

---

## **Supplementary information**

---

# **Evidence of horizontal indirect genetic effects in humans**

---

In the format provided by the  
authors and unedited

# **Evidence of horizontal indirect genetic effects in humans**

Charley XIA<sup>1</sup>, Oriol CANELA-XANDRI<sup>1,2</sup>, Konrad RAWLIK<sup>1</sup>, Albert TENESA<sup>1,2,3,\*</sup>

<sup>1</sup> The Roslin Institute, The Royal (Dick) School of Veterinary Studies, University of Edinburgh, Edinburgh, UK;

<sup>2</sup> MRC IGMM, Western General Hospital, University of Edinburgh, Edinburgh, UK;

<sup>3</sup> Usher Institute, NINE Edinburgh BioQuarter, University of Edinburgh, Edinburgh, UK;

\* Corresponding author: Albert Tenesa (albert.tenesa@ed.ac.uk; <https://orcid.org/0000-0003-4884-4475>)

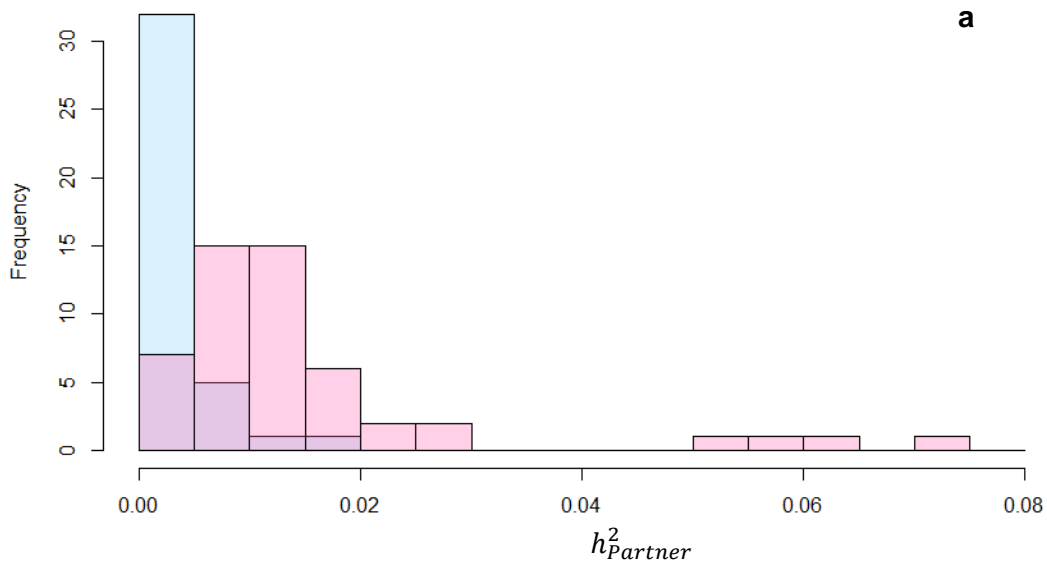
# Table of Contents

Supplementary Figures .....	4
Supplementary Figure 1. Estimate of partner indirect heritability and its comparison to own heritability. ....	4
Supplementary Figure 2. Comparing GREML results of the full model and the reduced model for 51 traits with significant $h^2_{partner}$ estimate ( $P < 0.05$ ). ....	5
Supplementary Figure 3. Testing permutation strategy using simulated data .....	6
Supplementary Figure 4. Comparing GREML results in randomised couples with real couples...8	
Supplementary Figure 5. Comparing $h^2_{partner}$ estimates from our models with household model.9	
Supplementary Tables .....	10
Supplementary Table 1. Descriptive summary for the phenotype. ....	10
Supplementary Table 2. Prevalence of binary traits in UK Biobank. ....	10
Supplementary Table 3. GREML results for non-sex-specific traits. ....	10
Supplementary Table 4. Prediction accuracy before and after modelling partner indirect genetic effects. ....	10
Supplementary Table 5. GREML results for sex-stratified and sex-specific analysis. ....	10
Supplementary Table 6. Phenotypic correlation between partners in European couples and permutations. ....	10
Supplementary Table 7. Comparing estimate of partner indirect heritability with its expected value under assortative mating.....	10
Supplementary Table 8. GREML results for 51 non-sex-specific traits after permutation.....	10
Supplementary Table 9. Couple and sibling replication results for 51 non-sex-specific traits. ....	10
Supplementary Table 10. Meta-analysis of all follow-up analyses. ....	10
Supplementary Table 11. Household model results for selected non-sex-specific traits .....	10
Supplementary Methods .....	11
Summary .....	11
Mathematical Derivation .....	12
Supplementary Figure 6. Diagram of our models.....	12
Supplementary Table 12. Strength of assortment for each case.....	13
Generic equations.....	13
Relations between phenotype, environment and genetics between partners under our models .....	13
Supplementary Table 13. The between traits between partners correlation under our models. ....	15
Derivation of $r_{cp}$ .....	15
Supplementary Table 14. Expectation of $r_{cp}$ under our models. ....	16
Derivation of $h^2_{own}$ .....	16
Supplementary Table 15. Expectation of $h^2_{own}$ under our models. ....	17
Derivation of $h^2_{partner}$ .....	17

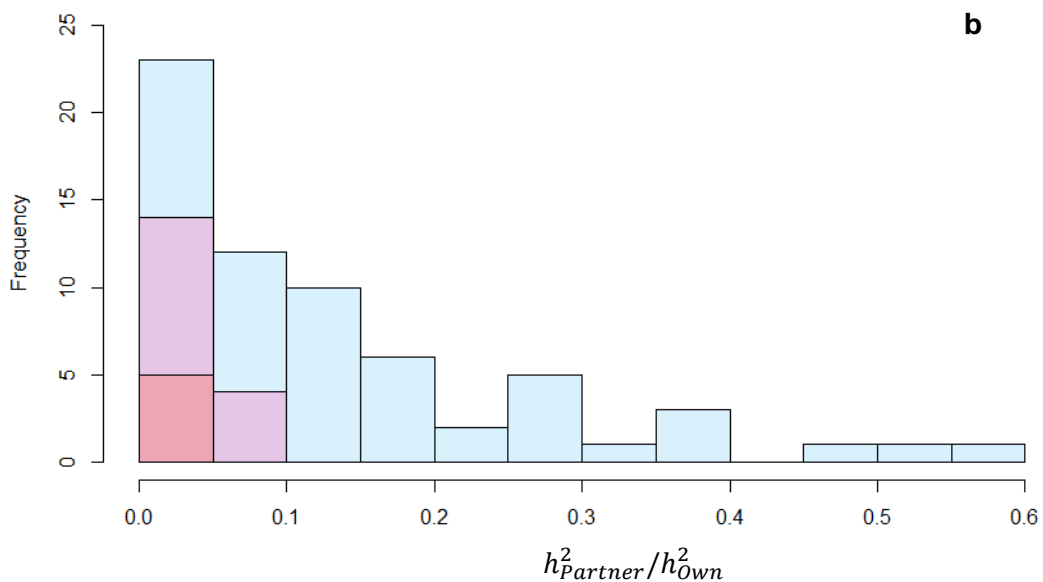
Supplementary Table 16. Expectation of $h_{partner}^2$ under our models. ....	18
Derivation of the ratio between $h_{partner}^2$ and $r_{cp}^2 h_{own}^2$ .....	18
Supplementary Table 17. Expectation of the ratio between $h_{partner}^2$ and $r_{cp}^2 h_{own}^2$ under our models. ....	18
Simulation .....	18
Validation of Analytical results.....	18
Supplementary Figure 7. Comparing the observed value from simulated data with its expected value under our model. ....	19
Simulation Details .....	20
Supplementary Results .....	23
Expectation of Results under Realistic Parameters .....	23
Supplementary Table 18: Ranges of relevant parameters observed across 51 non-sex-specific traits with significant estimates of $h_{partner}^2$ ( $P < 0.05$ ) in UK Biobank. ....	23
Supplementary Table 19. Selected values for each parameter based on UK Biobank observed values. ....	24
Results with no IGEs.....	24
Supplementary Table 20: Summary of the two scenarios not involving IGEs.....	24
Supplementary Figure 8. Expected values of $h_{partner}^2$ (y-axis) against $r_{cp}^2 h_{own}^2$ (x-axis) in the two scenarios not involving IGEs for the realistic range of model parameters.....	25
Results including IGEs.....	26
Supplementary Table 21. Summary of scenarios involving IGEs.....	27
Supplementary Figure 9. Plots of expected values of $h_{partner}^2$ against $r_{cp}^2 h_{own}^2$ in the eight scenarios involving IGEs for the realistic range of model parameters which covers the 1 <sup>st</sup> to 3 <sup>rd</sup> quartiles of each parameter observed in UK biobank.....	28
Supplementary Figure 10. Plots of expected values of $h_{partner}^2$ against $r_{cp}^2 h_{own}^2$ in the eight scenarios involving IGEs for the realistic range of model parameters which covers the minimum and the maximum of each parameter observed in UK biobank. ....	29
Supplementary Reference .....	30

## Supplementary Figures

### Supplementary Figure 1. Estimate of partner indirect heritability and its comparison to own heritability.

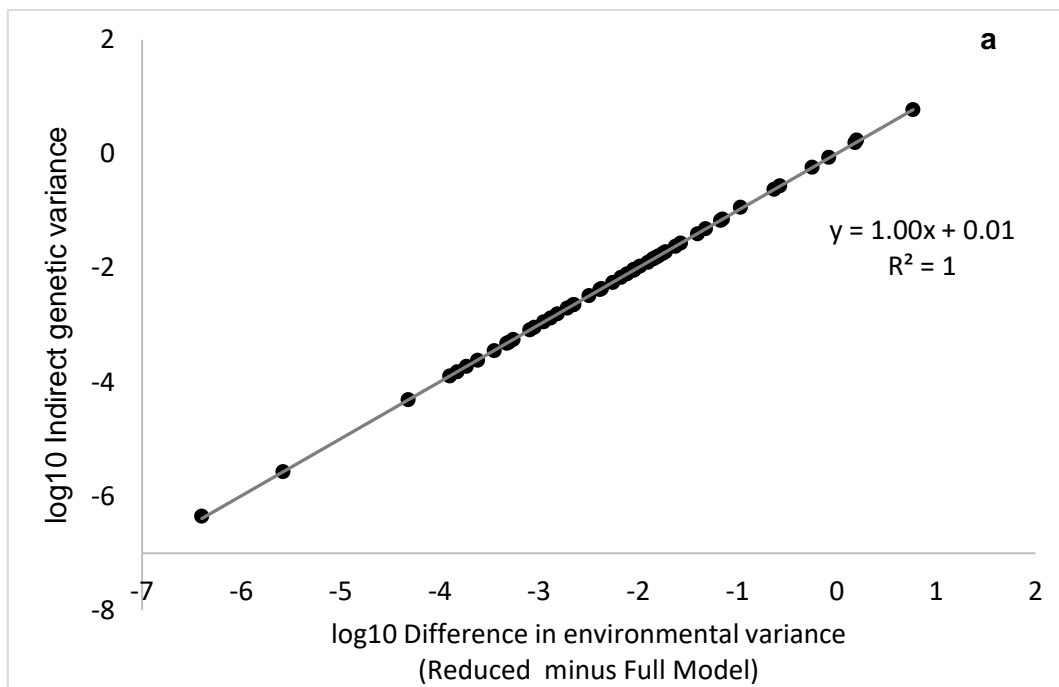


Panel **a** shows the histogram of estimates of partner indirect heritability across all 91 traits expressed and measured in both sexes. **Pink bars**: Estimates are significantly greater than 0 ( $P < 0.05$ ) and the mean of these estimates is  $\sim 0.015$ . **Blue bars**: Estimates are non-significant ( $P \geq 0.05$ ).

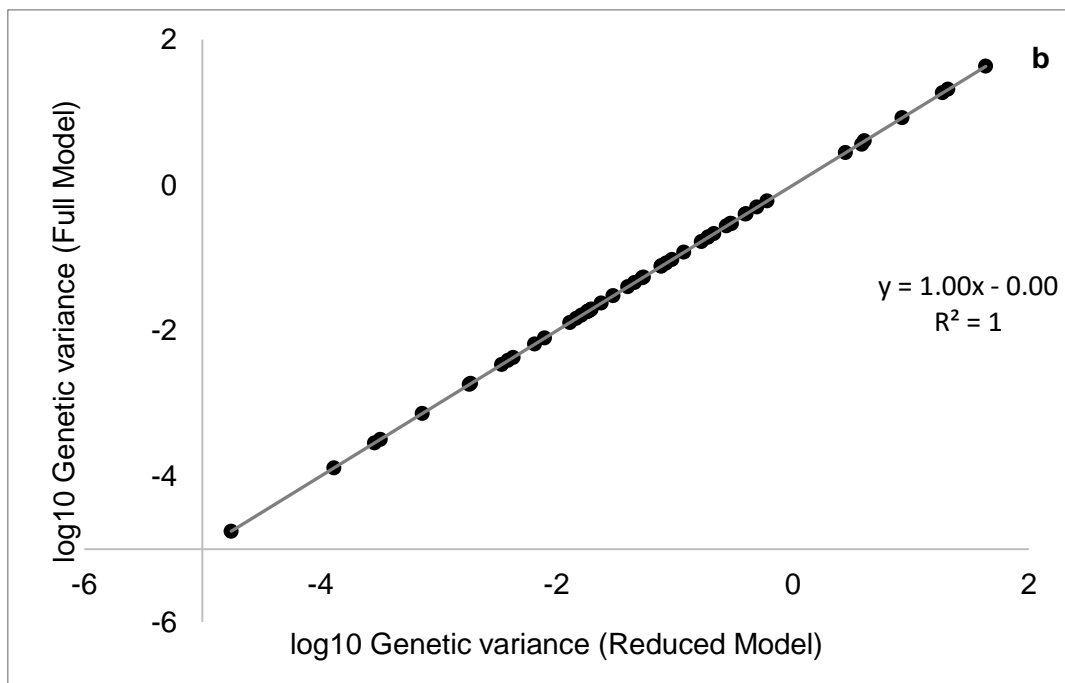


Panel **b** shows the histogram of ratio of partner indirect heritability to own heritability across all 51 traits which estimates of partner indirect heritability are greater than 0 ( $P < 0.05$ ). **Blue bar**: traits with own heritability lower than 0.15. **Pink bar**: traits with own heritability between 0.15 to 0.3. **Red bar**: traits with own heritability greater than 0.3. This plot indicates that partner indirect heritability is relatively large, compared to own heritability, for low heritability traits.

**Supplementary Figure 2. Comparing GREML results of the full model and the reduced model for 51 traits with significant  $h^2_{partner}$  estimate ( $P < 0.05$ ).**

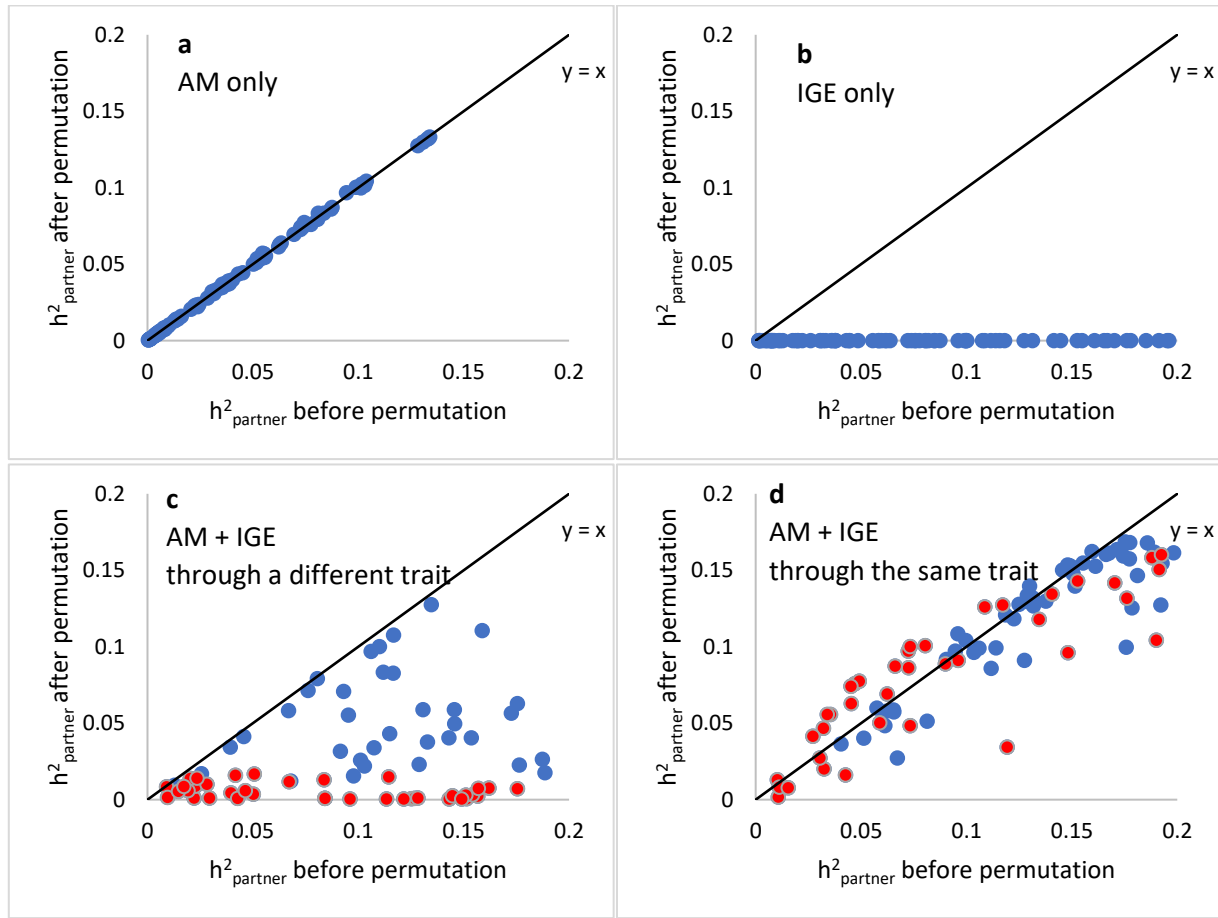


Panel **a** shows the comparison of the variance explained by the partner genotype in the full model and the difference in the environmental variance between the reduced and the full model. This plot shows that the partner genotype mainly removes variance from the individual's environment.



Panel **b** shows the comparison of the variance explained by one's own genotype in the full model and that in the reduced model. This plot shows that fitting the partner genotype in the model has little influence on the estimate of one's own additive genetic variance.

### Supplementary Figure 3. Testing permutation strategy using simulated data



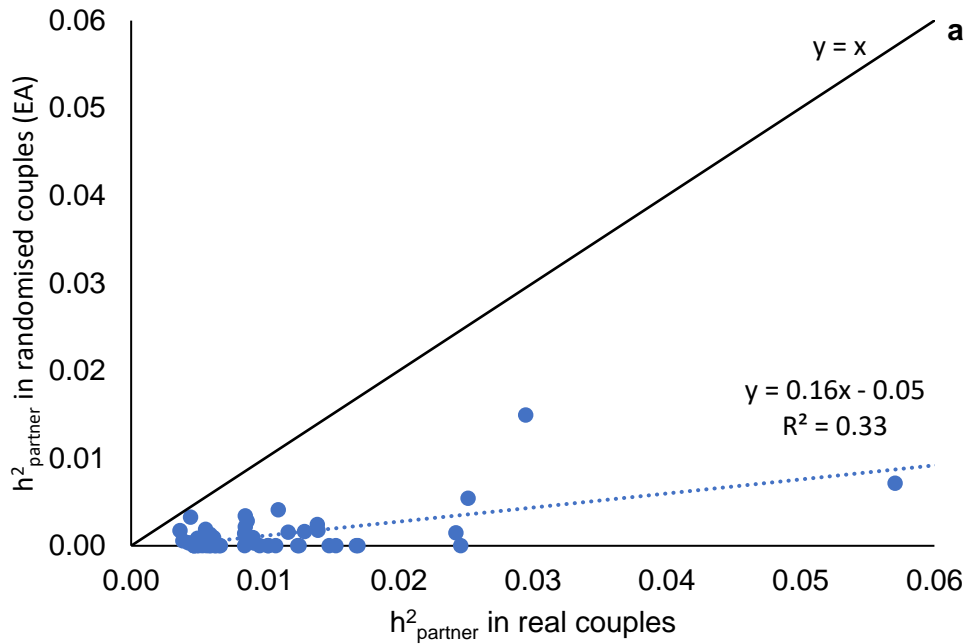
We tested how the estimate of partner heritability would change in randomised couples across different scenarios (AM, IGE and AM + IGE) using simulated data. 100 replications for each scenario. For each replication, we simulated up to two polygenic traits (one being the focal trait which might be under assortment and the other being the trait mediates IGEs) in 10,000 couples. We estimated partner heritability in the original 10,000 couples, randomised these couples using the same permutation strategy as in the manuscript (i.e. swapping partners within block of 10 based on the phenotype) and re-estimated the partner heritability in the randomised couples. The genetic variance, environmental variance, strength of assortment and IGE were simulated to be any value from 0.1 to 0.9 at interval of 0.1, chosen at random. More details on how to simulate traits and assortment see Supplementary Notes: Simulation Details. Red dots in panel **c** and **d** refer to traits under assortment with strength lower than 0.5.

Results indicate that, when couples are only affected by assortative mating (panel **a**), estimate of partner heritability is the same before and after permutation. In the presence of IGEs and when IGEs are mediated through a trait that is not correlated with the focal trait (panel **b** and **c**), we would

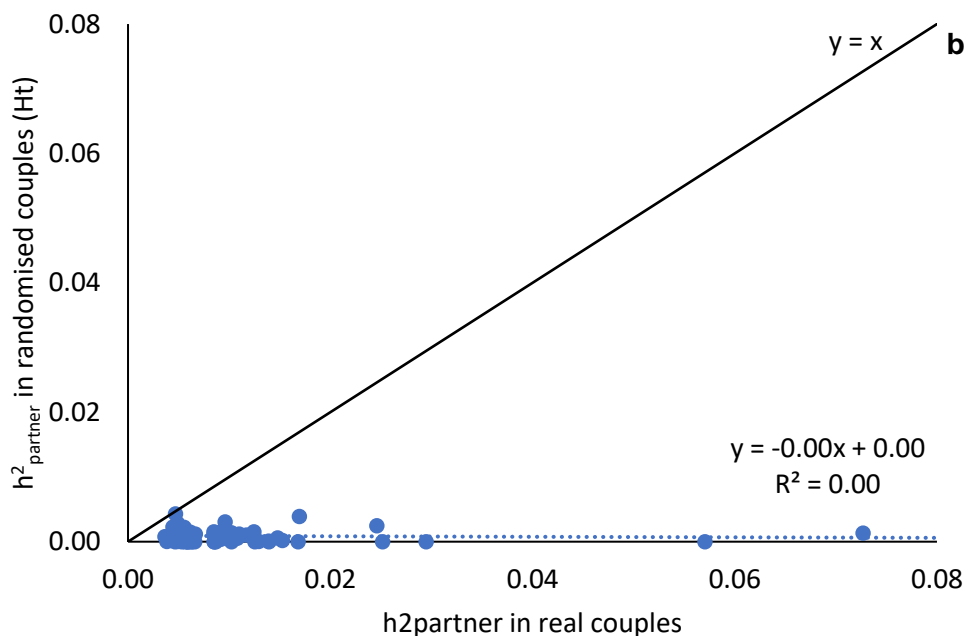
observe a drop in partner heritability in randomised couples comparing to the original couples. Since the strength of assortment is not large in real data (i.e.  $<0.5$ ) (red dots), the reduce in estimate is quite significant. However, when IGEs are mediated through a trait that is correlated with the assortative mating trait (e.g. the same trait) (panel **d**), the pattern of how partner heritability would change after permutation is unclear and the difference in partner heritability before and after permutation is relatively smaller. Therefore, our permutation method is hard to detect IGEs and separate them from direct assortment in this case (panel **d**).



**Supplementary Figure 4. Comparing GREML results in randomised couples with real couples.**

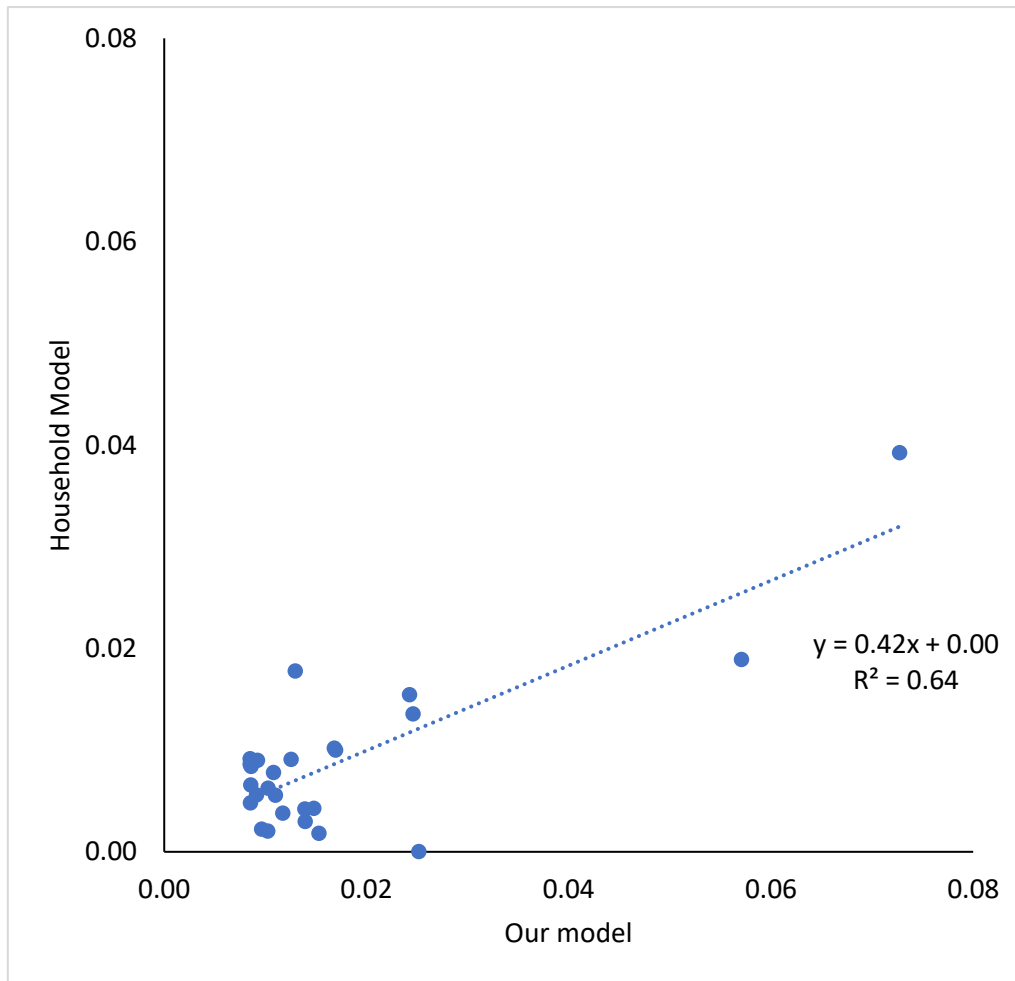


Panel **a** shows the comparison of the  $h^2_{\text{partner}}$  estimated in real couples (X-axis) and in randomised couples permutated based on couple correlation of educational attainment (EA) (Y-axis) for 50 traits which show significant  $h^2_{\text{partner}}$  estimates ( $P < 0.05$ ). On average,  $h^2_{\text{partner}}$  estimates from randomised couples are ~16% of those from real couples.



Panel **b** shows the comparison of the  $h^2_{\text{partner}}$  estimated in real couples (X-axis) and in randomised couples permutated based on couple correlation of standing height (Ht) (Y-axis) for 50 traits which show significant  $h^2_{\text{partner}}$  estimates ( $P < 0.05$ ). On average,  $h^2_{\text{partner}}$  estimates from randomised couples are ~0% of those from real couples.

**Supplementary Figure 5. Comparing  $h^2_{partner}$  estimates from our models with household model.**



We applied household model (our model plus a shared couple environment as a random effects) for 28 traits of which their  $h^2_{partner}$  estimates from our model (Supplementary Table 3) were significantly greater than 0 after Bonferroni correction ( $P < 0.05/91$ ). 2 traits failed to converge for the model and for the rest 26 traits, we found that the  $h^2_{partner}$  estimates from the household model (y-axis) are ~60% smaller than those from our model (x-axis) on average. Dotted line: fitted line.

## **Supplementary Tables**

Supplementary Tables 1-11 are provided in an additional excel file.

**Supplementary Table 1. Descriptive summary for the phenotype.**

**Supplementary Table 2. Prevalence of binary traits in UK Biobank.**

**Supplementary Table 3. GREML results for non-sex-specific traits.**

**Supplementary Table 4. Prediction accuracy before and after modelling partner indirect genetic effects.**

**Supplementary Table 5. GREML results for sex-stratified and sex-specific analysis.**

**Supplementary Table 6. Phenotypic correlation between partners in European couples and permutations.**

**Supplementary Table 7. Comparing estimate of partner indirect heritability with its expected value under assortative mating.**

**Supplementary Table 8. GREML results for 51 non-sex-specific traits after permutation.**

**Supplementary Table 9. Couple and sibling replication results for 51 non-sex-specific traits.**

**Supplementary Table 10. Meta-analysis of all follow-up analyses.**

**Supplementary Table 11. Household model results for selected non-sex-specific traits**

## Supplementary Methods

### Summary

We analytically derived the expected relationship between three parameters  $h_{partner}^2$ ,  $h_{own}^2$  and  $r_{cp}$  under direct and indirect assortative mating (AM) and in the presence of indirect genetic effects (IGEs). We then compared the obtained expressions with the experimental data. Our results support two conclusions drawn in the main text:

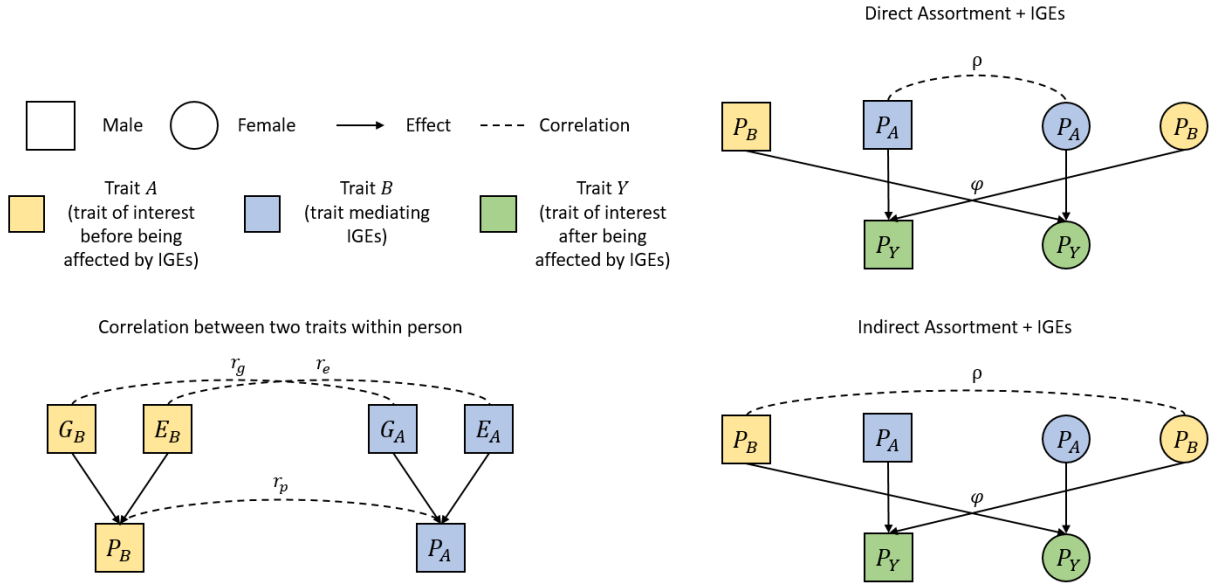
1. The observed data is not consistent with AM on its own.
2. The observed data is consistent with the presence of both AM and IGEs acting through two distinct traits with low or no genetic correlation.

To obtain the analytical expressions, we modelled a general scenario where we assumed that our trait of interest,  $P_Y$ , is a function of two traits. A trait  $P_A$ , which is observable and can be under direct or indirect assortative mating, and a trait  $P_B$ , which is hidden and represents the effects of IGEs. Considering a model with two traits  $P_A$  and  $P_B$  allows us to represent a wide range of different situations involving AM and IGEs and their interaction, with the relationship between the traits varying depending on the assumed scenario. As an example of a situation involving both AM and IGEs, we may consider trait  $P_A$  to be BMI at mate choice which partners assorted on, trait  $P_B$  to be dietary preferences, which affects the partners BMI, and  $P_Y$  is BMI 10 years after mate choice.

Using this model, we first derived analytical expressions for the heritability estimated from the individual's genotype ( $h_{own}^2$ ), the heritability estimated from the partners genotype while adjusting for an individual's own genotype ( $h_{partner}^2$ ) and the phenotypic correlation between partners ( $r_{cp}$ ). We validated the analytical expressions by comparing them with empirical estimates from simulated data (see Supplementary Figure 7, page 19). Finally we considered the expected relationship between  $h_{partner}^2$  and  $h_{own}^2 r_{cp}^2$ , under a realistic range of model parameters (i.e. based empirically on the range of parameters observed in the UK Biobank in Supplementary Table 18, page 24) and compared it with the observed data (Figure 3a in main text). We first conclude that the data are not consistent with the two possible scenarios based on AM without IGEs (see Supplementary Table 20 and Supplementary Figure 8, pages 24-25). Second, we considered an extended set of 8 further scenarios involving IGEs (see Supplementary Table 21, page 27). The two scenarios that best explained the observed data, were those involving both AM and IGEs acting through traits with either no or little genetic correlation (see Supplementary Figure 9, page 28).

## Mathematical Derivation

We considered a model under two general configurations, direct assortative mating on the trait affected by IGEs (e.g. BMI under assortment and affected by dietary preferences) or indirect assortative mating through assortment on the trait mediating IGEs (e.g. social economics status under assortment and affecting disease risk). Our objective is to obtain an analytic expression for  $h_{own}^2$ ,  $h_{partner}^2$  and  $r_{cp}$ , under different assumptions, which we will then compare with empirical results. The models are summarised in Supplementary Figure 6 (below).



**Supplementary Figure 6. Diagram of our models.**

To this end, we define a model with three traits. The trait of interest,  $P_Y$ , and two intermediate traits:  $P_A$  and  $P_B$  which would be the trait affected by IGEs and the trait mediating IGEs, respectively (see Fig. S1). These intermediate traits are composed of a genetic and an environmental component:  $P_A = G_A + E_A$  and  $P_B = G_B + E_B$ . Under our model, the observed phenotype,  $P_Y$ , is modelled as a linear combination of  $P_A$  and  $P_B$ . For example, the phenotypic value of partner 1 would be:  $P_Y^{p1} = P_A^{p1} + \phi P_B^{p2}$ , where  $\phi$  is a parameter weighting the contribution of  $P_B$  (i.e. if  $\phi$  is zero, there will not be IGEs) and  $p1$  and  $p2$  indicate the two partners.

We assume, without loss of generality, that  $\sigma^2(P_A) = \sigma^2(P_B) = 1$ . We furthermore assume the genetic-environmental correlations to be zero, i.e.,  $\text{corr}(G_A, E_B) = \text{corr}(E_A, G_B) = 0$ . Hence, denoting the genetic and environmental correlations between the intermediate traits as  $\text{corr}(G_A, G_B) = r_g$  and

$corr(E_A, E_B) = r_E$ , respectively, the correlation between the two intermediate phenotypes is  $corr(P_A, P_B) = r_P = r_G \sqrt{h_A^2 h_B^2} + r_E \sqrt{(1 - h_A^2)(1 - h_B^2)}$ .

We then consider two different scenarios. One where there is direct assortative mating through the trait  $P_A$ , and a second one where there is indirect assortative mating through the trait  $P_B$ . In each case, we define the strength of assortment as:

Direct assortment through $P_A$	$corr(P_A^{p1}, P_A^{p2}) = \rho$
Indirect assortment through $P_B$	$corr(P_B^{p1}, P_B^{p2}) = \rho$

**Supplementary Table 12. Strength of assortment for each case.**

### Generic equations

To develop our models, we will first define some previously described relations. Let's assume a generic trait that can be defined as a sum of a genetic ( $G$ ) and an environmental component ( $E$ ), i.e.  $P = G + E$ . If  $G$  and  $E$  are uncorrelated, then,

$$corr(G, P) = h,$$

$$corr(E, P) = \sqrt{1 - h^2}.$$

Under assortative mating, the expectations of the genetic, environmental, phenotypic, and genetic-by-environmental correlations between partners are<sup>2</sup>

$$corr(G^{p1}, G^{p2}) = \rho h^2,$$

$$corr(G^{p1}, E^{p2}) = corr(E^{p2}, G^{p1}) = \rho \sqrt{h^2} \sqrt{1 - h^2},$$

$$corr(E^{p1}, E^{p2}) = \rho(1 - h^2),$$

$$corr(P^{p1}, P^{p2}) = \rho.$$

### Relations between phenotype, environment and genetics between partners under our models

To develop our equations, we first need to develop some relations that we will use afterwards.

Under our models, we have these relations within the same individual:

$$\begin{aligned} \text{corr}(G_A, P_B) &= \frac{\text{cov}(G_A, P_B)}{\sigma(G_A)\sigma(P_B)} = \frac{\text{cov}(G_A, G_B) + \text{cov}(G_A, E_B)}{\sigma(G_A)\sigma(P_B)} = \frac{r_G\sigma(G_A)\sigma(G_B) + 0}{\sigma(G_A)\sigma(P_B)} = r_G h_B, \\ \text{corr}(P_A, G_B) &= \frac{\text{cov}(P_A, G_B)}{\sigma(P_A)\sigma(G_B)} = \frac{\text{cov}(G_A, G_B) + \text{cov}(E_A, G_B)}{\sigma(P_A)\sigma(G_B)} = \frac{r_G\sigma(G_A)\sigma(G_B) + 0}{\sigma(P_A)\sigma(G_B)} = r_G h_A. \end{aligned}$$

Since  $P_A$  and  $P_B$  are two random variables and correlated with a coefficient of  $r_p$ , the following linear relationships hold

$$\begin{aligned} P_B &= r_p \frac{\sigma_{P_B}}{\sigma_{P_A}} P_A + \epsilon_B, \\ P_A &= r_p \frac{\sigma_{P_A}}{\sigma_{P_B}} P_B + \epsilon_A, \end{aligned}$$

where  $\epsilon_B$  is the residual of  $P_B$  after adjusting for  $P_A$  that is uncorrelated with  $P_A$  (i.e.  $\text{corr}(P_A, \epsilon_B) = 0$ ), and  $\epsilon_A$  is the residual of  $P_A$  after adjusting for  $P_B$  that is uncorrelated with  $P_B$  (i.e.  $\text{corr}(P_B, \epsilon_A) = 0$ ). In addition, since the genetic and environmental components of the same trait are two uncorrelated independent random variables, the following relationships hold:  $\text{corr}(G_A, \epsilon_B) = \text{corr}(E_A, \epsilon_B) = \text{corr}(G_B, \epsilon_A) = \text{corr}(G_B, \epsilon_A) = 0$ .

Using these, we may derive the relationship between two traits across partners under our models, which are summarised in Supplementary Table 13 (page 15). As a detailed example for these derivations, assuming assortment is on trait  $P_A$ , we have

$$\begin{aligned} \text{corr}(G_A^{p1}, P_B^{p2}) &= \frac{\text{cov}\left(G_A^{p1}, r_p \frac{\sigma_{P_B}}{\sigma_{P_A}} P_A^{p2} + \epsilon_B^{p2}\right)}{\sigma(G_A)\sigma(P_B)} = \frac{r_p \frac{\sigma_{P_B}}{\sigma_{P_A}} \text{cov}(G_A^{p1}, P_A^{p2}) + \text{cov}(G_A^{p1}, \epsilon_B^{p2})}{\sigma(G_A)\sigma(P_B)} \\ &= \frac{r_p \frac{\sigma_{P_B}}{\sigma_{P_A}} \text{corr}(G_A^{p1}, P_A^{p2})\sigma(G_A)\sigma(P_A) + 0}{\sigma(G_A)\sigma(P_B)} \\ &= r_p \text{corr}(G_A^{p1}, P_A^{p2}) \\ &= \rho r_p h_A \end{aligned}$$

	Direct Assortment on $P_A$	Indirect Assortment on $P_B$
$corr(G_A^{p1}, P_B^{p2})$	$\rho r_P h_A$	$\rho r_G h_B$
$corr(G_B^{p1}, P_B^{p2})$	$\rho r_G r_P h_A$	$\rho h_B$
$corr(G_A^{p1}, P_A^{p2})$	$\rho h_A$	$\rho r_G r_P h_B$
$corr(G_B^{p1}, P_A^{p2})$	$\rho r_G h_A$	$\rho r_P h_B$
$corr(P_A^{p1}, P_A^{p2})$	$\rho$	$\rho r_P^2$
$corr(P_A^{p1}, P_B^{p2})$	$\rho r_P$	$\rho r_P$
$corr(P_B^{p1}, P_B^{p2})$	$\rho r_P^2$	$\rho$

**Supplementary Table 13. The between traits between partners correlation under our models.**

Using all these expressions we can finally obtain,

$$\begin{aligned}
\sigma^2(P_Y) &= \sigma^2(P_Y^{p2}) = \sigma^2(P_Y^{p1}) = \sigma^2(P_A^{p1} + \varphi P_B^{p2}) \\
&= \sigma^2(P_A^{p1}) + 2\varphi corr(P_A^{p1}, P_B^{p2})\sigma(P_A^{p1})\sigma(P_B^{p2}) + \varphi^2\sigma^2(P_B^{p2}) \\
&= 1 + 2\varphi\rho r_P + \varphi^2.
\end{aligned}$$

### Derivation of $r_{cp}$

The expectation of the correlation of  $P_Y$  between partners ( $r_{cp}$ ) is

$$\begin{aligned}
r_{cp} &= \frac{cov(P_Y^{p1}, P_Y^{p2})}{\sigma^2(P_Y)} = \frac{cov(P_A^{p1} + \varphi P_B^{p2}, P_A^{p2} + \varphi P_B^{p1})}{\sigma^2(P_Y)} \\
&= \frac{cov(P_A^{p1}, P_A^{p2}) + cov(P_A^{p1}, \varphi P_B^{p1}) + cov(\varphi P_B^{p2}, P_A^{p2}) + cov(\varphi P_B^{p1}, \varphi P_B^{p2})}{\sigma^2(P_Y)} \\
&= \frac{cov(P_A^{p1}, P_A^{p2}) + 2\varphi cov(P_A, P_B) + \varphi^2 cov(P_B^{p1}, P_B^{p2})}{\sigma^2(P_Y)} \\
&= \frac{corr(P_A^{p1}, P_A^{p2}) + 2\varphi corr(P_A, P_B) + \varphi^2 corr(P_B^{p1}, P_B^{p2})}{\sigma^2(P_Y)},
\end{aligned}$$



which under the assumptions of the two AM scenarios yields the specific equations in Supplementary Table 14 (below).

	Direct Assortment on $P_A$	Indirect Assortment on $P_B$
$r_{cp}$	$\frac{\rho + 2\phi r_P + \phi^2 \rho r_P^2}{1 + \phi^2 + 2\phi \rho r_P}$	$\frac{\rho r_P^2 + 2\phi r_P + \phi^2 \rho}{1 + \phi^2 + 2\phi \rho r_P}$ .

**Supplementary Table 14. Expectation of  $r_{cp}$  under our models.**

### Derivation of $h_{own}^2$

The expectation of  $h_{own}^2$  is the total variance explained ( $R^2$ ) by the linear equation of  $P_Y \sim G_A + G_B$ ,

$$h_{own}^2 = \frac{[corr(P_Y, G_A)]^2 + [corr(P_Y, G_B)]^2 - 2corr(P_Y, G_A)corr(P_Y, G_B)corr(G_A, G_B)}{1 - [corr(G_A, G_B)]^2},$$

where the individual terms are,

$$\begin{aligned} corr(P_Y, G_A) &= corr(P_Y^{p1}, G_A^{p1}) = corr(P_Y^{p2}, G_A^{p2}) \\ &= \frac{cov(P_A^{p1}, G_A^{p1}) + \phi cov(P_B^{p2}, G_A^{p1})}{\sigma(P_Y)\sigma(G_A)} \\ &= \frac{corr(P_A^{p1}, G_A^{p1}) + \phi corr(P_B^{p2}, G_A^{p1})}{\sigma(P_Y)}, \end{aligned}$$

$$\begin{aligned} corr(P_Y, G_B) &= corr(P_Y^{p2}, G_B^{p2}) = corr(P_Y^{p1}, G_B^{p1}) \\ &= \frac{cov(P_A^{p1}, G_B^{p1}) + \phi cov(P_B^{p2}, G_B^{p1})}{\sigma(P_Y)\sigma(G_B)} \\ &= \frac{corr(P_A^{p1}, G_B^{p1}) + \phi corr(P_B^{p2}, G_B^{p1})}{\sigma(P_Y)}, \end{aligned}$$

$$corr(G_A, G_B) = corr(G_A^{p2}, G_B^{p2}) = corr(G_A^{p1}, G_B^{p1}) = r_G.$$

Hence, using Supplementary Table 13, we can obtain the specific equations in Supplementary Table 15 (below) for our two scenarios.

	Direct Assortment on $P_A$	Indirect Assortment on $P_B$
$h_{own}^2$	$\frac{(1 + \varphi \rho r_p)^2 h_A^2}{1 + \varphi^2 + 2\varphi \rho r_p}$	$\frac{h_A^2 + 2\varphi \rho r_G h_A h_B + \varphi^2 \rho^2 h_B^2}{1 + \varphi^2 + 2\varphi \rho r_p}$

**Supplementary Table 15. Expectation of  $h_{own}^2$  under our models.**

### Derivation of $h_{partner}^2$

The expectation of  $h_{partner}^2$  is the  $R^2$  of the linear equation  $P_Y^{p1} \sim G_A^{p2} + G_B^{p2}$ ,

$$h_{partner}^2 = \frac{[corr(P_Y^{p1}, G_A^{p2})]^2 + [corr(P_Y^{p1}, G_B^{p2})]^2 - 2corr(P_Y^{p1}, G_A^{p2})corr(P_Y^{p1}, G_B^{p2})corr(G_A^{p2}, G_B^{p2})}{1 - [corr(G_A^{p2}, G_B^{p2})]^2},$$

where the individual terms are,

$$corr(P_Y^{p1}, G_A^{p2}) = \frac{cov(P_A^{p1}, G_A^{p2}) + \varphi cov(P_B^{p2}, G_A^{p2})}{\sigma(P_Y)\sigma(G_A)}$$

$$= \frac{corr(P_A^{p1}, G_A^{p2}) + \varphi corr(P_B^{p2}, G_A^{p2})}{\sigma(P_Y)},$$

$$corr(P_Y^{p1}, G_B^{p2}) = \frac{cov(P_A^{p1}, G_B^{p2}) + \varphi cov(P_B^{p2}, G_B^{p2})}{\sigma(P_Y)\sigma(G_B)}$$

$$= \frac{corr(P_A^{p1}, G_B^{p2}) + \varphi corr(P_B^{p2}, G_B^{p2})}{\sigma(P_Y)},$$

$$corr(G_A, G_B) = corr(G_A^{p2}, G_B^{p2}) = corr(G_A^{p1}, G_B^{p1}) = r_G.$$

Hence using Supplementary Table 13, we obtain the specific equations in Supplementary Table 16 (below) for our two scenarios.

	Direct Assortment on $P_A$	Indirect Assortment on $P_B$
$h_{Partner}^2$	$\frac{\rho^2 h_A^2 + 2\phi\rho r_G h_A h_B + \phi^2 h_B^2}{1 + \phi^2 + 2\phi\rho r_P}$	$\frac{(\rho r_P + \phi)^2 h_B^2}{1 + \phi^2 + 2\phi\rho r_P}$

**Supplementary Table 16. Expectation of  $h_{Partner}^2$  under our models.**

### Derivation of the ratio between $h_{Partner}^2$ and $r_{cp}^2 h_{Own}^2$

From Supplementary Table 14, Supplementary Table 15 and Supplementary Table 16 (pages 17-18), it is straightforward to obtain the ratio between  $h_{Partner}^2$  and  $r_{cp}^2 h_{Own}^2$  which for our two scenarios is given in Supplementary Table 17 (below).

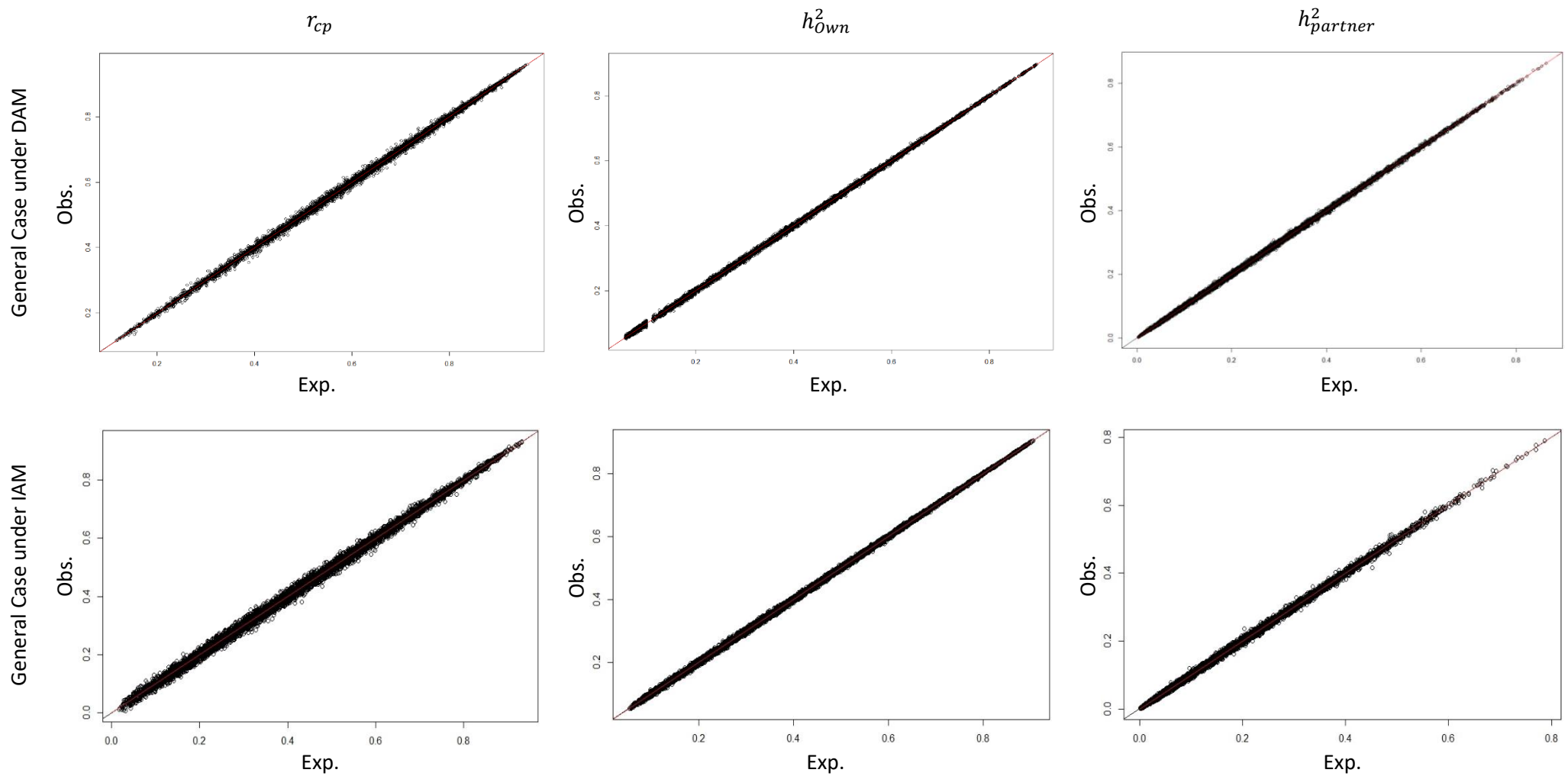
	$\frac{h_{Partner}^2}{r_{cp}^2 h_{Own}^2}$
Direct Assortment on $P_A$	$\left[ \frac{1 + \phi^2 + 2\phi\rho r_P}{(1 + \phi\rho r_P)(\rho + 2\phi r_P + \phi^2 \rho r_P^2)} \right]^2 \frac{\rho^2 h_A^2 + 2\phi\rho r_G h_A h_B + \phi^2 h_B^2}{h_A^2}$
Indirect Assortment on $P_B$	$\left[ \frac{(1 + \phi^2 + 2\phi\rho r_P)(\rho r_P + \phi)}{\rho r_P^2 + 2\phi r_P + \phi^2 \rho} \right]^2 \frac{h_B^2}{h_A^2 + 2\phi\rho r_G h_A h_B + \phi^2 \rho^2 h_B^2}$

**Supplementary Table 17. Expectation of the ratio between  $h_{Partner}^2$  and  $r_{cp}^2 h_{Own}^2$  under our models.**

## Simulation

### Validation of Analytical results

We compared empirical results from simulations with the analytically derived expectations. Specifically, we simulated populations of 10,000 couples based on the assumed model and estimated the three quantities  $h_{Own}^2$ ,  $h_{Partner}^2$  and  $r_{cp}$ . We compared these estimates across 10,000 replicates with varying parameters with the expectations given by the derived equations (Supplementary Table 14, Supplementary Table 15 and Supplementary Table 16, pages 17-18). As illustrated in Supplementary Figure 7 (below), results were in close agreement with the analytical results.



**Supplementary Figure 7. Comparing the observed value from simulated data with its expected value under our model.**

Expected values on x-axis and observed values on y-axis. In each plot, black line (fitted line) overlaps with red line ( $y=x$ ).

## Simulation Details

Simulations were performed by sampling 20,000 instances (10,000 instances for direct assortment + IGEs scenarios and 10,000 instances for indirect assortment + IGEs scenarios) of model parameters  $h_A^2$  and  $h_B^2$  (the heritabilities of phenotypes  $P_A$  and  $P_B$ );  $r_G$  and  $r_E$  (the genetic and environmental correlation between  $P_A$  and  $P_B$ );  $\rho$  (the strength of assortment); and  $\varphi$  (the strength of IGEs). For each replicate, we generated genotypes and effect sizes at 1,000 causal loci for each of the two traits based on the chosen model parameters (see below), of which  $[1000 \cdot r_G]$  were shared between the two traits. From the simulated loci effects and the genotypes, genetic components  $G_A$  and  $G_B$  were computed. Environmental components  $E_A$  and  $E_B$  were simulated based on the chosen correlation of  $r_E$  (see below) and then the phenotypes  $P_A$  and  $P_B$  were calculated as the sum of the corresponding genetic and environmental components. For each replicate, a total of 20,000 individuals were generated and paired into 10,000 couples based on the chosen strength of assortment,  $\rho$ , (see below) on the chosen trait (assortment on  $P_A$  for direct assortment + IGEs scenarios and on  $P_B$  for indirect assortment + IGEs scenarios). The phenotype  $P_Y$  for an individual was computed based on the individual's and their partner's phenotypic values for  $P_A$  and  $P_B$  and the chosen strength of IGEs,  $\varphi$ , (see below). Finally, the parameters of interest, specifically the heritability and partner heritability of  $P_Y$  and phenotypic correlation between couples in  $P_Y$ , were estimated. These individual steps are detailed in the following.

### *Simulation Parameters*

The values of parameters were sampled from the following values. The two heritabilities were sampled from the set of values between 0.1 and 0.9 inclusively at 0.1 increments. The correlations,  $r_G$  and  $r_E$ , were sampled from the set of values between 0 and 1 inclusively at 0.1 increments. The strength of assortment,  $\rho$ , was sampled from the set of values between 0.1 and 0.9 inclusively at 0.1 increments. The strength of IGEs,  $\varphi$ , was sampled from the values between 0 and 0.9 inclusively at 0.1 increments.

### *Simulating Genotypes and Genetic Values.*

The genotype at each locus for each individual was sampled independently assuming a biallelic variant in Hardy-Weinberg equilibrium with a reference allele frequency of 0.5.

For each locus, we then sampled allele substitution effects independently from a normal distribution with mean zero and variance one.

We first generated 1,000 loci with associated effects for trait  $P_A$ . Of these  $\lfloor 1000 \cdot r_G \rfloor$  were chosen randomly and combined with a further  $1000 - \lfloor 1000 \cdot r_G \rfloor$  independently generated loci and effects to form the set of causal loci for  $P_B$ .

Genetic values  $G_A$  and  $G_B$  for each individual were then computed as the sum of the product of the number of alternative alleles and allele substitution effect at each locus.

#### *Simulating Environmental Values $E_A$ and $E_B$*

Environmental values  $E_A$  and  $E_B$  for each individual were sampled independently from a bivariate normal distribution with mean zero and covariance matrix  $\begin{pmatrix} 1 & r_E \\ r_E & 1 \end{pmatrix}$ .

#### *Simulating Phenotypic Values $P_A$ and $P_B$*

Phenotypic values  $P_A$  and  $P_B$  for an individual were then computed as  $P_A = h_A \bar{G}_A + \sqrt{1 - h_A^2} \bar{E}_A$  and  $P_B = h_B \bar{G}_B + \sqrt{1 - h_B^2} \bar{E}_B$ , where  $\bar{G}_A$ ,  $\bar{G}_B$ ,  $\bar{E}_A$  and  $\bar{E}_B$  are within sample standardized quantities, i.e.,  $\bar{X} = (X - \text{mean}(X))/\text{std}(X)$ , where  $\text{mean}()$  and  $\text{std}()$  denote the sample mean and standard deviations.

#### *Pairing of Individuals into couples*

In order to pair the 10,000 simulated male individuals with the 10,000 simulated females so as to achieve a required phenotypic correlation of  $\rho$  for the trait under assortment, we proceeded as follows. We first generated for each male individual a new ‘target’ female phenotype. This was sampled conditional on the male’s phenotype from a normal distribution parameterised so as to yield a correlation of  $\rho$  between the target phenotype and the male phenotype. Then we performed a rank-match between the target phenotypes and the actual phenotypes of the simulated female individuals. For example, if the target phenotype for the 1<sup>st</sup> male has the 800 largest value among all 10,000 target phenotypes and, among the simulated females, the 800 largest simulated phenotypic value belongs to the 999<sup>th</sup> female, we paired the 1<sup>st</sup> male and the 999<sup>th</sup> female and so on.

#### *Computation of Phenotypic Values $P_Y$*

Following pairing of individuals,  $P_Y$  for an individual was computed as  $P_Y = P_A^{\text{own}} + \varphi P_B^{\text{partner}}$ .

### *Estimation of Parameters*

We estimated the heritability of  $P_Y$  as the total variance explained by the linear regression  $P_Y \sim G_A + G_B$ . This is conceptually equivalent to running a standard GREML to estimate the heritability for the observed phenotype using a GRM computed by using the individual own genotypes. Analogously, we estimated the partner indirect heritability of the observed phenotype as the total variance explained in the linear regression  $P_Y \sim G_A^{partner} + G_B^{partner}$ . This is conceptually equivalent to running a GREML to estimate the heritability for the observed phenotype using a GRM computed by using the individual's partner genotypes. In both cases if both  $r_G$  and  $r_E$  were one, i.e.,  $P_A$  and  $P_B$  are the same trait, we excluded  $G_B$  from the two regressions. Finally, the phenotypic correlation between partners in  $P_Y$  was calculated as the Pearson correlation between the generated phenotypic values of partners.

## Supplementary Results

### Expectation of Results under Realistic Parameters

Using the derived formulae, we computed the distribution of results we would expect for a realistic spectrum of parameters. Specifically, we computed  $h_{Partner}^2$ ,  $h_{Own}^2$  and  $r_{cp}$ , under different assumptions and for different values of the traits heritability (on the observed scale), the phenotypic correlation between traits, the breeding value (BV) correlation between two traits and the observed couple correlation. The range of considered values for these parameters, summarised in Supplementary Table 18 (below), was based on their estimates amongst the 51 non-sex-specific traits with significant estimate of  $h_{Partner}^2$  ( $P < 0.05$ ). In order to compare the computed expected results with the empirically observed results we plot  $h_{Partner}^2$  against  $h_{Own}^2 r_{cp}^2$  and compare to the pattern observed in Figure 3a of the main text.

Parameter	Minimum	1 <sup>st</sup> . quartiles	Median	3 <sup>rd</sup> quartiles	Maximum
Own heritability	0.012	0.05	0.08	0.19	0.52
$r_{cp}$	-0.066	0.08	0.19	0.28	0.50
$r_P$	-0.52	-0.01	0.013	0.06	0.87
$r_{BV}$	-0.64	-0.01	0.019	0.07	0.91

**Supplementary Table 18: Ranges of relevant parameters observed across 51 non-sex-specific traits with significant estimates of  $h_{Partner}^2$  ( $P < 0.05$ ) in UK Biobank.**

For each considered scenario, we calculated the expected value of  $h_{Own}^2$ ,  $h_{Partner}^2$  and  $r_{cp}$  based on assumptions about AM and IGEs and equations in Supplementary Table 20 and Supplementary Table 21 (page 24 and page 27). We evaluated two situations, where the parametric values selected covered the 1<sup>st</sup> and 3<sup>rd</sup> quartiles (e.g. the middle 50% distribution) of the observed values in UK biobank data or the selected parametric values covered the minimum to maximum of the observations (Supplementary Table 19, below). We then plotted  $h_{Partner}^2$  against  $h_{Own}^2 r_{cp}^2$  to see whether it matched the pattern observed in Figure 3a (in the main text).



Parameter	Covers 1 <sup>st</sup> and 3 <sup>rd</sup> quartiles of the observed values	Covers minimum and maximum of the observed values
$h_A^2; h_B^2$	0.05, 0.1, 0.15, 0.2	0.05, 0.15, 0.3, 0.5
$\rho$	0.05, 0.1, 0.2, 0.3	-0.05, 0.1, 0.3, 0.5
$\varphi$	$\pm 0.1, \pm 0.2$	$\pm 0.1, \pm 0.2$
$r_G$	-0.05, 0, 0.05, 0.1	-0.6, -0.3, $\pm 0.1$ , 0.4, 0.7, 0.9
$r_P$	-0.05, 0, 0.05, 0.1	-0.5, -0.3, $\pm 0.1$ , 0.4, 0.7, 0.9

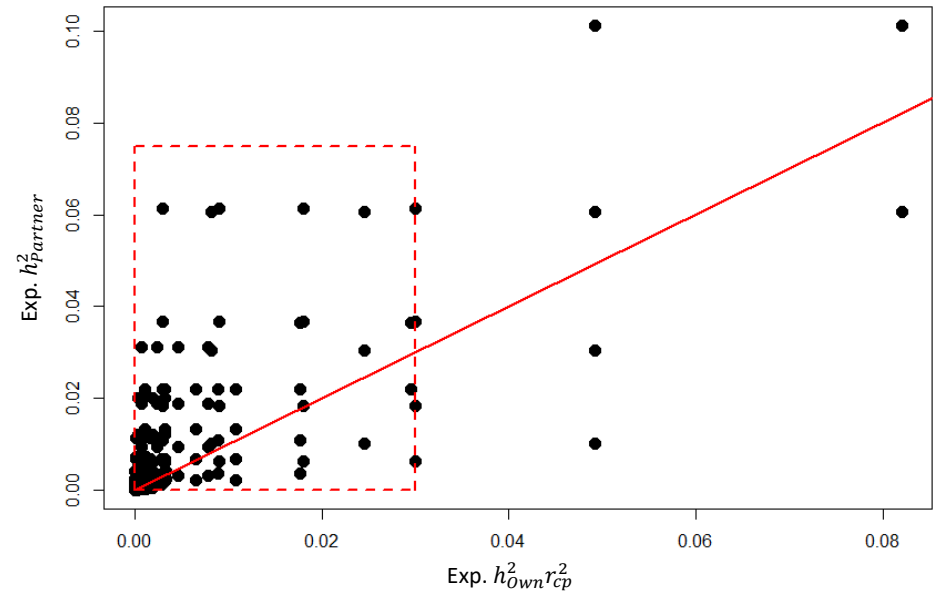
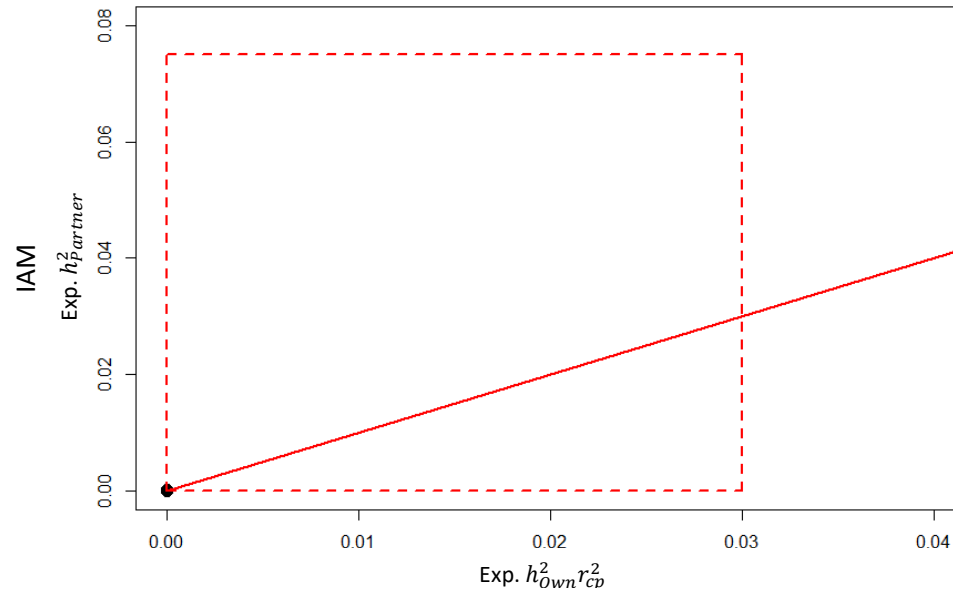
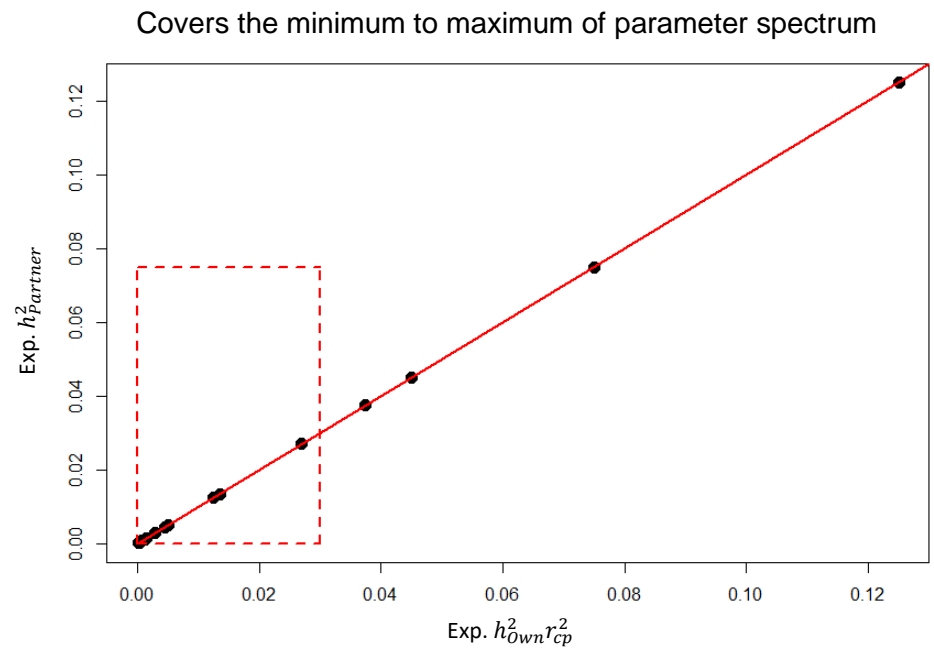
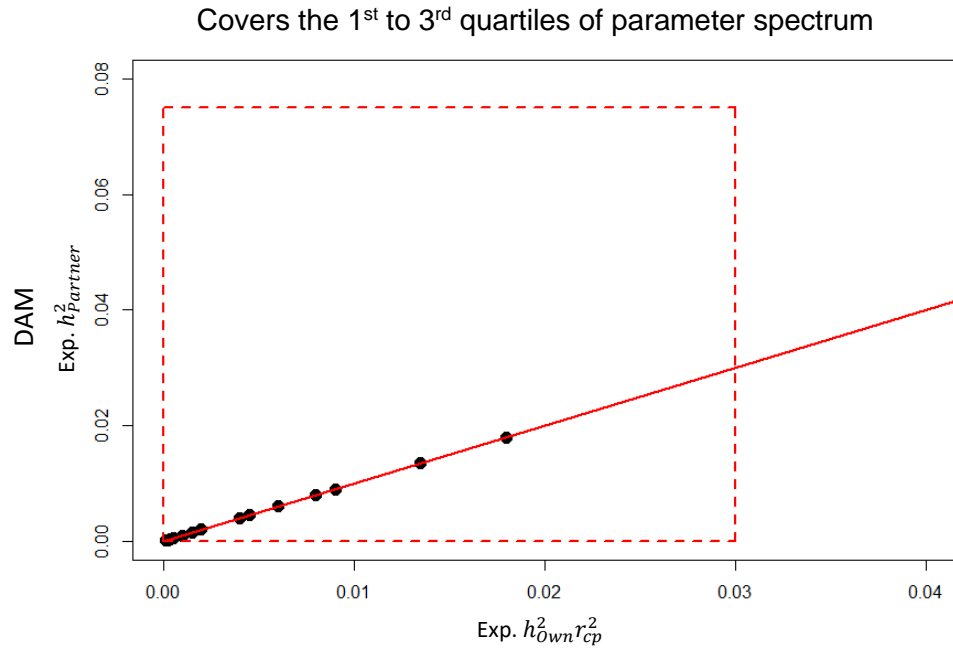
**Supplementary Table 19. Selected values for each parameter based on UK Biobank observed values.**

### Results with no IGEs

We first consider the two scenarios without any IGE, i.e., phenotypic partner correlations arise purely due to direct or indirect assortative mating. These two scenarios are summarised in Supplementary Table 20 (below), including the specific form the general equations take in these cases. The range of results for the considered realistic range of parameters in these two scenarios is shown in Supplementary Table 19 (above). As can be seen these results (Supplementary Figure 8, page 25) do not provide a good match to the observed results from real couple data (Figure 3a of the main text).

Scenario	Assortment	IGEs	Pleiotropy	$r_{cp}$	$h_{own}^2$	$h_{partner}^2$
1	Trait <i>A</i>	No	One Trait	$\rho$	$h_A^2$	$\rho^2 h_A^2$
2	Trait <i>B</i>		Correlated traits	$\rho r_P^2$	$h_A^2$	$\rho^2 r_P^2 h_B^2$

**Supplementary Table 20: Summary of the two scenarios not involving IGEs.**



**Supplementary Figure 8. Expected values of  $h^2_{Partner}$  (y-axis) against  $h^2_{0wn}r^2_{cp}$  (x-axis) in the two scenarios not involving IGEs for the realistic range of model parameters. Red line:  $y = x$ . Red dotted rectangle: the boundary of y and x from real data as shown in Figure 3a in the main text.**

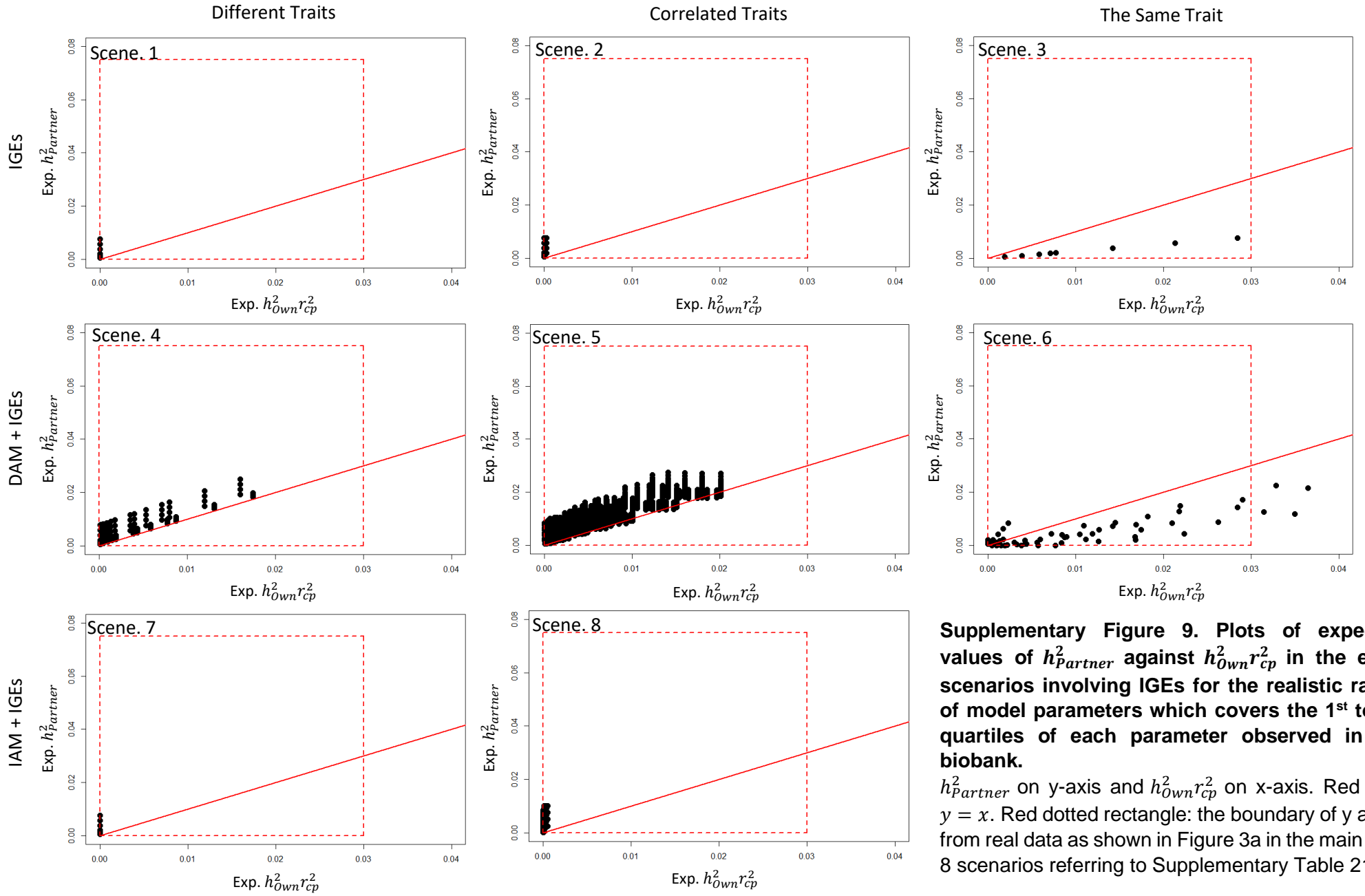
## Results including IGEs

We then considered eight further scenarios which all included IGEs. These are summarised in Supplementary Table 21 (below). The range of results for the considered realistic range of parameters in these eight scenarios is shown in Supplementary Table 19 (page 24).

As can be seen, scenarios 4 and 5 provide the best fit to the observed results while looking at the middle range of the parameter spectrum (Supplementary Figure 9, page 28). That is, when both direct assortment and IGEs are involved in the model and when the trait mediating IGEs is a completely uncorrelated trait (scenario 4) or a trait has a little correlation (e.g.  $r_p \leq 0.1$ ) with the focal trait under assortment (scenario 5). However, once we extend our parameter spectrum to the full range, only scenario 4 remained a similar pattern (Supplementary Figure 10, page 29).

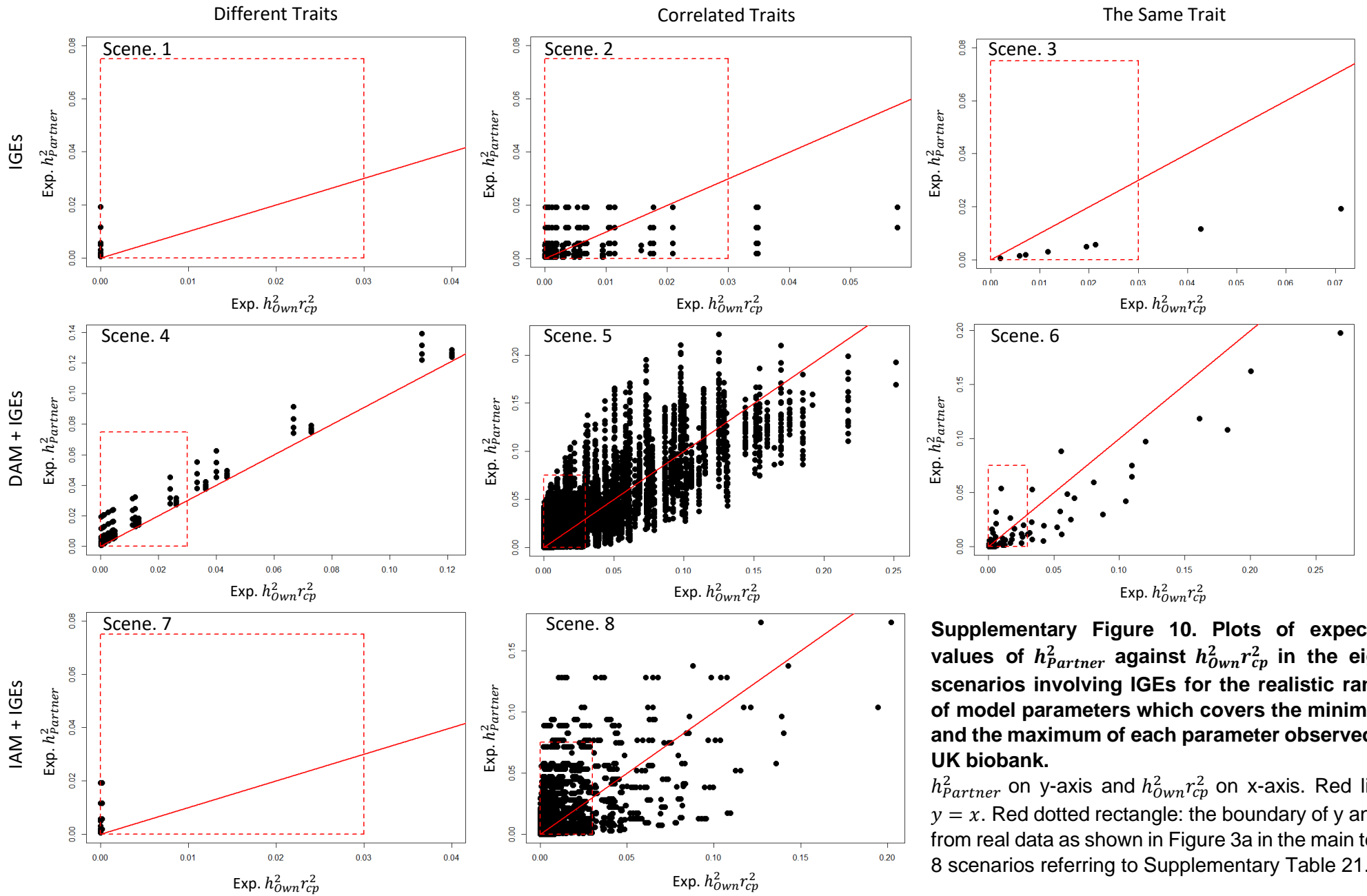
	Assortment	IGEs	Pleiotropy	$r_{cp}$	$h_{Own}^2$	$h_{Partner}^2$
1	No	Yes	Different traits	0	$\frac{h_A^2}{1 + \varphi^2}$	$\frac{\varphi^2 h_B^2}{1 + \varphi^2}$
2			Correlated traits	$\frac{2\varphi r_P}{1 + \varphi^2}$	$\frac{h_A^2}{1 + \varphi^2}$	$\frac{\varphi^2 h_B^2}{1 + \varphi^2}$
3			One trait	$\frac{2\varphi}{1 + \varphi^2}$	$\frac{h_A^2}{1 + \varphi^2}$	$\frac{\varphi^2 h_A^2}{1 + \varphi^2}$
4	Trait $A$		Different traits	$\frac{\rho}{1 + \varphi^2}$	$\frac{h_A^2}{1 + \varphi^2}$	$\frac{\rho^2 h_A^2 + \varphi^2 h_B^2}{1 + \varphi^2}$
5			Correlated traits	$\frac{\rho + 2\varphi r_P + \varphi^2 \rho r_P^2}{1 + 2\varphi \rho r_P + \varphi^2}$	$\frac{(1 + 2\varphi \rho r_P + \varphi^2 \rho^2 r_P^2) h_A^2}{1 + 2\varphi \rho r_P + \varphi^2}$	$\frac{\rho^2 h_A^2 + 2\varphi \rho r_G h_A h_B + \varphi^2 h_B^2}{1 + 2\varphi \rho r_P + \varphi^2}$
6			One trait	$\frac{\rho + 2\varphi + \varphi^2 \rho}{1 + 2\varphi \rho + \varphi^2}$	$\frac{(1 + 2\varphi \rho + \varphi^2 \rho^2) h_A^2}{1 + 2\varphi \rho + \varphi^2}$	$\frac{(\rho^2 + 2\varphi \rho + \varphi^2) h_A^2}{1 + 2\varphi \rho + \varphi^2}$
7	Trait $B$		Different traits	$\frac{\varphi^2 \rho}{1 + \varphi^2}$	$\frac{h_A^2 + \varphi^2 \rho^2 h_B^2}{1 + \varphi^2}$	$\frac{\varphi^2 h_B^2}{1 + \varphi^2}$
8			Correlated traits	$\frac{\rho r_P^2 + 2\varphi r_P + \varphi^2 \rho}{1 + \varphi^2 + 2\varphi \rho r_P}$	$\frac{h_A^2 + 2\varphi \rho r_G h_A h_B + \varphi^2 \rho^2 h_B^2}{1 + \varphi^2 + 2\varphi \rho r_P}$	$\frac{(\rho r_P + \varphi)^2 h_B^2}{1 + \varphi^2 + 2\varphi \rho r_P}$

**Supplementary Table 21. Summary of scenarios involving IGEs.**



**Supplementary Figure 9. Plots of expected values of  $h^2_{Partner}$  against  $h^2_{Own}r^2_{cp}$  in the eight scenarios involving IGEs for the realistic range of model parameters which covers the 1<sup>st</sup> to 3<sup>rd</sup> quartiles of each parameter observed in UK biobank.**

$h^2_{Partner}$  on y-axis and  $h^2_{Own}r^2_{cp}$  on x-axis. Red line:  $y = x$ . Red dotted rectangle: the boundary of y and x from real data as shown in Figure 3a in the main text. 8 scenarios referring to Supplementary Table 21.



**Supplementary Figure 10. Plots of expected values of  $h^2_{Partner}$  against  $h^2_{Own}r^2_{cp}$  in the eight scenarios involving IGEs for the realistic range of model parameters which covers the minimum and the maximum of each parameter observed in UK biobank.**

$h^2_{Partner}$  on y-axis and  $h^2_{Own}r^2_{cp}$  on x-axis. Red line:  $y = x$ . Red dotted rectangle: the boundary of  $y$  and  $x$  from real data as shown in Figure 3a in the main text. 8 scenarios referring to Supplementary Table 21.

## Supplementary Reference

1. Okbay, A. *et al.* Genome-wide association study identifies 74 loci associated with educational attainment. *Nature* **533**, 539-542 (2016). (Cited in Supplementary Table 1)
2. Falconer, D.S. & Mackay, T.F.C. *Introduction to Quantitative Genetics*. 4<sup>th</sup> edition., (Prentice Hall, Harlow, England, 1996). (Cited in Supplementary Methods)