The contribution of mate-choice, couple convergence and confounding to assortative mating

Jennifer Sjaarda1,2,3, Zoltán Kutalik1,2,3

1 University Center for Primary Care and Public Health, Lausanne, Switzerland

2 Department of Computational Biology, University of Lausanne, Switzerland

3 Swiss Institute of Bioinformatics, Lausanne, Switzerland

Abstract

Increased phenotypic similarity between partners, termed assortative mating (AM), has been observed for many traits. However, it is currently unclear if these observations are due to mate choice for certain phenotypes, post-mating convergence, or a result of confounding factors such as shared environment. To dissect these underlying phenomena, we adapted Mendelian randomisation (MR) to 51,664 couples in the UK biobank for a panel of 118 phenotypes under AM. We found that over half (64 of 118) of the tested traits were found to have a causal relationship between partners, with females having on average larger effect. Forty traits, including systolic blood pressure, basal metabolic rate, weight and height, showed significantly larger phenotypic correlation than MR-estimates, suggesting the presence of confounders. Subsequent analyses revealed smoking, overall health rating, household income, and education as major overall confounders, accounting for XX, XX, 29.8 and 11.6% of phenotypic correlations, respectively. We found some limited evidence for couple-correlation convergence (e.g. increased similarity with respect to smoking and medication use), measured by stratifying couples by their time spent together. Finally, we found that the vast majority (77%) of identified cross-trait causal associations among partners can be explained by same-trait AM combined with a causal effect between these traits (within the same individual), with negligible direct cross-trait effects. In summary, this study has revealed many novel causal effects within couples and shed light on the impact of confounding on couple phenotypic similarity.

Introduction

In human populations, increased phenotypic similarity exists between partners compared to random pairs, a phenomenon known as (positive) assortative mating (AM). This has been observed across a wide variety of traits, including anthropometric measures (such as BMI and height), socioeconomic factors, various behavioural and lifestyle measures, (including diet, smoking habits, hobbies, among others), and even disease risk1–8. These observations can be explained by several factors. First, people tend to and actively seek out partners who are more similar to themselves with respect to certain phenotypes9,10. Second, phenotypic similarity can reflect post-mating convergence due to shared household and/or partner influence and interaction over time11–13. Finally, non-random assortment with respect to a phenotype can be due to confounders (at the moment of mate choice) such as shared environment and sociocultural and/or geographical barriers14–16.

The causes and consequences of phenotypic assortment remain unresolved and have implications in the study of human behaviour, population genetics, and public health. For instance, increased phenotypic similarity could naturally imply genetic similarity. This can result in variants that are otherwise independent to become correlated, which can consequently result in elevated resemblance between siblings and increased variation between different families, which could ultimately result in a concentration of (genetic) resources17–19. Indeed, evidence for genetic indirect effects among couples have been observed. For instance, it has been shown that the genome of an individual can predict the traits of their partner20. Another study found evidence of horizonal effects between an individual’s genome and the trait of their partner, suggesting that the partner heritability of a trait cannot be solely explained by between partner trait correlation21.

Any trait influenced by shared confounders will show assortment, hence it is crucial to separate traits under direct assortment from those that show partner-similarity due to being driven by another trait/factor under direct assortment. Thus, traits can be under direct- and/or indirect mate choice, and additionally modified by post-mating convergence. These phenomena can be rephrased for modelling purposes as follows: we treat direct assortment as a causal effect acting between the traits of the couple (index to partner) and indirect assortment as a confounder effect. These cross-partner causal effects can have two types: direct mate choice and direct influence of a partner during co-habitation. Confounder effects can emerge due to shared factors (e.g. socio-economic, geographic) and/or traits under direct assortment, and these can occur pre-mate choice or intensified post-mate choice due to shared household/habits.

Despite limited pioneering work, it is currently unknown to what extent the observed phenotypic similarity between partners is due to the three outlined components, specifically resolving the impact of confounding from casual factors. Analogous to challenges in classical epidemiological studies, where it is difficult, if not impossible, to discern causal factors from confounders, mere phenotypic similarity among couples is susceptible to the same interpretational limitations and challenges. Mendelian randomization (MR) is an alternative approach which is used to assess causality with large-scale observational with available genetic instruments. MR takes advantage of the random allocation of genetic variants, to infer causality between an exposure and an outcome22. This random allocation of genetic variants minimizes the possibility of reverse causality and confounding. To date, MR has proven to be a reliable causal inference method, revealing thousands of novel, causal relationships between exposures and outcomes.

In this work, we sought to adapt MR by examining causality between individuals, where the exposure and outcomes traits are measured in different individuals (whereas classical MR designs involve a single individual, e.g. BMI risk on CAD). This approach has been attempted for exploring couple effects with respect to alcohol consumption, and it was shown that while the observed phenotypic correlation does not tend to increase with age, the observed correlation and the estimated direct causal effect differed substantially23. Here, we examined a large number of complex traits and applied MR to estimate the direct causal effects impacting mate-choice, explored the impact of time couples live together on their similarity, and examined the cumulative role of a wide range of potential confounders on trait correlations between partners. Finally, we explored how cross-trait AM emerges by dissecting them to direct and indirect (same-trait AM combined with classical (same-sample cross-trait) causal effects) counterparts.

Methods

Sample selection and couple definition

This study used the UK Biobank (UKBB) cohort, a prospective population-based study with over 500,000 adult participants. Couples were identified and selected within the UKBB according to the following procedure. The initial UKBB sample comprised 502,616 individuals. First, participants were filtered to only genotyped, white, unrelated individuals according to the genetic QC file (specifically participants were retained if they had the following values in the QC file: “excess.relatives” = 0, “putative.sex.chromosome.aneuploidy” = 0, “in.white.british.ancestry.subset” = 1 and “used.in.pca.calculation” = 1). Redacted samples and participants that removed consent were also excluded. After filtering, 337,138 participants remained. Within this sample, we included individuals coming from households with exactly two unrelated, opposite-sex participants, leaving 108, 898 participants. Finally, using the data at data-field 6141, “How are people in household related to participant” pairs were filtered to only include couples who had both responded “Husband, wife, or partner”, leaving 103,328 participants, comprising 51,664 couples for downstream analyses (Supplementary Figure 1).

Mendelian Randomisation

MR is based on the principal that genetic variants at birth to assess the causal relationship between an exposure of interest and an outcome. ﻿The random distribution of genetic variants at birth reduces the possibility of confounding or reverse causation as explanations for the link between the exposure and outcome in the same way that the random allocation of a therapy in a randomized controlled trial minimizes this risk. MR relies on three core assumptions of the genetic variants, or instrumental variables (IVs), used to inform the causal relationship between an exposure and outcome. First, the IVs must be associated with the exposure of interest (the relevance assumption). Second, IVs must not be associated with any confounder in the exposure-outcome relationship (the exchangeability assumption). Third, IVs must not affect the outcome except through the exposure (the exclusion restriction assumption). There are several methods to estimate the causal effect using MR, the simplest being the ratio method, whereby a ratio is taken between the variant-outcome association and the variant-risk factor association, known as the Wald method. A natural extension of this, combining multiple IVs, is known as the inverse-variance weighted (IVW) method, which we applied in this report24, where the causal effect of on using genetic variant is given by the formula , with standard error .

Phenotype selection and processing

We performed an agnostic, phenome-wide approach for selecting phenotypes in this study. Specifically, we first selected phenotypes which were analysed by the Neale group and which had both male, female and joint summary statistics available. Next, after intersecting this list of phenotypes with our internal database (application number #16389), we had 1,278 phenotypes available for analysis. Phenotypes were processed in the filtered QC-data set (N = 337,138) according to a slightly modified version of the PHESANT pipeline to accommodate the phenotypes that we had available in our database25. Continuous variables were transformed to a normal distribution using a rank-preserving inverse normal quantile transformation (INQT), while ordinal and binary traits were re-categorized according to PHESANT documentation (for e.g. categories with less than 10 participants were removed). We then filtered these phenotypes as follows. First, we computed the raw phenotypic correlation amongst couples and removed phenotypes with a Pearson correlation < 0.1, in order to focus on traits with some indication of assortment. To ensure that INQT was not significantly impacting the correlations of each trait, we also calculated the Spearman correlation for each trait, and found consistent correlation estimates (Supplementary Figure 2). Next, we removed phenotypes which had less than 5 valid IVs for MR. IVs were defined based on an association *p* < 5 x 10-8 in the joint Neale summary statistics, after pruning for independence (based on a clumping procedure performed in PLINK with the options --clump-kb 10000 and --clump-r2 0.001 using the 1000 Genomes European samples as a reference). Third, using the sex-specific summary statistics, the IV heterogeneity between sexes was calculated. IVs that showed (Bonferroni (BF) corrected) significant evidence of heterogeneity between sexes were excluded (*p* < 0.05/[number of IVs]). After filtering IVs with significant sex-heterogeneity, phenotypes were again filtered to only those with at least five valid IVs remaining. Fourth, dietary phenotypes were removed due to high correlation amongst these phenotypes (due to the shared household), insufficient power, problems with reverse causation and difficult interpretation26. Finally, we manually removed several duplicated and redundant phenotypes. Specifically, all left-side body traits (highly correlated with right-side) were removed and we also retained only one of the duplicated phenotypes for BMI and weight (retaining data fields 21001 and 21002, respectively). Additionally, all “qualifications” data was removed (corresponding to field [6138](https://biobank.ctsu.ox.ac.uk/showcase/field.cgi?id=6138)) due to the availability of finer-scale correlated variables, such as “age completed full time education” (data field 845). After this process, 118 phenotypes remained for analysis (see Supplementary Figure 3).

Estimation of single-trait causal effects in couples

To investigate the causal effect of a trait in one individual on the same trait of their partner, we performed a couple-specific MR analysis. Specifically, the trait in the index case was used as the exposure, and the same trait in the partner was used as the outcome trait. The effect of genetic variants on the exposure were obtained from the Neale lab summary statistics, using the full UK Biobank sample. Instruments for each trait were selected as described above, i.e. being both genome-wide (GW) significant (*p* < 5x10-8) and pruned for independence. Next, we estimated the effects of SNPs on the outcomes of interest by testing the association between each genetic instrument measured in the index individual with the phenotype measured in the partner using the UKBB partner data set described above. In other words, for each phenotype, the corresponding genetic data for the IVs were obtained from the index case while the phenotypes (dependent variable) were taken from the corresponding partner. All SNP-trait estimates were estimated in males and females separately (i.e. using the sex-specific Neale results or two separate models in the couple data), adjusting for age and the first 40 genetic PCs of both the index and partner. We performed linear regression of SNP effects on phenotypes, regardless of data type (including binary), to mimic the Neale models as closely as possible. Continuous phenotypes were scaled to have mean 0 and SD of 1 before regression, while ordinal and binary phenotypes were left as processed by PHESANT.

To estimate the causal effect of a trait from an index case to a partner (, we combined the effects of genetic instruments on the exposure (from Neale) with effects on the outcomes (measured among couples) in an MR framework using the IVW method (Figure 1a)24. To estimate the causal effects in both sexes combined, SNP-effects were first meta-analysed across sexes using fixed effects models prior to performing MR (rather than meta-analysing the MR estimates directly) to minimize weak instrument bias27. Effects of the genetic estimates on both the exposure and outcome were first standardized (such that the squared effect size represents the explained variance) to allow for seamless comparison across traits and to the raw phenotype correlation. Significance was determined by adjusting for the number of effective tests based on the correlation matrix of phenotypes tested28, resulting in 66 independent tests. The significance threshold was adapted accordingly as *p* < 0.05/66.

After estimating single trait causal effects in couples, we used a two-tailed z-test to identify traits with a significant difference between the MR-estimate and the phenotypic correlation in couples. For each of trait with discrepant estimates, we tested the causal effect of each of the remaining phenotypes in our pipeline () on the focal trait of interest () using MR (). These same-person MR estimates were calculated using meta-analysed sex-specific Neale estimates for both the SNP-exposure and SNP-outcome effects using the IVW-method. Before performing each same-person MR, genetic variants were first filtered for evidence of reverse causality at a threshold of *p* < 0.001 (Steiger filter)29, whereby SNPs were removed if the standardized SNP effect with the outcome was stronger than the effect with the exposure based on a one-tailed t-test at a significance level of p < 0.001. SNP-effects were standardized prior to calculating MR effects.

We then explored those potential confounders, , with a significant impact on (*p* < 0.05/66). As the confounding impact of each involves a within couple effect (), as illustrated in Figure 2, we further filtered the remaining traits, to those with a significant within couple MR effect (*p* < 0.05/[number of remaining ). After identifying potential confounder traits, we combined these with the within couple causal effect (), and calculated the correlation due to confounding as to determine the contribution each trait () confounds the within couple correlation for trait (). We subsequently calculated the ratio of this correlation () and the correlation of in partners a .

**Figure 1:** **Mendelian randomisation schematic within couples.** **A** illustrates the causal effect among couples with a single trait (, where represents genetic variant(s), represents a single trait (in an index and a partner), and represents confounding factors which are not associated with genetic variance owning to the random distribution of alleles at conception. **B** represents the expended causal network involving two traits and the various estimated causal paths from an index case () to a phenotype in the partner () given by , , and . Causal effects from to () can be summarized by three possible (non-independent) scenarios: (1) could exert a causal effect on , followed by having a causal effect on in the partner alone (); (2) the reverse could occur whereby has a causal effect on in the index alone, followed by a causal effect of case on (); (3) there could be other mechanisms, either acting directly or through other unmeasured/considered variables. To quantify , we first estimated the causal effect of on in multivariable MR (not illustrated), to exclude any residual effect of on phenotype from index to partner. These three scenarios could also act in some combination. Therefore, the estimate would capture the paths of , and other mechanisms combined.

Diagram

Description automatically generated

**Figure 2:** DAG illustrates the impact a confounder (trait ) could have on the phenotypic correlation between partners for a given trait (). Correlation due to confounding can be calculated as .

Assessing the role of confounders on trait correlation in couples

We sought to explore the impact of potential confounders on mate-choice by calculating the trait correlations between partners that are due to confounding. We considered the impact of the following confounders () on the trait correlations of the remaining traits 117 traits in the pipeline: average household income, age completed full-time education, sports club or gym user, and North and East birth place coordinates (data fields 738, 845, 6160, 129, and 130 respectively). Using the single-trait causal effects in couples and the same-person MR-estimates, correlation due to founding was calculated for each pair ( as (Figure 2). These confounding estimates were finally contrasted to the actual couple correlation values to explore the extent that each may confound couple correlations by examining the ratio between the two estimates (i.e *)*). Birthplace coordinates (east-west and north-south) were considered together, and their invoked trait correlations were summed up, as they are orthogonal by definition.

Investigating the effect of time and age on correlations and causal relationships in couples

Trait similarity in couples can be driven by both mate choice and/or trait convergence over time spent together. To tease out the contribution of these different sources, we explored whether the cross-partner causal effects change as a function of the length of the relationship and age. The length of relationship for each couple was proxied by the minimum value in the two partners of “length of time at current address” (data field 699). To estimate the effect of age, we took the median age of couples. For each of the two derived variables for each couple (length of relationship and median age), we split the couples into five roughly equal sized bins (using the “smart\_cut” function from the cutr R package). We first estimated the phenotypic correlation of each trait, within couples of each bin. Next, for each single-trait MR described above, analyses were run in the full sample as well as in the different bins. Of the significant results identified in the sex-combined analysis above, we tested to see if there was any significant difference in MR-estimates amongst the bins. Binned MR-estimates were computed using SNP-outcome effect estimated in each bin separately, and the SNP-outcome effects used the same SNP-exposure effects from Neale. Analyses were run in each sex separately and combined (meta-analysed at the SNP level). As above, SNP effects were standardized prior to calculating MR estimates. To assess for the presence of a trend across bins, we tested the significance of the slope of a linear model of bin-specific correlations and bin-specific MR-estimates, inversely weighted by the SE, versus the  bin centre (more precisely, the median  age or time-spent-together for the given bin). Multiple testing was, as described above, adapted based on the effective number of tests, restricted to traits which showed significant causal effects in the joint (both sexes combined), non-binned MR.

Estimation of two-trait causal effects in couples

Using the same process described above involving a single trait, we also sought to investigate causal effects within couples involving two traits (). In other words, two different traits were used as exposure (in the index individual) and outcome (in the partner) to determine the causal effects in couples involving different traits (for example, effect of education in index case on BMI in the partner). Here, the number of tests could be as high as the number of traits squared, but we only considered trait combinations with phenotypic correlation < 0.8 (estimated in the entire UKBB, N = 337,138), in order to avoid too closely related traits. The same set of SNPs were used as in the same-person MR (i.e. first filtered for the presence of reverse causality). As in the single trait MR, SNP-exposure effects were obtained from the Neale summary statistics and SNP-outcome effects were estimated in the couple derived dataset. MR models were run in both sexes separately and jointly (meta-analysing the SNP effects before performing MR analyses). Significance was determined based on the effective number of tests (calculated previously as 66), square, for each pair (i.e. *p* < 0.05/[662]), and highly correlated trait pairs were also removed.

Comparison of paths from index to partner

There are several independent paths through which a trait in an index case could exert a causal effect on a trait in the partner, and we wanted to explore if one path was more dominant, in general, and if there was evidence for the presence of other traits involved. Restricting to only BF-significant trait pairs (with phenotypic correlation < 0.8) from the couple MR, we sought to explore the various paths from a phenotype in an index case () to a phenotype in the partner () as illustrated in Figure 1b. Logically, is less likely to have a direct effect on another , with the exception of exposure traits that directly alter the environment of their partner, such as smoking creating the presence of second-hand smoke. For instance, increased BMI in an index case is not expected to directly increase cardiovascular disease risk in their partner, but more probably to act first on BMI itself in the partner. To explore whether this intuition holds, we dissected the causal effect from to () into three possible (non-independent) mechanisms. First, could exert a causal effect on , followed by having a causal effect on in the partner alone (). Second, the reverse could occur whereby has a causal effect on in the index alone, followed by a causal effect of case on (). Third, there could be other mechanisms, either acting directly or through other unmeasured/considered variables. These three scenarios could also act in some combination. In this way, the estimate would capture the paths of , and other mechanisms combined.

Using the same-person MR estimates () that were calculated as described above, we estimated and representing the various paths from to . To quantify , the single-trait couple causal effect estimate (i.e. from the regression ) were multiplied by the same-individual causal estimate (i.e. from ). To quantify , we first estimated the causal effect of on in multivariable MR (MVMR), to exclude any residual effect of on phenotype from index to partner. Specifically, was used as the independent variable with both and as independent variables (i.e. the MVMR was ). We included both IVs from and , pruned for independence (performed in PLINK with the options --clump-kb 10000 and --clump-r2 0.001 using the 1000 Genomes European samples as a reference). We took the coefficient of as the direct causal effect from to () and multiplied this by the same-individual causal estimate (). Finally, we estimated directly from our two-trait couple MR framework (). We compared the estimates of , , and using a z-test to assess their difference and assessed their relationship using linear regression with the intercept forced through the origin. Finally, we assessed the proportion of that could not be explained merely by the paths quantified by and . As and are not perfectly independent, potentially due to correlation between and or pleiotropic limitations of MR, we estimated the extent of dependence via the correlation between and across the different trait pairs. To account for the duplicate signals due to this correlation, we removed the effects of from by keeping the residuals from the linear regression . We then estimated the proportion of variance explained () of jointly by and the residualised .

Results

Relationship between causal effects and raw phenotypic correlation in couples

First, we asked whether discrepancies between observational couple correlations and causal effects could be explained by specific confounder traits. For this, a total of 118 phenotypes were selected (based on their elevated correlation between partners and sufficient [> 5] valid IVs suitable for MR analysis) and subsequently tested for a causal effect from index to partner within couples, using MR. We then compared the causal effects obtained using MR with the raw phenotypic correlation observed among couples to identify any traits where the correlation was different than the MR-estimate, using a two-tailed z-test to test for a significant difference between the estimates. Significant differences would be indicative of the presence of confounders (either negative or positive) driving the observed phenotypic correlation. After adjusting for the effective number of tested traits (p < 0.05/66), we identified 43 traits which showed different phenotypic correlation compared to MR-estimate (see Figure 3, Supplementary Table 1). Of these, 3 traits corresponded to larger MR-estimate compared to correlation (time spent watching television, comparative height size at age 10, and overall health rating), while the remaining 36 represented traits with larger correlation compared to MR-estimate. Among these included place of birth, North-coordinate (NC); systolic blood pressure (SBP); height; and various blood cell counts, including basophil count (place of birth NC: correlation () = 0.58, MR-estimate () = 0.33, one tailed p-value for difference (*p*) < 1 x 10-8; SBP: = 0.16, = 0.054, *p* = 1 x 10-8; height: = 0.25, = 0.21, *p* = 6.6 x 10-6; basophil count: = 0.47, = 0.0013, *p* < 1 x 10-8).

Of these 43 traits where couple correlation was significantly different than MR causal estimates, we sought to identify potential confounders which may, in part, explain the discrepant estimates. For the three traits where correlation was less than MR-estimate, we searched for negative confounders (i.e. negative ), and conversely for traits where the correlation was greater than the MR-estimate we searched for positive confounders. We found no potential negative confounder to explain the cases where correlation was smaller than the MR-estimate. On the other hand, for the 40 traits with phenotypic correlation larger than MR-estimate, we found many potential positive confounders. Namely, the mean number of potential confounders from our set of 117 candidates was 22.56, with a maximum of 39, for one trait we did not identify any potential confounders (Supplementary Table 2). For instance, for systolic blood pressure, we identified 29 (correlated) potential confounders which may explain the larger phenotypic correlation as compared to MR effect. These potential confounders included physical activity, BMI, lung fitness measures, overall health rating among many others. For weight, we found 30 potential confounders, including anthropometric traits (such as leg, trunk, arm fat mass), various behavioural traits which are reflective of exercise patterns, such as time spent watching television, walking pace, phone use, among many others. Many of the 40 traits with larger phenotypic correlation compared to MR-estimates included blood cell counts and/or percentages (such as white blood cell (leukocyte) count, neutrophil count, monocyte count and percentage, reticulocyte percentage and count). The potential confounders for these traits were highly overlapping, including physical activity level, anthropometric traits, smoking and health rating. Other notable confounders included measures of physical activity for forced vital capacity; smoking status and fitness measures for basal metabolic rate; and measures of body size for hand grip strength. Finally, for each confounder we calculated the correlation due to confounding (C) for each potential confounder as described above (see Figure 2). We then compared the difference in estimates to the maximum C-value for each trait (Figure 4) since due to the high correlation between confounders hindered the sensible estimation of their cumulative effect.

Chart

Description automatically generated

**Figure 3:** Scatter plot shows the within couple standardized MR-estimates () versus the phenotypic correlation among couples (. A two-tailed z-test was used to test for a significant difference between the estimates. After adjusting for the number of effective tests (p < 0.05/66), 39 significant differences were identified (shown in dark blue), where 3 traits showed larger MR-estimates compared to correlation, and 36 traits showed larger correlation compared to MR-estimates.

Chart, scatter chart

Description automatically generated

**Figure 4:** Scatter plot shows the difference in phenotypic correlation and MR-estimate versus the maximum C for each trait where the phenotypic correlation was greater than the MR-estimate (number of traits = 39). The identity line is shown in black.

Impact of potential confounders on trait correlation in couples

Next, we assessed the impact of potential confounders on trait correlation in couples. Our first observation was that geographical location (using place of birth North/East coordinates) has a negligible impact on phenotypic correlations. However, we found that household income, age completed full time education and physical activity levels (measured using the variable “leisure/social activities: sport club or gym”) had an important confounding impact on raw phenotypic correlation among couples (Figure 4). Specifically, when we calculated the ratio of correlation due to confounding over the raw phenotypic correlation among couples and took the average across all traits tested, we found that the impact was far from non-zero (29.8, 11.6 and 17.1%, respectively) while only 1.2% for birthplace location (Table X).

Diagram

Description automatically generated

**Figure 5:** Figure displays couple correlation due to confounding versus the phenotypic trait correlation among couples for selected potential confounder traits (). For each trait in the pipeline, we tested the contribution four confounder traits (average household income, age completed full-time education, and sports club or gym user, PCs) could impact the phenotypic couple correlation. The couple correlation due to confounding for each trait was calculated for each confounder as . In the case of birthplace coordinates, C-values were summed across the two (independent) North and East coordinates. The identity line is shown in black.

Effect of sex, age and time together causal effects in couples

Among the 118 phenotypes tested, we identified 64 significant causal effects in partners after adjusting for the effective number of tests (*p* < 0.05/66) (see Supplementary Table X). We also examined the Cochran’s heterogeneity Q-statistic to identify traits with high heterogeneity and found no evidence of heterogeneity in the MR-estimates (all *p* > 0.05/66). Of the 64 significant results, we then tested to see if there was any difference between sexes. After adjusting for the effective number of tests among the remaining 60 traits based on their pair-wise correlation matrix (*p* < 0.05/29), no traits showed significant difference between sexes. However, 15 traits showed a nominally significant difference between sexes (*p* < 0.05), as shown in Supplementary Table X), which is 4.7-times higher than expected (*pbinomial* = 7.45x10-8). Applying a paired t-test among these 15 traits revealed that female-to-male MR-estimates are on average larger than male-to-female estimates (*p* = 0.014).

Next, we explored the impact of age and time-spent-together among the 64 significant traits in both males and females separately and both sexes combined, using linear regression to assess the trend across different bins (measured using a linear regression of MR-estimates versus median bin). There were no significant results in the sex-combined results after adjustment for number of effective tests (p < 0.05/29) in both the results binned by age or time-spent-together (proxied by the amount of time at same address).

Finally, we examined the Pearson phenotypic correlation within the different bins and assessed for the presence of a trend, again using linear models (phenotypic correlation versus median bin). After adjusting for number of tests (*p* < 0.05/29), we identified two traits which showed a significant trend across the bins according to time-spent-together, namely body fat percentage and hand grip strength (right). In both cases, the correlation decreased as time-spent-together increased. We found another two traits which were showed a significant trend across the bins by median age, smoking status: previous and aspirin use. In this case, for both phenotypes, the slope increased as age median-age increased (Figure 6 and Supplementary Table X). We found consistent findings using the Spearman correlation (all *p* < 0.05/29).

Chart, scatter chart

Description automatically generated

**Figure 6:** Scatter plots show the phenotypic correlation among couples within different bins. Couples were binned by median age (panel a and b) and time-spent-together (proxied by the time lived at same household, panel c and d).

Identification of underlying mechanisms for cross-trait assortment

We sought to compare three estimated paths from a phenotype in the index case () to another phenotype in its partner () as illustrated in Figure 3. The total causal effect between and (denoted by ) can be split up into three components: 1) assortative mating through (i.e. ) and then a causal effect between and in the partner (i.e. ), their product is denoted by ; 2) causal effect between and in the index individual (i.e. ), followed by assortative mating through (i.e. ), their product is denoted by ; 3) any remaining effect of on .  We computed within-couple cross-trait causal effect estimates (i.e. ) for all combinations of trait pairs (). Of these, we identified 1327 significant MR effects ( < 0.05/[662]) among couples, which was reduced to 1088 pairs after removing those with phenotypic correlation > 0.8 (a summary of a set of pruned traits can be found in Supplementary Table X). Among these, identified several relationships which were almost completely dominated by (assortative mating through the outcome), and others dominated by (assortative mating through the exposure). Specifically, we found causal relationships between partners for leg fat percentage (right) and time spent watching television; body mass index and overall health rating; all dominated by . On the other hand, we found the other causal relationships between partners which were primarily dominated by (assortative mating through the exposure), including: comparative height at age 10 (i.e. “When you were 10 years old, compared to average would you describe yourself as: shorter, taller, average”) and forced vital capacity (FVC); and standing height on hand grip strength. Finally, we found other pairs where neither nor appeared to capture the relationship (i.e. was significantly larger than both estimates), including BMI and SBP/DBP.

We then estimated the contribution of the first two components contributing to these significant cross-trait effects, and , and compared their contribution to the total effect using standard linear regression (Figure 7, Table X: omega\_regression\_estimates). Paired t-test comparing and effect estimates revealed that (assortative mating through ) is stronger (*p* = 1.1 x 10-5) in general compared to (assortative mating through ). When we summed up the effects of and , we found that the sum was significantly larger than . However, these two effects seemed to be correlated, carrying potentially shared signals. Hence, we first residualized for the effects of (), to ensure complete independence between the two estimates, and then added to (). We found no significant difference between and the sum of in this analysis and points in general were near the identity line. Linear regression results revealed that 76% of the total effect () can be explained by the two paths () and that the is on average very close to the total effect.

Chart

Description automatically generated

**Figure 7:** Top panel shows the regression between the various paths from the index case () to another phenotype in its partner () for the 1088 trait pairs with significant MR-effects among couples ( < 0.05/[662]) and had correlation < 0.8. the index case () to another phenotype in its partner (). To calculate , we residualized for the effects of (), to ensure complete independence between the estimates, and then added to (). Lower panel displays a boxplot comparing the coefficients of the estimates among the trait pairs, after remaining 19 trait-pairs where the sign did not match between any combination of the four coefficients.

Discussion

Phenotypic similarity among couples is a well-establish phenomena in the field of epidemiology, most notably among behavioural, cognitive, anthropometric, and cultural/religious traits. Such similarity is predominantly driven by two main factors. First, people tend to choose “mates” with those that are similar to themselves, known as assortative mating. Secondly, and less well-studied, couples may influence each other overtime which induces further phenotypic similarity. However, phenotypic correlations may be subject to confounding and therefore not representative of true causal relationships, but rather indirect due to being correlates of traits under direct mate-choice.

In this report, we sought to investigate causal relationships among couples within the UKBB using MR. We analysed 118 traits, representing a wide range of anthropometric-, behavioural-, and disease-related traits. Only three pairs had significantly larger causal effect than observed correlation, which would suggest a highly unlikely negative confounder. A more probable explanation is incorrect MR estimates due to assumption violations. On the other hand, we identified 40 traits which showed significantly larger phenotypic correlation compared to MR causal estimates. These results suggest that raw phenotypic correlations among couples are likely subject to positive confounding due to correlation with other traits. Indeed, we identified many potential confounders of these relationships, BMI, overall health rating, fitness measures, income, smoking, to name a few. Together these results suggest phenotypic correlations among couples may be capturing other important aspects of AM including measures of socioeconomic status and various measures of health and fitness.

When we investigated the impact of common confounders on our entire panel of phenotypes (i.e. fixing a confounder and assessing its widespread impact on all single-trait AM), we found that household income, age completed education, participant of a sport club or gym are indeed important confounders of observational phenotypic correlations among couples, explaining on average 29.8, 11.6 and 17.1% of the phenotypic couple correlations among traits tested, respectively. These results also suggest that phenotypic correlations in couples are indeed significantly confounded and point to a relatively few key traits which are driving AM observations.

Among the 118 phenotypes tested, we found widespread evidence of causal effects among partners. In particular, we identified 64 same-trait causal effects within partners (out of 118 traits), and we also found no evidence of heterogeneity among same-trait couple MR estimates (). This suggests that associations between the index genotype and partner’s phenotype are primarily acting indirectly through the causal relationship between the traits, rather than the presence of a direct effect for index genotype to the partner’s phenotype. If we assume that genetic effects to partner traits can only happen via first altering a trait of the index case, pleiotropic instruments would only emerge from indirect genetic effects (through another trait), which could be tested and excluded via pheWAS.

We found that in general, female-to-male estimates were stronger than male-to-female estimates. These results point to AM being stronger among females compared to males. While no individual trait yielded a detectable sex-difference, among the 64 traits, we found an almost 5-fold enrichment for nominally significant sex-different causal effects than expected by chance. When we investigated phenotypic correlations across different ages and amount of time-spent-together, our results suggest that fitness and anthropometric measures are important initially, however the correlation decreases with time. This phenomenon may reflect that these traits are important during mate-choice, but the longer people stay together the less important it becomes to stay similar in those aspects. On the other hand, we found that smoking and medication use (aspirin, specifically) become more concordant among couples as age increases. As age and time-spent-together are highly correlated variables, it is difficult to distinguish whether this is an effect of convergence or suggestive of an age-dependent mate-choice. We did not identify any significant trends when we investigated the impact of causal effects by time-spent-together or by age. While this could be due to limitations such as statistical power, this is consistent with previous reports which suggest that initial mate choice is a more dominant factor in contributing to phenotypic similarity compared to convergence7,30–32.

Our findings investigating cross-trait assortment suggest that causal effects from to are primarily driven by assortative mating through (i.e. ) followed by a causal effect within the partner from to in the partner (i.e. ). In other words, among trait pairs which show a significant causal effect in couples, the exposure is passed from the index to the partner before a causal effect from to in the partner. In contrast, the less likely path would be the inverse, whereby the presence of a causal effect from to in an index case is then followed by being passed directly from index to partner. This result was expected, as it is more reasonable for couples to influence each other at the exposure level rather than the outcome level. Furthermore, our results revealed that the majority of the effect from to goes through assortative mating through and ( and , but primarily ) rather than directly from to or through another (third) variable.

We found 1088 significant cross-trait causal effects within couples. In particular, general these cross-trait effects (given by ) can be summarized by three categories: (1) driven by assortment on the exposure () (2) driven by assortment on the outcome (), and (3) not explained by either (i.e. being greater than both and ). Of note, there were fewer cases in category three, where the causal effect from to was not captured by or , suggestive of either a direct effect to or the presence of a confounder variable. As an example from the first group, we found a positive causal effect of time spent watching television on BMI is driven by the fact that partners causally influence each other with respect time spent watching television which in turn has an impact on BMI at the individual level. On the other hand, representing the second category, we found a positive causal relationship from height to education, with a stronger path through , representing a path whereby height (as a proxy for “dynastic” wealth) increases educational attainment (found previously33) within a single individual, and AM then occurs via education level. Finally, as an example of an effect in category three, we found a negative causal effect of being a never smoker on white blood cell leucocyte count within partners, in other words leucocyte count was higher among individuals with partners who smoked. While we also identified a significant effect through (AM through smoking), the effect was much stronger through . These findings suggest that there could be a direct effect from index partner by way of second-hand smoke. These results are consistent with previous work showing higher WBC count in smokers34, which might already be achieved by second-hand smoking.

This study has limitations which should be taken into account. First, with the current data, we were not able to find strong evidence for couple convergence over time. We did make use of both age and time-together data (proxied by time at the same address) to help shed light on this question, and were able to show that certain traits indeed appear to converge as a couple spends more time together while other traits appear to be more important in the selection process (i.e. true assortative mating). However, to properly assess the question, longitudinal data including measures before couples were together would be best suited to disentangle the complex relationship between assortative mating and convergence. Secondly, while assortative mating through the exposure () and the outcome ( represent independent paths from to , our results suggest that the computed effects using MR estimates are not perfectly independent. This could potentially be due to overlap in genetic instruments, bidirectional causal effect between them or the fact that both estimates depend on the causal from *X* to *Y*. To the best of our ability, we tried to mitigate this bias. In the calculation of rho, we used a MVMR approach to remove effects of X on Y. Also, when summing gamma and rho we first residualised gamma for effects of rho to ensure independence. Finally, to increase statistical power and robustness, we limited our traits to those with significant correlation amongst couples and >5 valid IVs (among a few other filtering criteria). As a result, anthropometric traits constituted a larger proportion of our traits under study and represent a large percentage of our significant findings. Other phenotypes, such as behavioural and lifestyle traits, were included where possible but, in general, with less statistical power due to both lower couple correlation and less IVs. Additionally, we were limited to the available traits and white British samples in the UKBB. Assortative mating is highly population-specific; hence our findings are not necessarily generalisable to other populations.

In summary, we have surveyed a large number of complex traits with significant couple correlation in the UK Biobank and explored to which extent the observed couple similarity is due to couple convergence or confounding (i.e. assortment for correlated phenotypes). We have also demonstrated that cross-trait assortment can largely be explained by single-trait assortments between either trait and substantial causal effects between these traits. Our findings provide new insights into possible mechanisms underlying observed assortative mating patterns at an unprecedented scale and resolution.

References

1. Silventoinen, K., Kaprio, J., Lahelma, E., Viken, R. J. & Rose, R. J. Assortative mating by body height and BMI: Finnish twins and their spouses. *Am. J. Hum. Biol.* **15**, 620–627 (2003).

2. Maes, H. H., Neale, M. C. & Eaves, L. J. Genetic and environmental factors in relative body weight and human adiposity. *Behav. Genet.* **27**, 325–51 (1997).

3. Keller, M. C. *et al.* The Genetic Correlation between Height and IQ: Shared Genes or Assortative Mating? *PLoS Genet.* **9**, (2013).

4. Mare, R. D. Five Decades of Educational Assortative Mating. *Am. Sociol. Rev.* **56**, 15–32 (1991).

5. Agrawal, A. *et al.* Assortative mating for cigarette smoking and for alcohol consumption in female Australian twins and their spouses. *Behav. Genet.* **36**, 553–566 (2006).

6. Buss, D. M. Marry Someone Who Is Similar To Us in Almost Every Variable. *Am. Sci.* **73**, 47–51 (1985).

7. Watson, D. *et al.* Match makers and deal breakers: Analyses of assortative mating in newlywed couples. *J. Pers.* **72**, 1029–1068 (2004).

8. Hippisley-Cox, J. Married couples’ risk of same disease: cross sectional study. *BMJ* **325**, 636–636 (2002).

9. Buss, D. M. *et al.* International Preferences in Selecting Mates. *J. Cross. Cult. Psychol.* **21**, 5–47 (1990).

10. Buss, D. M. & Barnes, M. Preferences in Human Mate Selection. *J. Pers. Soc. Psychol.* **50**, 559–570 (1986).

11. Anderson, C., Keltner, D. & John, O. P. Emotional Convergence Between People over Time. *J. Pers. Soc. Psychol.* **84**, 1054–1068 (2003).

12. Gonzaga, G. C., Campos, B. & Bradbury, T. Similarity, Convergence, and Relationship Satisfaction in Dating and Married Couples. *J. Pers. Soc. Psychol.* **93**, 34–48 (2007).

13. Humbad, M. N., Donnellan, M. B., Iacono, W. G., McGue, M. & Burt, S. A. Is spousal similarity for personality a matter of convergence or selection? *Pers. Individ. Dif.* **49**, 827–830 (2010).

14. Risch, N. *et al.* Ancestry-related assortative mating in Latino populations. *Genome Biol.* **10**, (2009).

15. Sebro, R., Peloso, G. M., Dupuis, J. & Risch, N. J. Structured mating: Patterns and implications. *PLoS Genet.* **13**, 1–22 (2017).

16. Abdellaoui, A. *et al.* Genetic correlates of social stratification in Great Britain. *Nat. Hum. Behav.* **3**, 1332–1342 (2019).

17. Yengo, L. *et al.* Imprint of assortative mating on the human genome. *Nat. Hum. Behav.* **2**, 948–954 (2018).

18. Border, R. *et al.* Assortative mating biases marker-based heritability estimators. *Nat. Commun.* **13**, (2022).

19. Border, R. *et al.* Cross-trait assortative mating is widespread and inflates genetic correlation estimates. *bioRxiv* 2022.03.21.485215 (2022) doi:10.1101/2022.03.21.485215.

20. Robinson, M. R. *et al.* Genetic evidence of assortative mating in humans. *Nat. Hum. Behav.* **1**, 0016 (2017).

21. Xia, C., Canela-Xandri, O., Rawlik, K. & Tenesa, A. Evidence of horizontal indirect genetic effects in humans. *Nat. Hum. Behav.* **05**, 399–406 (2020).

22. Lawlor, D. A. *et al.* Mendelian randomization: using genes as instruments for making causal inferences in epidemiology. *Stat. Med.* **27**, 1133–1163 (2008).

23. Howe, L. J. *et al.* Genetic evidence for assortative mating on alcohol consumption in the UK Biobank. *Nat. Commun.* **10**, (2019).

24. Burgess, S., Butterworth, A. & Thompson, S. G. Mendelian randomization analysis with multiple genetic variants using summarized data. *Genet. Epidemiol.* **37**, 658–665 (2013).

25. Millard, L. A. C., Davies, N. M., Gaunt, T. R., Smith, G. D. & Tilling, K. Software application profile: PHESANT: A tool for performing automated phenome scans in UK Biobank. *Int. J. Epidemiol.* **47**, 29–35 (2018).

26. Pirastu, N. *et al.* Using genetic variation to disentangle the complex relationship between food intake and health outcomes. *bioRxiv* 829952 (2020) doi:10.1101/829952.

27. Burgess, S., Small, D. S. & Thompson, S. G. A review of instrumental variable estimators for Mendelian randomization. *Stat. Methods Med. Res.* **26**, 2333–2355 (2017).

28. Gao, X., Starmer, J. & Martin, E. R. A multiple testing correction method for genetic association studies using correlated single nucleotide polymorphisms. *Genet. Epidemiol.* **32**, 361–369 (2008).

29. Hemani, G., Tilling, K. & Davey Smith, G. Orienting the causal relationship between imprecisely measured traits using GWAS summary data. *PLOS Genet.* **13**, e1007081 (2017).

30. Mascie-Taylor, C. G. N. Spouse similarity for IQ and personality and convergence. *Behav. Genet.* **19**, 223–227 (1989).

31. Caspi, A., Herbener, E. S. & Ozer, D. J. Shared experiences and the similarity of personalities: a longitudinal study of married couples. *J. Pers. Soc. Psychol.* **62**, 281–91 (1992).

32. Yengo, L. *et al.* No Evidence for Social Genetic Effects or Genetic Similarity Among Friends Beyond that Due to Population Stratification: A Reappraisal of Domingue et al (2018). *Behav. Genet.* **50**, 67–71 (2019).

33. Tyrrell, J. *et al.* Height, body mass index, and socioeconomic status: Mendelian randomisation study in UK Biobank. *BMJ* **352**, (2016).

34. Pedersen, K. M. *et al.* Smoking and Increased White and Red Blood Cells: A Mendelian Randomization Approach in the Copenhagen General Population Study. *Arterioscler. Thromb. Vasc. Biol.* **39**, 965–977 (2019).