Polygenic prediction of educational attainment within and between families from genome-wide association analyses in 3 million individuals. *Nature genetics* **54**, 437-449, doi:10.1038/s41588-022-01016-z (2022).
- 58 Akimova, E. T., Wolfram, T., Ding, X., Tropf, F. C. & Mills, M. C. Polygenic predictions of occupational status GWAS elucidate genetic and environmental interplay for intergenerational status transmission, careers, and health. *Nature Human Behaviour* (2024).
- 59 Kweon, H. *et al.* Associations between common genetic variants and income provide insights about the socioeconomic health gradient. *Nature Human Behaviour* (2024).
- 60 Hill, W. D. *et al.* Genome-wide analysis identifies molecular systems and 149 genetic loci associated with income. *Nature communications* **10**, 5741 (2019).
- 61 Chen, T.-T. *et al.* Shared genetic architectures of educational attainment in East Asian and European populations. *Nature Human Behaviour*, 1-14 (2024).
- 62 Fisher, R. A. The correlation between relatives on the supposition of Mendelian inheritance. *Philosophical Transactions of the Royal Society of Edinburgh* **52**, 399-433 (1918).
- 63 Plomin, R., Haworth, C. M. & Davis, O. S. Common disorders are quantitative traits. *Nature reviews genetics* **10**, 872-878 (2009).
- 64 Visscher, P. M. & Goddard, M. E. From RA Fisher's 1918 paper to GWAS a century later. *Genetics* **211**, 1125-1130 (2019).
- 65 Tan, T. *et al.* Family-GWAS reveals effects of environment and mating on genetic associations. *medRxiv*, 2024.2010.2001.24314703, doi:10.1101/2024.10.01.24314703 (2024).
- 66 Wilding, K., Wright, M. & von Stumm, S. Using DNA to Predict Education: a Meta-analytic Review. *Educational psychology review* **36**, 102 (2024).
- 67 Abdellaoui, A. & Verweij, K. J. Dissecting polygenic signals from genome-wide association studies on human behaviour. *Nature Human Behaviour* **5**, 686-694 (2021).
- 68 Darwin, C. R. *On the origin of species by means of natural selection, or the preservation of favoured races in the struggle for life.* (John Murray, 1859).
- 69 Hugh-Jones, D. & Abdellaoui, A. Human capital mediates natural selection in contemporary humans. *Behavior genetics* **52**, 205-234 (2022).
- 70 Mills, M. C. & Mathieson, I. The challenge of detecting recent natural selection in human populations. *Proceedings of the National Academy of Sciences* **119**, e2203237119 (2022).
- 71 Zeng, J. *et al.* Signatures of negative selection in the genetic architecture of human complex traits. *Nature genetics* **50**, 746-753 (2018).
- 72 Zeng, J. *et al.* Widespread signatures of natural selection across human complex traits and functional genomic categories. *Nature communications* **12**, 1164 (2021).
- 73 O'Connor, L. J. *et al.* Extreme polygenicity of complex traits is explained by negative selection. *The American Journal of Human Genetics* **105**, 456-476 (2019).
- 74 Lek, M. *et al.* Analysis of protein-coding genetic variation in 60,706 humans. *Nature* **536**, 285-291 (2016).
- 75 Cassa, C. A. *et al.* Estimating the selective effects of heterozygous protein-truncating variants from human exome data. *Nature genetics* **49**, 806-810 (2017).
- 76 Mathieson, I. *et al.* Genome-wide analysis identifies genetic effects on reproductive success and ongoing natural selection at the FADS locus. *Nature Human Behaviour* **7**, 790-801, doi:10.1038/s41562-023-01528-6 (2023).
- 77 Gardner, E. J. *et al.* Reduced reproductive success is associated with selective constraint on human genes. *Nature* **603**, 858-863, doi:10.1038/s41586-022-04549-9 (2022).
- 78 Ganna, A. *et al.* Quantifying the impact of rare and ultra-rare coding variation across the phenotypic spectrum. *The American Journal of Human Genetics* **102**, 1204-1211 (2018).
- 79 Ganna, A. *et al.* Ultra-rare disruptive and damaging mutations influence educational attainment in the general population. *Nature neuroscience* **19**, 1563-1565 (2016).
- 80 Paynter, R. The archaeology of equality and inequality. *Annual review of anthropology* **18**, 369-399 (1989).

- 81 Smith, M. E. The archaeological study of neighborhoods and districts in ancient cities. *Journal of Anthropological Archaeology* **29**, 137-154 (2010).
- 82 Abdellaoui, A. *et al.* Population structure, migration, and diversifying selection in the Netherlands. *European journal of human genetics : EJHG* **21**, 1277-1285, doi:10.1038/ejhg.2013.48 (2013).
- 83 Kerminen, S. *et al.* Fine-scale genetic structure in Finland. *G3: Genes, Genomes, Genetics* **7**, 3459-3468 (2017).
- 84 Leslie, S. *et al.* The fine-scale genetic structure of the British population. *Nature* **519**, 309-314 (2015).
- 85 Abdellaoui, A. *et al.* Genetic correlates of social stratification in Great Britain. *Nature human behaviour* **3**, 1332-1342 (2019).
- 86 Kuznetsov, I. A. *et al.* Assessing the impact of 20th century internal migrations on the genetic structure of Estonia. *bioRxiv*, (2023).
- 87 Kong, A. *et al.* The nature of nurture: Effects of parental genotypes. *Science (New York, N.Y.)* **359**, 424-428 (2018).
- 88 Howe, L. J. *et al.* Within-sibship genome-wide association analyses decrease bias in estimates of direct genetic effects. *Nature genetics* **54**, 581-592 (2022).
- 89 Abdellaoui, A., Dolan, C. V., Verweij, K. J. H. & Nivard, M. G. Gene–environment correlations across geographic regions affect genome-wide association studies. *Nature genetics* **54**, 1345-1354, doi:10.1038/s41588-022-01158-0 (2022).
- 90 Ding, X., Brazel, D. M. & Mills, M. C. Factors affecting adherence to non-pharmaceutical interventions for COVID-19 infections in the first year of the pandemic in the UK. *BMJ open* **11**, e054200 (2021).
- 91 Wade, L. From Black Death to fatal flu, past pandemics show why people on the margins suffer most. *Science (New York, N.Y.)* (2020).
- 92 Dobson, A. P. & Carper, E. R. Infectious diseases and human population history. *Bioscience* **46**, 115-126 (1996).
- 93 Lindahl, J. F. & Grace, D. The consequences of human actions on risks for infectious diseases: a review. *Infection ecology & epidemiology* **5**, 30048 (2015).
- 94 Lee, R. D. Population dynamics of humans and other animals. *Demography*, 443-465 (1987).
- 95 Clark, G. & Cummins, N. Malthus to modernity: wealth, status, and fertility in England, 1500–1879. *Journal of Population Economics* **28**, 3-29 (2015).
- 96 Clark, G. & Hamilton, G. Survival of the richest: the Malthusian mechanism in pre-industrial England. *The Journal of Economic History* **66**, 707-736 (2006).
- 97 Jaadla, H., Potter, E., Keibek, S. & Davenport, R. Infant and child mortality by socio-economic status in early nineteenth-century England. *The Economic History Review* **73**, 991-1022 (2020).
- 98 Buss, D. M. & Schmitt, D. P. Mate preferences and their behavioral manifestations. *Annual review of psychology* **70**, 77-110 (2019).
- 99 Skirbekk, V. Fertility trends by social status. *Demographic research* **18**, 145-180 (2008).
- 100 Matysiak, A. & Vignoli, D. Fertility and Women’s Employment: A Meta-analysis: Fécondité et travail des femmes: une méta-analyse. *European Journal of Population/Revue européenne de Démographie* **24**, 363-384 (2008).
- 101 Cleland, J. Potatoes and pills: An overview of innovation-diffusion contributions to explanations of fertility decline. *Diffusion processes and fertility transition: Selected perspectives*, 39-65 (2001).
- 102 Stulp, G., Sear, R., Schaffnit, S. B., Mills, M. C. & Barrett, L. The reproductive ecology of industrial societies, Part II: the association between wealth and fertility. *Human Nature* **27**, 445-470 (2016).
- 103 Barban, N. *et al.* Genome-wide analysis identifies 12 loci influencing human reproductive behavior. *Nature genetics* **48**, 1462-1472 (2016).
- 104 Beauchamp, J. P. Genetic evidence for natural selection in humans in the contemporary United States. *Proceedings of the National Academy of Sciences* **113**, 7774-7779 (2016).
- 105 Kong, A. *et al.* Selection against variants in the genome associated with educational attainment. *Proceedings of the National Academy of Sciences* **114**, E727-E732 (2017).

Sanjak, J. S., Sidorenko, J., Robinson, M. R., Thornton, K. R. & Visscher, P. M. Evidence of directional and stabilizing selection in contemporary humans. *Proceedings of the National Academy of Sciences* **115**, 151-156 (2018).

Doepke, M., Hannusch, A., Kindermann, F. & Tertilt, M. The economics of fertility: A new era. *National Bureau of Economics Research* (2023).

Kolk, M. & Barclay, K. Cognitive ability and fertility among Swedish men born 1951–1967: evidence from military conscription registers. *Proceedings of the Royal Society B* **286**, 20190359 (2019).

Kolk, M. The relationship between life-course accumulated income and childbearing of Swedish men and women born 1940–70. *Population Studies* **77**, 197-215 (2023).

Wright, S. Systems of mating. III. Assortative mating based on somatic resemblance. *Genetics* **6**, 144 (1921).

Schwartz, C. R. Trends and variation in assortative mating: Causes and consequences. *Annual Review of Sociology* **39**, 451-470 (2013).

Horwitz, T. B., Balbona, J. V., Paulich, K. N. & Keller, M. C. Evidence of correlations between human partners based on systematic reviews and meta-analyses of 22 traits and UK Biobank analysis of 133 traits. *Nature Human Behaviour* **7**, 1568-1583 (2023).

Abdellaoui, A., Verweij, K. J. & Zietsch, B. P. No evidence for genetic assortative mating beyond that due to population stratification. *Proceedings of the National Academy of Sciences of the United States of America* **111**, E4137, doi:10.1073/pnas.1410781111 (2014).

Hugh-Jones, D., Verweij, K. J. H., Pourcain, B. S. & Abdellaoui, A. Assortative mating on educational attainment leads to genetic spousal resemblance for polygenic scores. *Intelligence* **59**, 103-108 (2016).

Kemper, K. E. *et al.* Phenotypic covariance across the entire spectrum of relatedness for 86 billion pairs of individuals. *Nature communications* **12**, 1050 (2021).

Robinson, M. R. *et al.* Genetic evidence of assortative mating in humans. *Nature Human Behaviour* **1**, 0016 (2017).

Sunde, H. F. *et al.* Genetic similarity between relatives provides evidence on the presence and history of assortative mating. *Nature communications* **15**, 2641 (2024).

Young, A. S. Estimation of indirect genetic effects and heritability under assortative mating. *bioRxiv* (2023).

Sunde, H. F., Eilertsen, E. M. & Torvik, F. A. Understanding indirect assortative mating and its intergenerational consequences. *bioRxiv*, 2024.2006.2021.600029, doi:10.1101/2024.06.21.600029 (2024).

Greenwood, J., Guner, N., Kocharkov, G. & Santos, C. Marry your like: Assortative mating and income inequality. (National Bureau of Economic Research, 2014).

Dribe, M. & Lundh, C. Status homogamy in the preindustrial marriage market: partner selection according to age, social origin, and place of birth in nineteenth-century rural Sweden. *Journal of Family History* **34**, 387-406 (2009).

Zijdeman, R. L. & Maas, I. Assortative mating by occupational status during early industrialization. *Research in Social Stratification and Mobility* **28**, 395-415 (2010).

Clark, G. The inheritance of social status: England, 1600 to 2022. *Proceedings of the National Academy of Sciences* **120**, e2300926120 (2023).

Abdellaoui, A. *et al.* Trading social status for genetics in marriage markets: evidence from Great Britain and Norway. School of Economics Working Paper (2023).

Uchiyama, R., Spicer, R. & Muthukrishna, M. Cultural evolution of genetic heritability. *Behavioral and Brain Sciences* **45**, e152 (2022).

Abdellaoui, A. The evolutionary dance between culture, genes, and everything in between. *The Behavioral and brain sciences* **45**, e153 (2022).

Tropf, F. C. *et al.* Hidden heritability due to heterogeneity across seven populations. *Nature Human Behaviour* **1**, 757-765, doi:10.1038/s41562-017-0195-1 (2017).

Yengo, L. *et al.* Imprint of assortative mating on the human genome. *Nature Human Behaviour* **2**, 948-954, doi:10.1038/s41562-018-0476-3 (2018).

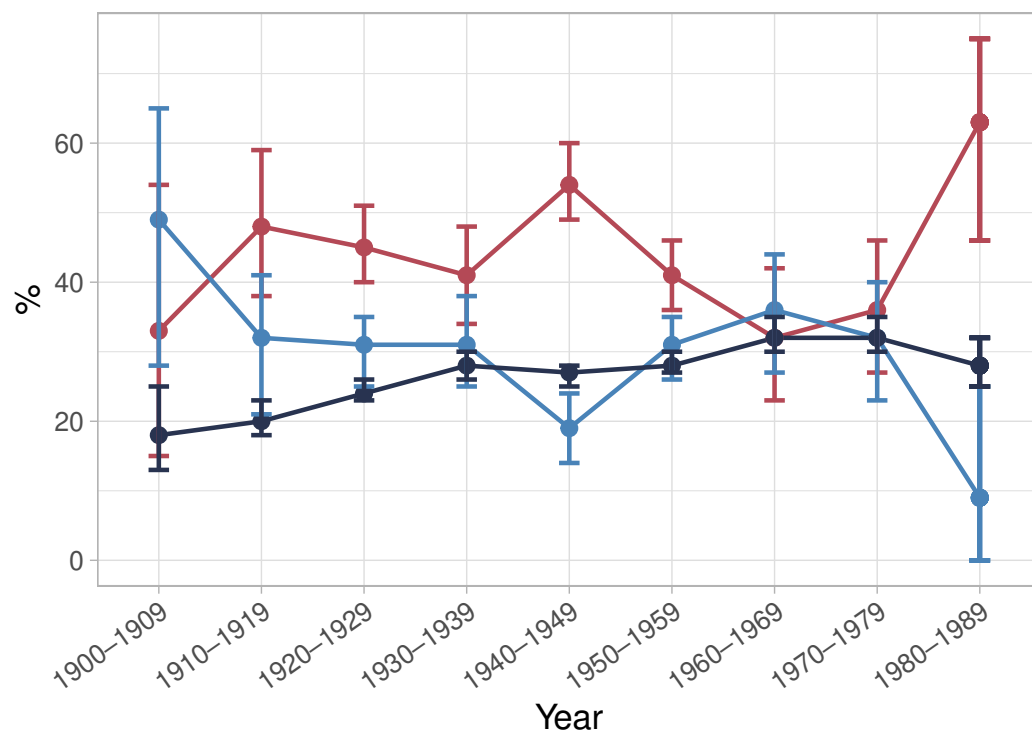
Young, A. I. *et al.* Mendelian imputation of parental genotypes improves estimates of direct genetic effects. *Nature genetics* **54**, 897-905 (2022).



- 130 Mills, M. C. & Rahal, C. The GWAS Diversity Monitor tracks diversity by disease in real time. *Nature genetics* **52**, 242-243 (2020).
- 131 Mills, M. C. & Rahal, C. A scientometric review of genome-wide association studies. *Communications biology* **2**, 1-11 (2019).
- 132 Schoeler, T. *et al.* Participation bias in the UK Biobank distorts genetic associations and downstream analyses. *Nature Human Behaviour* **7**, 1216-1227, doi:10.1038/s41562-023-01579-9 (2023).
- 133 van Alten, S., Domingue, B. W., Faul, J., Galama, T. & Marees, A. T. Reweighting UK Biobank corrects for pervasive selection bias due to volunteering. *International journal of epidemiology* **53**, doi:10.1093/ije/dyae054 (2024).
- 134 Cheesman, R. *et al.* Comparison of adopted and nonadopted individuals reveals gene–environment interplay for education in the UK Biobank. *Psychological science* **31**, 582-591 (2020).
- 135 Selzam, S. *et al.* Comparing Within- and Between-Family Polygenic Score Prediction. *American journal of human genetics* **105**, 351-363, doi:10.1016/j.ajhg.2019.06.006 (2019).
- 136 Nivard, M. G. *et al.* More than nature and nurture, indirect genetic effects on children’s academic achievement are consequences of dynastic social processes. *Nature Human Behaviour* **8**, 771-778, doi:10.1038/s41562-023-01796-2 (2024).
- 137 Abdellaoui, A. Regional differences in reported Covid-19 cases show genetic correlations with higher socio-economic status and better health, potentially confounding studies on the genetics of disease susceptibility. *MedRxiv* (2020).

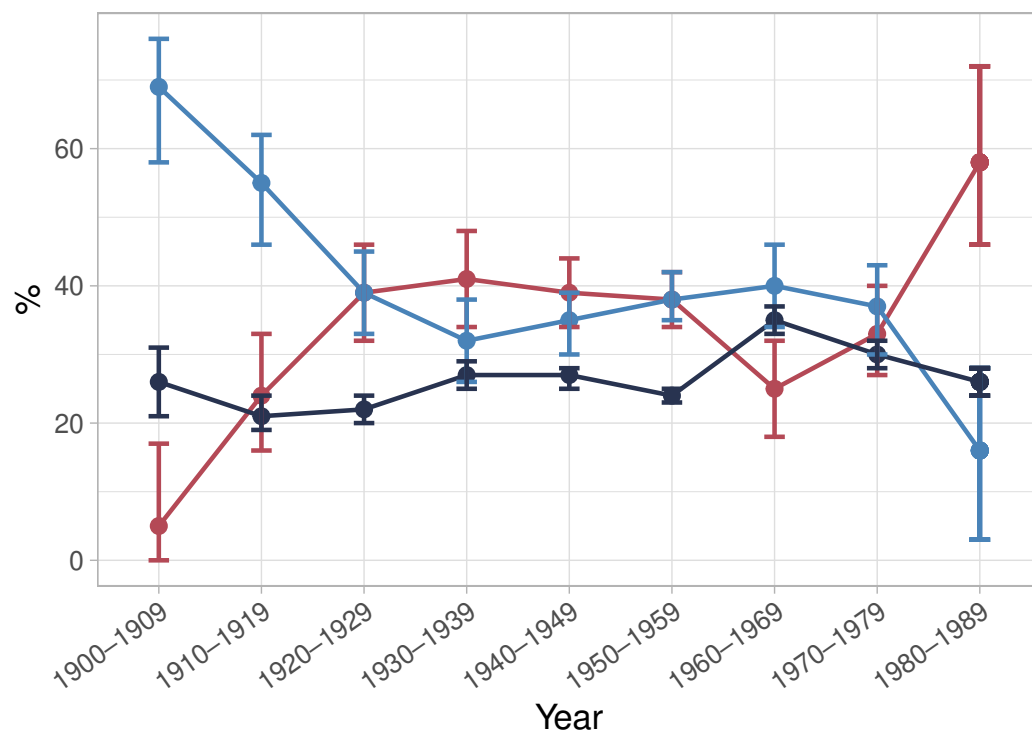
## Males

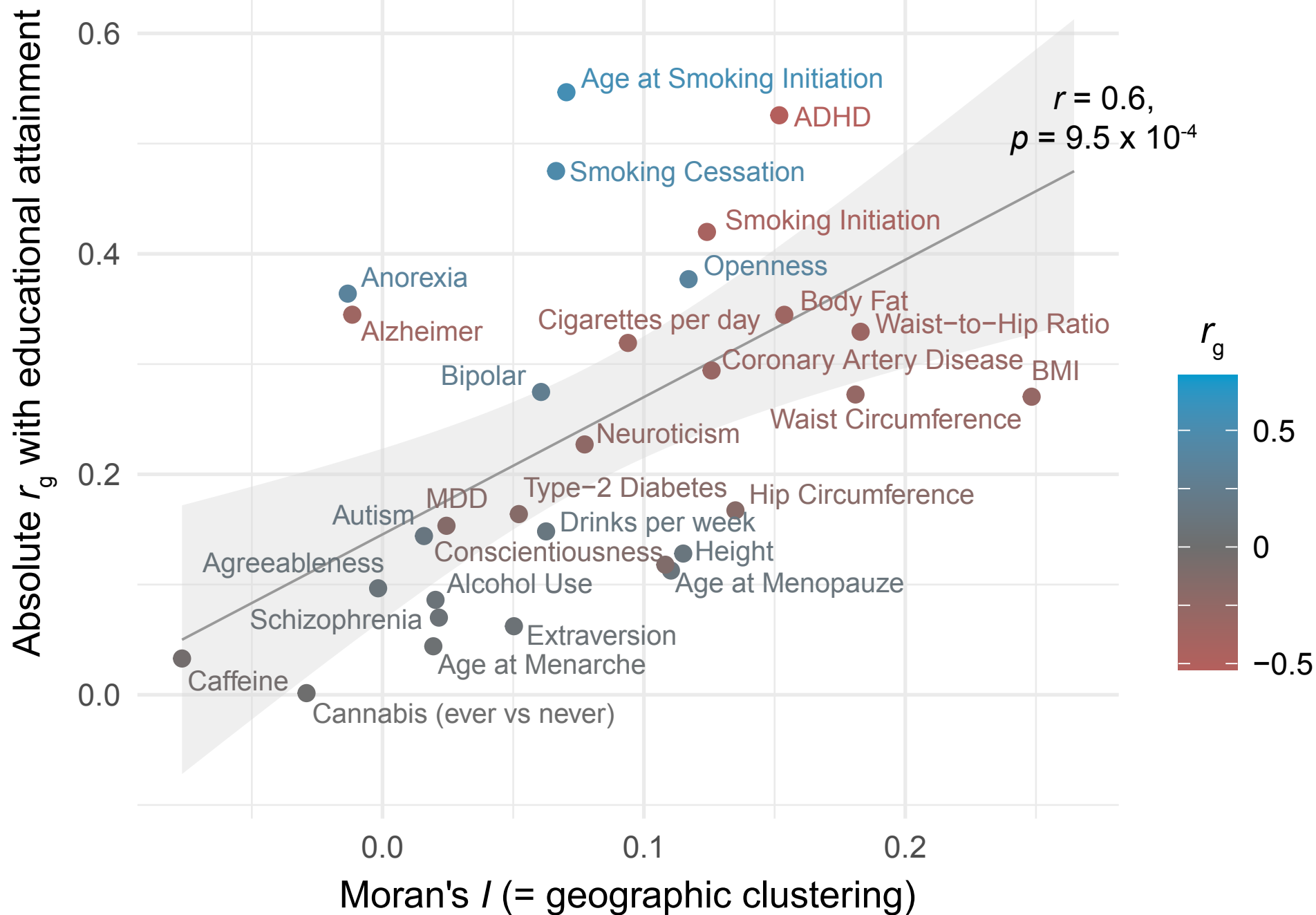
Genetic Shared Environment Unique Environment



## Females

Genetic Shared Environment Unique Environment





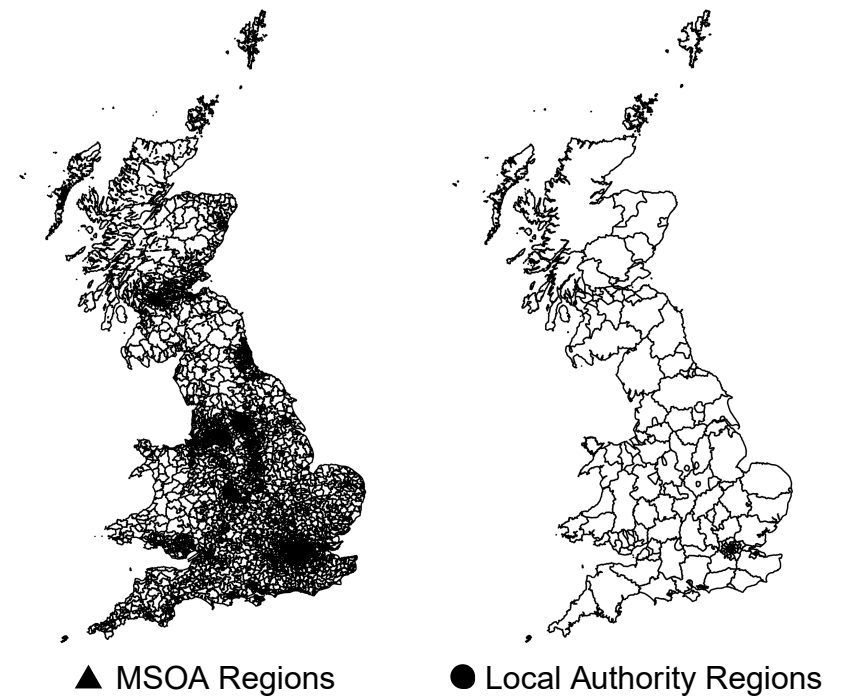
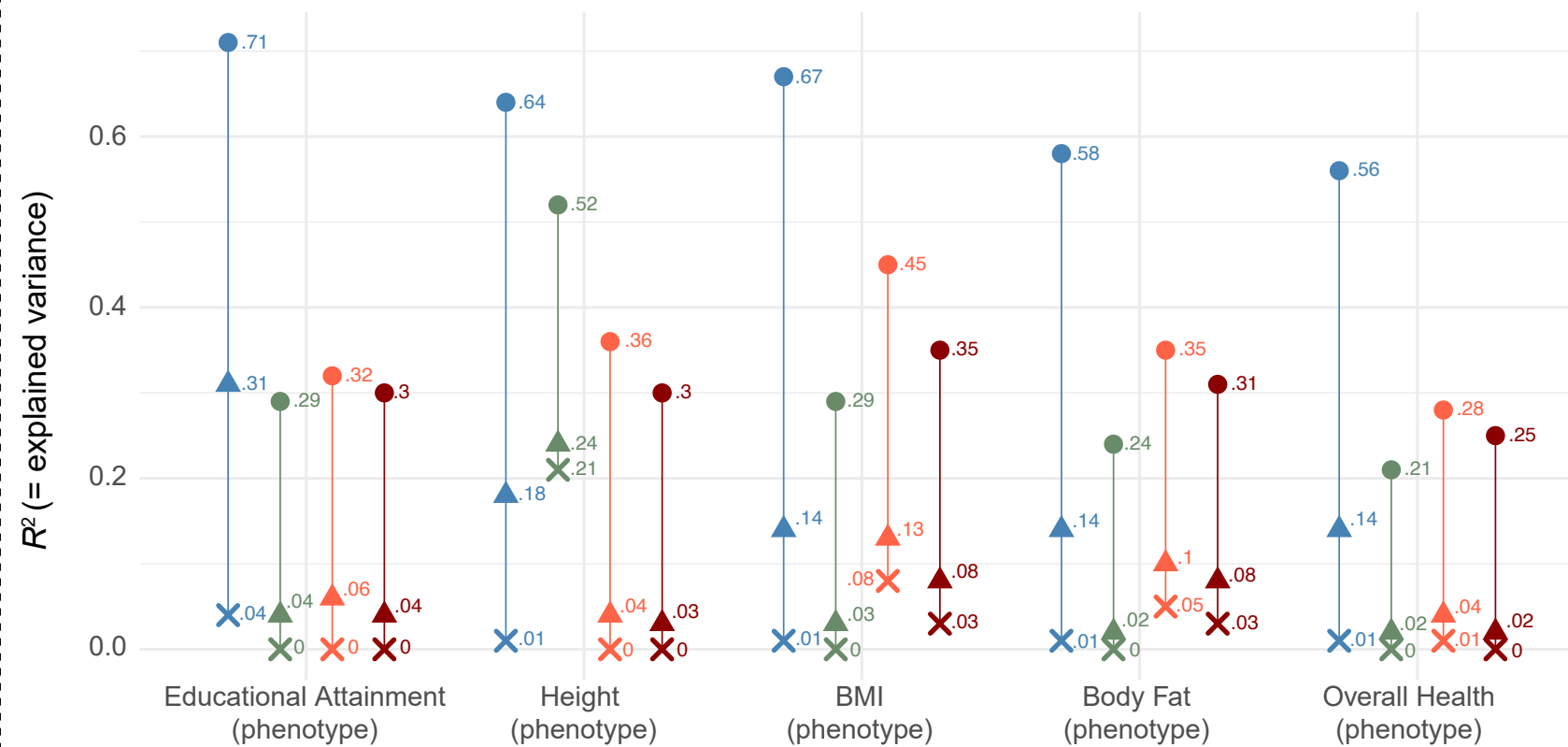
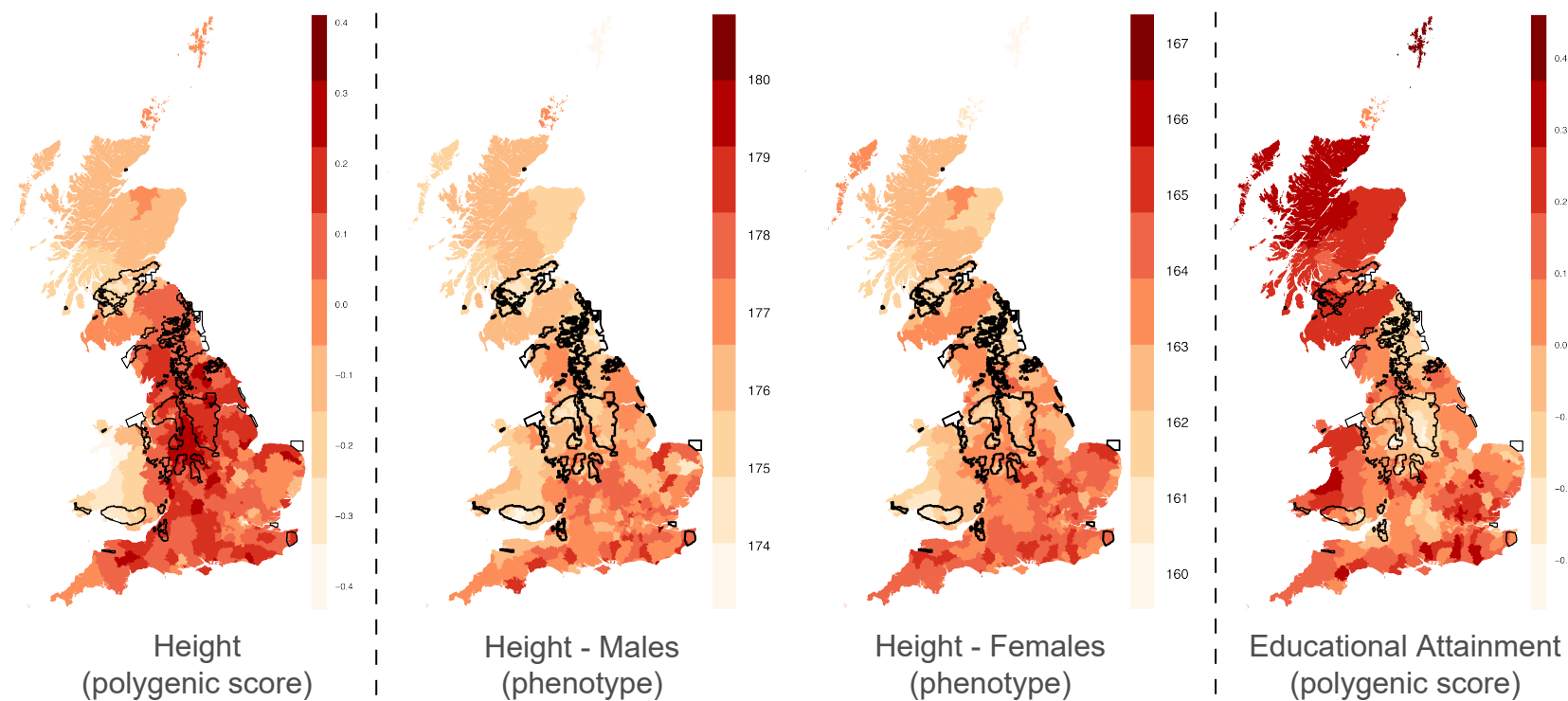
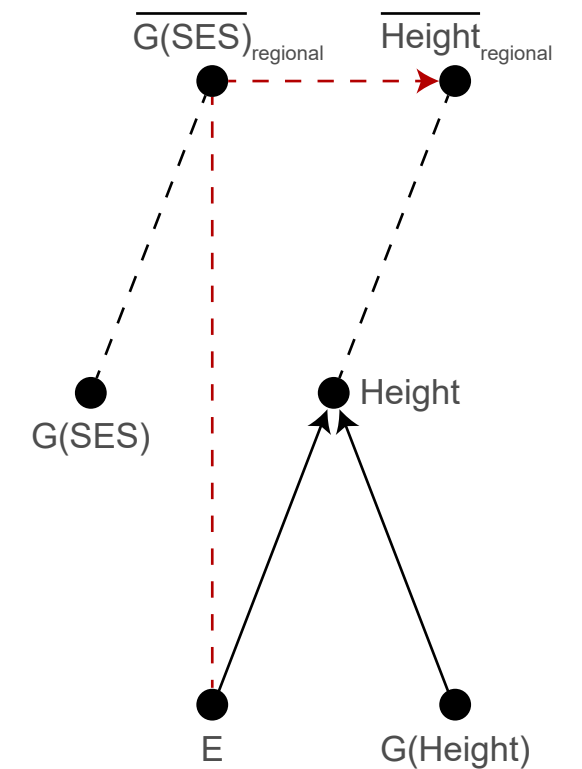
**a**

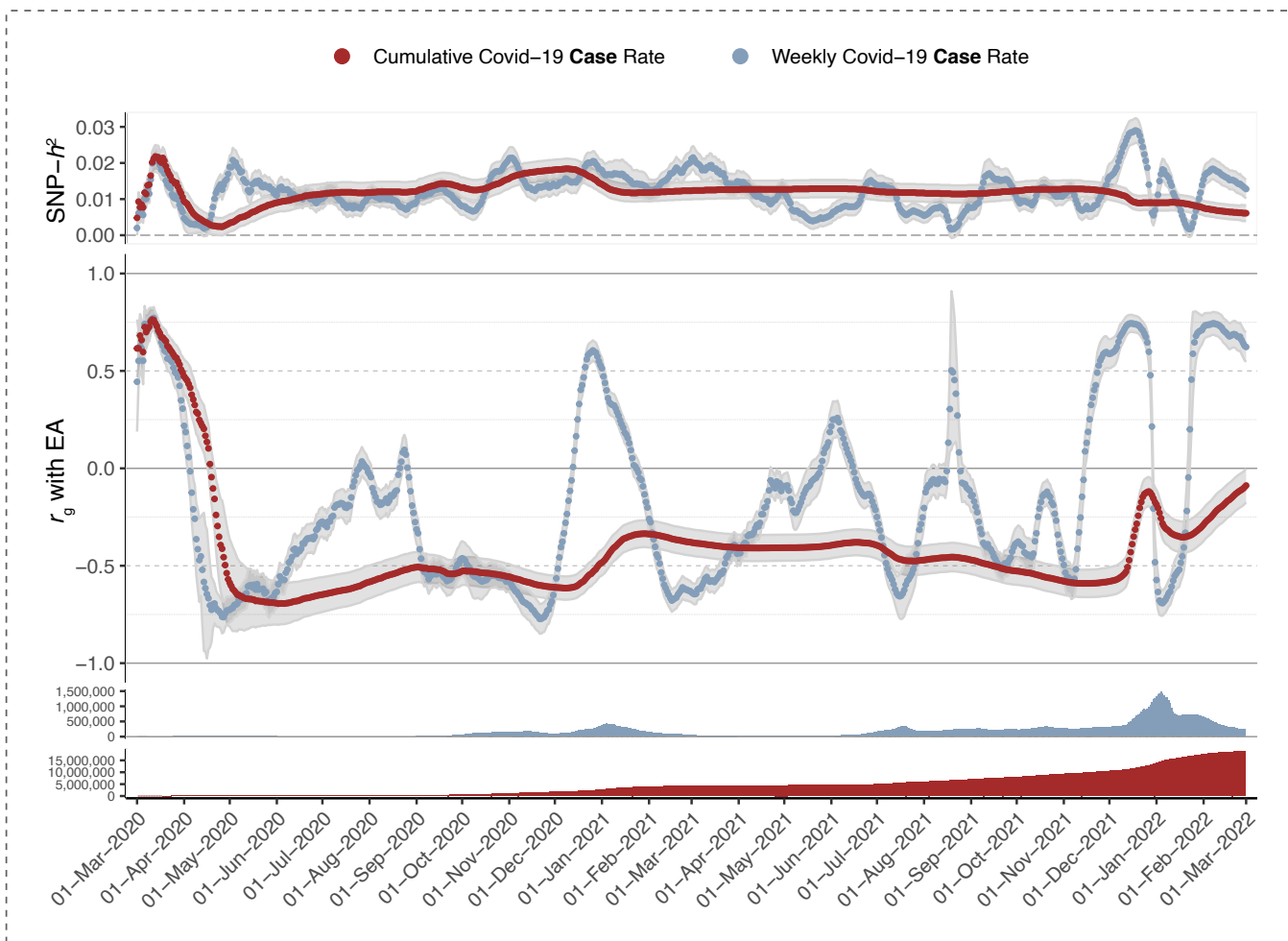
### Polygenic Scores vs Phenotypes

Individual-level vs Regional (MSOA + Local Authority) level

- Educational Attainment (polygenic score)
- Height (polygenic score)
- BMI (polygenic score)
- Body Fat (polygenic score)

- Regional Averages: Local Authority Regions
- ▲ Regional Averages: MSOA Regions
- ✕ Individual-Level

**b****c**

**a****b**