

Supplemental Text and Figures for: Estimates of the heritability of human longevity are substantially inflated due to assortative mating.

J. Graham Ruby, Kevin M. Wright, Kristin A. Rand, Amir Kermany, Keith Noto, Don Curtis, Neal Varner, Daniel Garrigan, Dmitri Slinkov, Ilya Dorfman
Julie M. Granka, Jake Byrnes, Natalie Myres, Catherine Ball

Contents

1	A simple structural equation model for assortative mating and phenotypic correlation	3
1.1	Nominal heritability: the foundational model	3
1.2	Consideration of non-genetic inheritance	4
1.3	Consideration of assortative mating	5
1.4	Consideration of shared household environments	7
1.5	Modeling the correlation between siblings-in-law	9
1.6	Applying the structural equation model to remote blood relatives	10
1.7	Applying the structural equation model to remote in-law relatives	12
2	Over-estimation of heritability due to assortative mating	13
2.1	Modification of expected correlations by the addition of assortative mating	13
2.2	Error from linear extrapolation through multiple relative-types	15
3	Solutions for terms in our structural equation model	16
3.1	Co-siblings-in-law: two informative relative types	16
3.2	Definition of a useful meta-term: A_{sib}	18
3.3	Calculation of the transferable variance term t^2	18
3.4	Calculation of the assortative mating term a (two methods)	19
3.5	Calculation of the inheritance term β	19
3.6	Shared non-transferrable environment: calculation of c_{sib} and c_{sp}	20
3.7	Estimation of t^2 by the assortment-correction method	21
4	Identity-by-descent analysis of SAP accuracy	22
4.1	Genotyping and IBD-calculation methods	22
4.2	IBD between in-law relatives	22
5	Evaluation of standard-error accuracies within the network structure of the SAP	23

5.1	Method: empirical estimation of correlation errors	23
5.2	Result: observed errors were approximately as-expected	24
6	Gompertzian mortality hazard through recent human history	25
6.1	Method: measurement of age-specific mortality hazard	25
6.2	Lifespan differences between genders varied through history	25
6.3	Gompertzian hazard was consistent through history	26
	References	27
7	Supplemental Figure S1: Identity-by-descent evaluation of SAP relationships.	29
7.1	Figure S1 Legend	29
8	Supplemental Figure S2: Evaluation of standard-error accuracies	30
8.1	Figure S2 Legend	30
9	Supplemental Figure S3: Consistent Gompertzian mortality hazard through history	31
9.1	Figure S3 Legend	32

1 A simple structural equation model for assortative mating and phenotypic correlation

The term *assortative mating* refers to sexual selection for mates with similar phenotypes. To the extent that phenotypes are genetically driven or associated, it can decrease the genetic variance/increase the genetic homogeneity within families. With respect to a given phenotype, it can take several forms. *Primary* assortative mating refers to direct selection on the phenotype of interest, while *secondary* assortative mating (homogamy) refers to the indirect effects of population-structure, inbreeding and social stratification (homogamy) on phenotypic similarity, as well as the effects of primary assortative mating around secondary phenotypes on the primary phenotype (Rao et al, 1979)[1]. As we discuss in the Main Text, primary assortative mating is impossible for our phenotype of interest (lifespan): the selection of mates being an activity of living beings, while complete lifespan can only be directly observed directly after death.

1.1 Nominal heritability: the foundational model

The simplest consideration of the concept of heritability (h^2) is in terms of its definition: the fraction of phenotypic variance attributable to genetic variance, given that all phenotypic variance can be divided between “genetic” (V_G) and “environmental/other” (V_E):

$$h^2 = \frac{V_G}{V_P} = \frac{V_G}{V_G + V_E} \quad (1)$$

In this framework, h^2 reflects the correlation that would be expected between two individuals whose genetic states are identical (with respect to phenotypically-relevant genetics), but for whom all other factors that may affect the phenotype are randomly-distributed. Identical twins are examples where the first criterion is met (identical genotypes), but not the second (random environmental/other factors). Nonetheless, if the contribution of shared environmental and other factors could be teased out of the overall correlation between twins, the remaining correlation would reflect h^2 (also excluding dominance and epistasis variance). The assumption of nominal heritability estimates is that as relatives become more distant, their shared environment drops to zero and their amount of shared genetics drops by half with each generation. In nature, the latter assumption is a rule imposed by the mechanism of meiosis, and the consequential fraction of shared genetics is referred to as the *additive relatedness*. Under all of the assumptions listed above, the heritability of a trait can be calculated easily from the correlations of phenotypes (ρ) between remote relatives (piblings/avuncular, ρ_{pib} ; first cousins, ρ_{cuz1} ; first cousins once-removed, ρ_{cuz1r1} ;

etc.):

$$h^2 = 4 \times \rho_{pib} \quad (2)$$

$$= 8 \times \rho_{cuz1} \quad (3)$$

$$= 16 \times \rho_{cuz1r1} \quad (4)$$

$$= [...etc...]$$

In the sections below, additive relatedness will be described in terms of a variable (β) rather than in terms of the constant $\frac{1}{2}$. The nominal heritabilities described in the equations above would therefore take the following form, with the further assumption that $\beta = \frac{1}{2}$:

$$\rho_{pib} = 2\beta^3 h^2 \quad (5)$$

$$\rho_{cuz1} = 2\beta^4 h^2 \quad (6)$$

$$\rho_{cuz1r1} = 2\beta^5 h^2 \quad (7)$$

The reason for the extra “2” on the right side of each equation will become apparent when pathway diagrams are discussed.

1.2 Consideration of non-genetic inheritance

As described in the trio of papers on “Multifactorial Inheritance with Cultural Transmission and Assortative Mating” by Cloninger, Rice, and Reich [2, 3, 4], the classical formula for phenotypic variance, $V_P = h^2 + e^2 = 1$ can be modified to include aspects of the environment that are passed down through families in a similar manner as genetics (Cloninger 1979 [3], eq.2):

$$V_P = h^2 + b^2 + e^2 + 2whb = 1 \quad (8)$$

where V_P is total phenotypic variance (set to 1), h , b , and e are the extents to which additive genetic factors, “transmissible sociocultural factors”, and “non-transmissible environment” influence phenotype, respectively (quotes: Cloninger et al, 1979[3], Table 2). [Note: our manuscript uses the term “transferable” rather than “transmissible”, but with the same intended meaning.] Genotypes are partially shared between relatives, as are sociocultural factors (money, geo-location, etc), making the relative contributions of the two difficult to distinguish. Therefore, they are sometimes modeled as a single, combined term t that relates to h and b as follows:

$$t^2 = h^2 + b^2 + 2whb \quad (9)$$

where w indicates the degree to which genotype is correlated with sociocultural status. Use of t^2 simplifies equation 8 to:

$$V_P = t^2 + e^2 = 1 \quad (10)$$

The variable t^2 is therefore akin to h^2 , but rather than describing the “genetic variance” of the phenotype, it describes the variance due to all transferable factors, or “transferable variance”.

1.3 Consideration of assortative mating

The manner in which assortative mating can increase phenotypic correlation between family members (a father F , a mother M , and two offspring $O1$ and $O2$) is summarized by Rice 1979 [2], Figure 2 (reproduced below, with variable names and topology modified to be consistent with the rest of our manuscript):

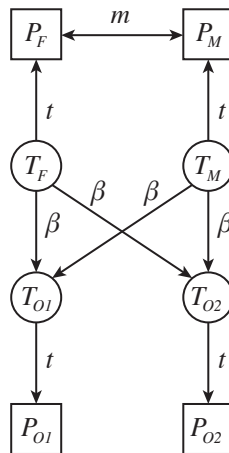


Diagram S1

This path diagram, depicting the phenotypic correlations within a nuclear family with two children, describes transferrable states as latent variables for each family member (T) and the phenotypes (P) that arise as a consequence of those variables. The transferable latent states represent the genotypes as well as the more-abstract/less-defined “sociocultural” type. The extent of inheritance from one parent to one child is described by β , and the extent to which T influences phenotype is represented by t . This model can apply to total transference, but can also represent genetics or sociocultural factors individually by substituting h or b for t , respectively. Consideration of the independent contributions of genetics versus sociocultural factors in an integrated manner leads to path diagrams of increased complexity (e.g. Cloninger 1979 [3], Fig. 3), but we will continue to use a combined model as illustrated in Diagram S1.

Assortative mating is modeled by Rice, Cloninger, and Reich [2, 3] in terms of the correlation of parental phenotypes m . Under this framework, the correlation between the

phenotypes P_{O1} and P_{O2} (i.e. ρ_{sib}) is:

$$\begin{aligned}\rho_{sib} &= 2(\beta^2 t^2) + 2(\beta^2 t^4 m) \\ &= 2\beta^2 t^2 (1 + t^2 m)\end{aligned}\tag{11}$$

While useful for its guiding principles, the precise model in Diagram S1 makes an assumption about assortative mating that does not hold: the model focuses on primary assortative mating, which is based “directly on phenotypic preference” (Cloninger, 1979[3], p.180). In that case, the variable t (or h or b) is again relevant to the estimation of assortment (through the path mt^2). However, in the case of lifespan, it is **by definition** not possible for mates to observe the phenotype at the time of selection (i.e. dead people are generally ineligible to get newly-married or to procreate). This caveat, also stated in the assumptions of Rao et al (1976)[5]: “We suppose that parental phenotypes and the indices of their childhood environments are determined prior to marriage” (p.228), does not remove the possibility of assortative mating, but limits it to *secondary* assortative mating (assortment based on phenotypic indices of relevance to the primary phenotype but nonetheless distinct from it; described by Cloninger, 1979[3]) and the sociocultural and genetic homogamy of the local community from which a mate is selected (this includes inbreeding). This slight variation on the Rice et al (1979)[2] model from Diagram S1 is illustrated below, in Diagram S2:

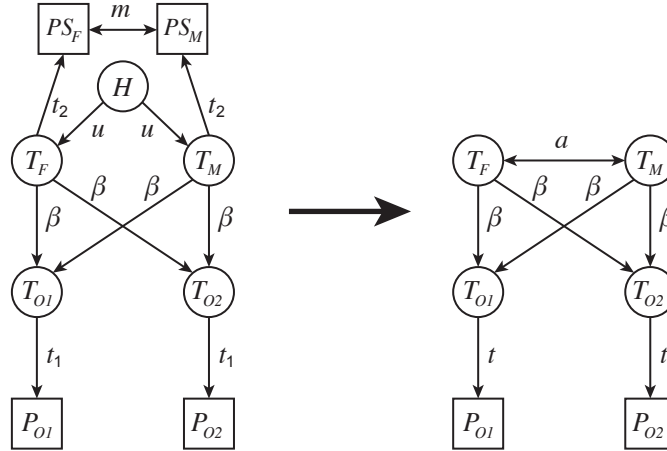


Diagram S2

where PS represents some combined set of secondary phenotypes that are governed by similar genetics as the focal phenotype, to the extent described by t_2 . Mate-assortment around those secondary phenotypes PS creates a correlation between them, described by m . In addition, the cultural and genetic homogamy of the available mate pool can be modeled in terms of a genetic and sociocultural latent state, H , that influences the latent

states of the mother and father (T_M and T_F) to the extent described by the variable u (as in Rao et al, 1979[1]). The correlation between the latent states of the siblings (T_{O1} and T_{O2}) increases as a consequence of these two types of assortment, just as they did due to primary assortment in equation 11:

$$\begin{aligned}\rho_{sib} &= 2\beta^2 t_1^2 + 2\beta^2 t_1^2 (u^2 + t_2^2 m) \\ &= 2\beta^2 t_1^2 (1 + u^2 + t_2^2 m)\end{aligned}\tag{12}$$

Note an important difference between the effect of primary assortment as modeled in equation 11 and secondary assortment as modeled in equation 12: in the first case, one of the factors defining the extent to which assortative mating can enhance the correlation between siblings' latent states is the influence of the latent state on the *focal* phenotype: t (t_1 in eq. 11). That inclusion of t_1 in the assortative-mating addendum ($1 + t_1^2 m$) limits the effect of assortment on phenotypic correlation: a trait with low transferable variance (or heritability) can never substantially increase the correlation between the latent states of spouses (or, by extension, siblings), no matter how strong the assortment on the phenotype. Such is not the case for secondary assortment. The affect of the latent state on the *focal* phenotype (t_1) is absent from the assortative-mating addendum from equation 12: ($1 + u^2 + t_2^2 m$). For non-primary assortative mating, there is no limit to the extent that assortative mating can elevate the correlations of the latent states.

Moving forward, for simplicity and seeing as we do not attempt to separately address the values of u versus m versus t_2 , we abbreviate that transference pathway as a single correlation value, a (Diagram S2, right side):

$$a = u^2 + t_2^2 m\tag{13}$$

$$\rho_{sib} = 2\beta^2 t_1^2 (1 + a)\tag{14}$$

Note that such a definition does not preclude primary assortment. In the case that primary assortment were possible, it would be similarly covered by a :

$$a = u^2 + t_2^2 m_2 + t_1^2 m_1\tag{15}$$

1.4 Consideration of shared household environments

Not all aspects of one's environment are transferrable: the models of Rice, Cloninger, and Reich [2, 3] therefore additionally include nodes for siblings' childhood environments (E_{ch}), which are correlated to the extent described by $c_{e:sib}$ and influence the phenotype to the extent described by e_{ch} , below (as in Rice et al, 1978[2] Fig 6; or Cloninger et al, 1979[3] Fig 3):

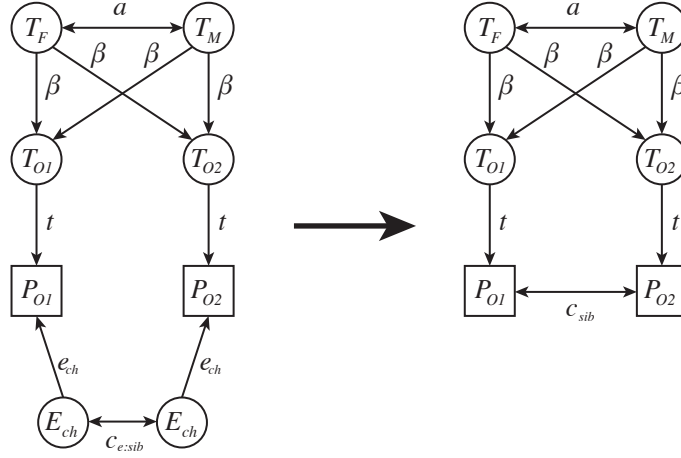


Diagram S3

For simplicity, we abbreviate that path as c_{sib} (right side). Now, with non-transferrable childhood environment taken into consideration, the correlation between siblings' phenotypes becomes:

$$c_{sib} = c_{e:sib}e^2 \quad (16)$$

$$\rho_{sib} = 2\beta^2t^2(1 + a) + c_{sib} \quad (17)$$

If equation 17 were to be fully-extended, and expressed in terms of b^2 and h^2 rather than t^2 , then it would still be missing two sources of genetic variance not shared between other relative-types except via in-breeding: dominance variance and epistasis variance. The h^2 term built into equation 17 only includes additive genetic variance. We do not separately parameterize these terms here, so c_{sib} should be considered to encompass them in addition to the effect of shared childhood environment.

The sharing of adult household environment is often ignored by structural equation models because it can be: by defining the phenotype as pre-cohabitation, the effects of this non-transferrable variable can be ignored (in the words of Rao et al, 1976[5], "We suppose that parental phenotypes... are determined prior to marriage[;] i.e., cohabitation does not increase similarity of mates", p.228-229). However, in the case of lifespan, the phenotype is not defined until death, at which time spouses will likely have spent the majority of their lives sharing household environment. We therefore cannot ignore it, and we model it similarly to the shared childhood household environments of siblings, as a single term (here, c_{sp}):

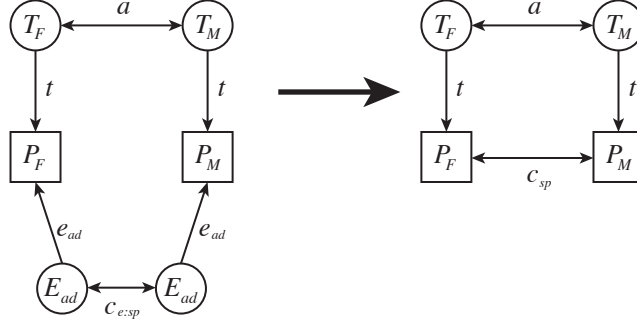


Diagram S4

As for c_{sib} , the correlation of spousal household environments includes both a term for the extent to which the environmental latent states of the adult household E_{ad} are correlated, described by the variable $c_{e:sp}$, and the extent to which those latent states influence the phenotype, described by the variable e_{ad} . We combine those terms into a single term for the resulting phenotypic correlation, c_{sp} :

$$c_{sp} = c_{e:sp}e^2 \quad (18)$$

$$\rho_{sp} = t^2a + c_{sp} \quad (19)$$

1.5 Modeling the correlation between siblings-in-law

We combine the models from Diagram S3 and S4 to construct a model for siblings-in-law. Here, the phenotype of the intermediate sibling/spouse (individual $O2$) is no longer in play, only that individual's latent state T_{O2} . This extended relative-type does not generally share a household environment, so our model only includes transferrable factors, passed between family members through inheritance (β) and assortative mating (a):

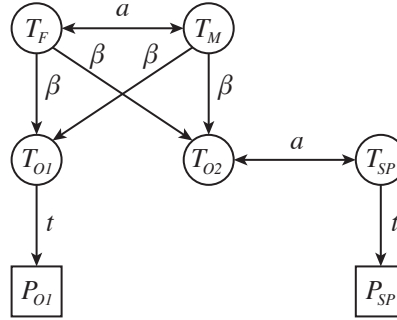


Diagram S5

The full correlation between sibling-in-laws' phenotypes (ρ_{sibL}) can be described as the product of the correlations between each sibling-in-law's phenotype with the latent state of the intermediate sibling/spouse:

$$\rho(P_{O1}, T_{O2}) = 2\beta^2 t(1 + a) \quad (20)$$

$$\rho(T_{O2}, P_{SP}) = ta \quad (21)$$

$$\begin{aligned} \rho_{sibL} &= \rho(P_{O1}, T_{O2}) \times \rho(T_{O2}, P_{SP}) \\ &= 2\beta^2 t(1 + a) \times ta \\ &= 2\beta^2 t^2(1 + a)a \end{aligned} \quad (22)$$

1.6 Applying the structural equation model to remote blood relatives

As with siblings-in-law, remote relatives are assumed to not share non-transferable environment. This assumption is also made by Cloninger et al, (1979). A pathway diagram for an avuncular relationship (referred to as a “pibling” pair, short for “parent’s sibling”), therefore takes the form shown in Diagram S6 below:

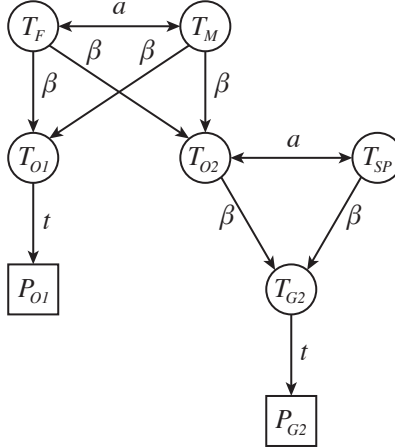


Diagram S6

The path between P_{O1} and P_{G2} always passes through the latent transferrable state T_{O2} , making the overall $\rho(P_{O1}, P_{G2})$ (referred to as ρ_{pib}) equal to the product of $\rho(P_{O1}, T_{O2})$

and $\rho(T_{O2}, P_{G2})$:

$$\rho(P_{O1}, T_{O2}) = 2\beta^2 t(1 + a) \quad (23)$$

$$\begin{aligned} \rho(T_{O2}, P_{G2}) &= \beta t + \beta t a \\ &= \beta t(1 + a) \end{aligned} \quad (24)$$

$$\begin{aligned} \rho_{pib} &= \rho(P_{O1}, T_{O2}) \times \rho(T_{O2}, P_{G2}) \\ &= 2\beta^2 t(1 + a) \times \beta t(1 + a) \\ &= 2\beta^3 t^2(1 + a)^2 \end{aligned} \quad (25)$$

For first-cousins, the situation is similar:

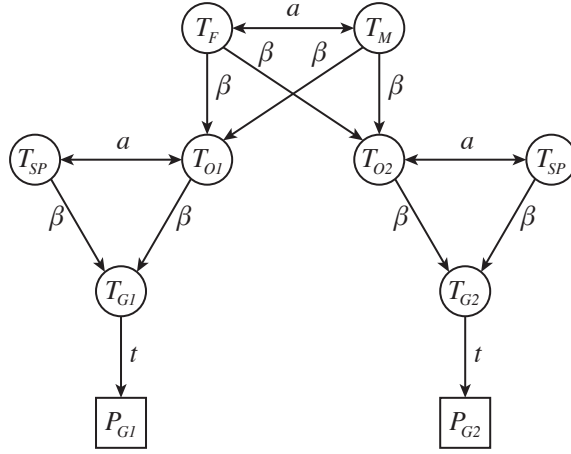


Diagram S7

with all paths between P_{G1} and P_{G2} passing through both T_{O1} and T_{O2} , making the overall $\rho(P_{G1}, P_{G2})$ (referred to as ρ_{cuz1}) equal to the product of $\rho(P_{G1}, T_{O1})$, $\rho(T_{O1}, T_{O2})$, and $\rho(T_{O2}, P_{G2})$:

$$\rho(P_{G1}, T_{O1}) = \beta t(1 + a) \quad (26)$$

$$\rho(T_{O1}, T_{O2}) = 2\beta^2(1 + a) \quad (27)$$

$$\rho(T_{O2}, P_{G2}) = \beta t(1 + a) \quad (28)$$

$$\begin{aligned} \rho_{cuz1} &= \rho(P_{G1}, T_{O1}) \times \rho(T_{O1}, T_{O2}) \times \rho(T_{O2}, P_{G2}) \\ &= \beta t(1 + a) \times 2\beta^2(1 + a) \times \beta t(1 + a) \\ &= 2\beta^4 t^2(1 + a)^3 \end{aligned} \quad (29)$$

Similarly to the situation described by Cloninger et al (1979)[3] for remote relatives in their equations 29-33, the equations for piblings and siblings can be generalized to the

descendants of siblings (specifically, the i th descendant of a person and the i' th descendant of that person's sibling):

$$\rho_{ii'} = 2\beta^{n+1}(1+a)^nt^2 \quad (30)$$

where $n = i + i' + 1$.

1.7 Applying the structural equation model to remote in-law relatives

A similarly general model can be generated for the spouses of remote relatives (i.e. remote relatives-in-law). Considering piblings-in-law, which can be either the pibling of one's spouse (described by $\rho(P_{O1}, P_{SP2})$) or the spouse of one's pibling (described by $\rho(P_{SP1}, P_{G2})$) in Diagram S8:

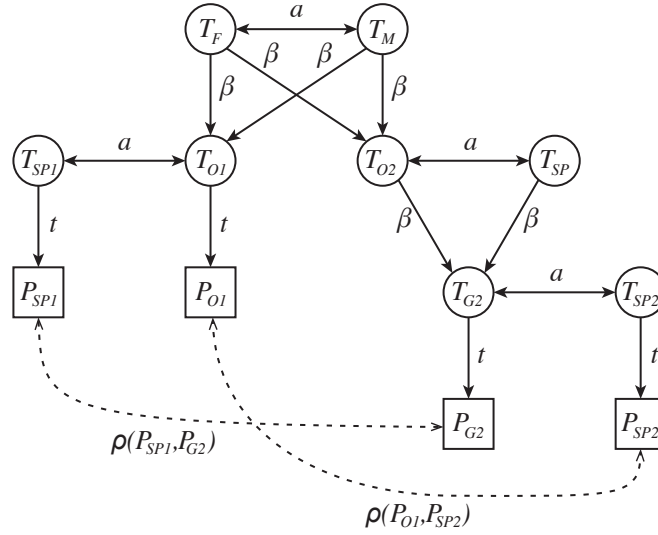


Diagram S8

In both the cases of $\rho(P_{O1}, P_{SP2})$ and $\rho(P_{SP1}, P_{G2})$, all connecting paths pass through $\rho(T_{O1}, T_{G2})$, flanked by a direct effect on a phenotype on one side (t) and a path through a spouse's latent transferrable state (T_{SP1} or T_{SP2}) to their phenotype on the other (in both cases, at). The pibling-in-law correlation (ρ_{pibL}) is therefore given by:

$$\begin{aligned} \rho_{pibL} &= \rho(P_{O1}, P_{SP2}) = \rho(P_{SP1}, P_{G2}) \\ &= \rho(T_{O1}, T_{G2}) \times t \times at \\ &= 2\beta^3(1+a)^2 \times t \times at \\ &= 2\beta^3 t^2 (1+a)^2 a \end{aligned} \quad (31)$$

This is the same formula as given above for ρ_{pib} with an additional multiplication by a . The pattern of multiplying the additive term for a relationship-type by a to get the additive term for the in-law equivalent relationship type was previously seen for siblings (prior to consideration of shared household environment; compare eq. 14 to eq. 22). That pattern can be extended further, to yet-more remote relatives-in-law (i.e. the spouse of a remote relative, as defined above):

$$\rho_{ii'L} = 2\beta^{n+1}(1+a)^n t^2 a \quad (32)$$

where again, $n = i + i' + 1$, and the in-law relative is either the spouse of the i th descendant of one sibling or the i' th descendant of the other.

As a reminder: the models above consider total transference, but can also represent genetics or sociocultural factors individually by substituting h or b for t , respectively. Simultaneous considerations of both h and b generally include one or more terms to describe the correlation between an individual's genetic and sociocultural states (for instance, w in Cloninger et al, 1979[3]; or s and a in Rao et al, 1979[1]). Such considerations greatly complicate these models and add yet more parameters to an already over-parameterized analysis. Below, we maintain that simplicity by considering total transference.

2 Over-estimation of heritability due to assortative mating

In the model framework described above, equations 30 and 32 provide a generalized framework for considering transference of phenotypically-relevant factors between remote relatives and relatives-in-law, respectively. The consequences of assortment on heritability estimates are worth considering. Below, we illustrate the ability of assortative mating to inflate estimates of heritability if they are calculated using the assumptions of nominal heritability. Our discussion below is written in terms of h^2 (i.e. assuming no transferrable-environment contribution, so that $b^2 = 0$ and $t^2 = h^2$).

2.1 Modification of expected correlations by the addition of assortative mating

Assume a trait with 30% heritability ($h^2 = 0.3$), and initially assume there to be no assortative mating. A plot of additive relatedness versus phenotypic correlation across relative-types (omitting any contribution from shared household environment) produces a linear array of points that extrapolates from the origin to the actual heritability, at additive relatedness = 1:

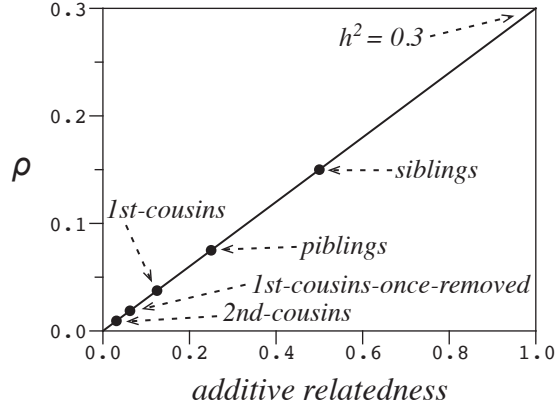


Diagram S9

However, when assortative mating is included, and its coefficient a increases, the plot of additive-relatedness-versus-phenotypic-correlation becomes increasingly curved, with correlations between increasingly distant relatives being increasingly inflated:

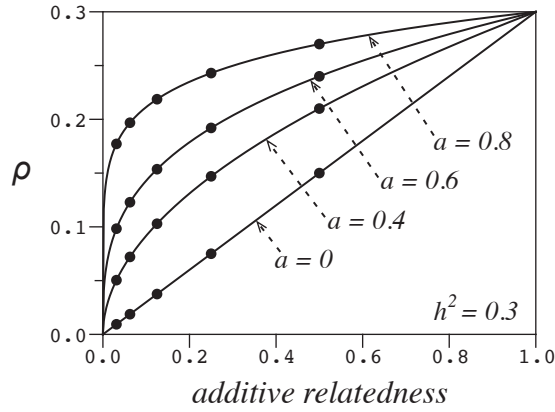


Diagram S10

As a consequence of that inflation, heritability estimates taken from any relative-type's phenotypic regression/correlation would also end up inflated. That trend is exhibited below for piblings:

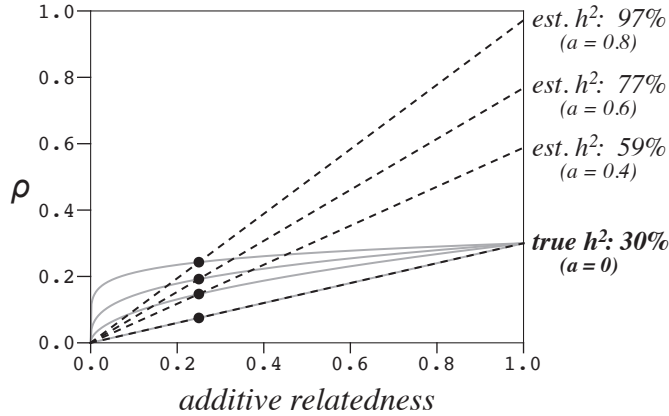


Diagram S11

2.2 Error from linear extrapolation through multiple relative-types

While examination of an additive-relatedness versus phenotypic-correlation plot across relative-types should reveal any substantially confounding effects of assortative mating on the estimation of heritability, a danger to interpretation is posed by close relatives: the relative-types that lay farthest to the right on these plots (siblings, parents, and monozygotic twins) also share household environment. For siblings, for example, that extra non-transferrable environmental correlation is expressed as c_{sib} in equation 17. The danger arises because in all but the most extreme cases, the curvature of the additive-relatedness-versus-phenotypic-correlation plot for distant relatives will be subtle versus the position of the nearest-relative's datum. The combined increased-sharpness of the slope at the far left of the plot, together with the shared-environment-driven inflation of the rightmost datum, will produce a series of points that may fall along an approximately-straight line, whose y-intercept is close to the origin. The curvature that arises from assortative mating is scarcely observable, and the shared-household during childhood term (c_{sib}) is erased in its contribution to the overall slope.

Below, in Diagram S12, is a modeled example illustrating the potential error described above. The model includes a true heritability (h^2) of 0.3, with an assortative-mating coefficient a of 0.3. The true relationship between additive relatedness and phenotypic correlation, **excluding** shared childhood environment, is shown in grey, with the distant-relative data (from piblings, 1st-cousins, 1st-cousins-once-removed, and 2nd-cousins) calculated from equation 30. The sibling value is calculated from equation 17, with c_{sib} set to 0.05. The dotted line shows a linear regression through those five points: its seemingly-consistent slope and y-intercept close to zero give the impression of a phenotype with high heritability, low assortative mating, and negligible shared-childhood environment effect (none of which

is true):

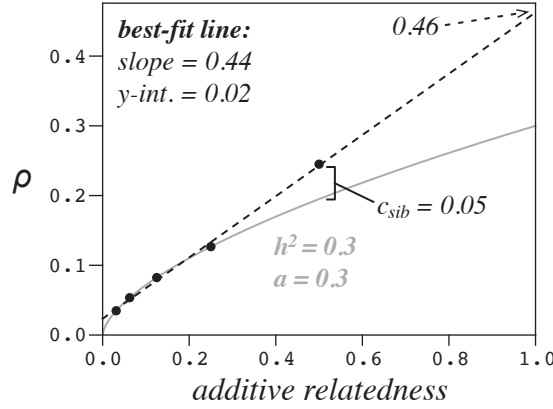


Diagram S12

In summary: without direct examination of assortative mating, the additive-relatedness versus phenotypic-correlation plot of a low-heritability but highly-assorted trait with a large influence from household environment will look close enough to that of a highly-heritable trait with little-to-no assortment nor household environment effect.

3 Solutions for terms in our structural equation model

In this section, we described methods for estimating the core parameters of our structural equation model as described in Section 1. These include the terms for transferrable variance, assortative mating and inheritance (t^2 , a , and β) as well as the two non-transferrable shared household environment terms used here (c_{sib} and c_{sp}).

3.1 Co-siblings-in-law: two informative relative types

The term “co-sibling-in-law” describes two distinct relationship types, each of which is independently informative for finding solutions to the terms in our structural equation model. The first is the sibling of a sibling’s spouse (abbreviated “sib-law-sib”, or *sibLsib*), with the correlation of phenotypes described by the pathway diagram below:

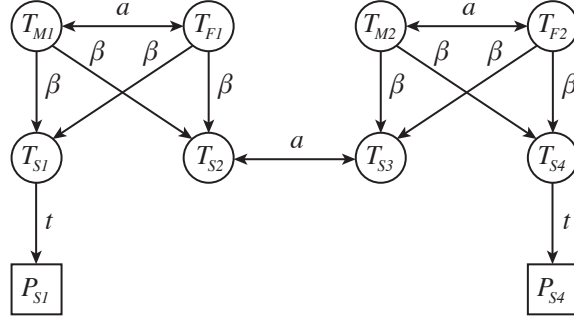


Diagram S12

This relationship includes all of the terms of siblings-in-law, as described in equation 22, but with equivalent paths between siblings' latent states traversed twice, and that correlation term therefore squared:

$$\rho(P_{S1}, T_{S2}) = \rho(T_{S3}, P_{S4}) = 2\beta^2(1 + a)t \quad (33)$$

$$\rho(T_{S2}, T_{S3}) = a \quad (34)$$

$$\begin{aligned} \rho_{sibLsib} &= \rho(P_{S1}, T_{S2}) \times \rho(T_{S2}, P_{S3}) \times \rho(T_{S3}, P_{S4}) \\ &= 2\beta^2(1 + a)t \times a \times 2\beta^2(1 + a)t \\ &= [2\beta^2(1 + a)]^2 at^2 \end{aligned} \quad (35)$$

The second co-sibling-in-law relationship type is the spouse of a spouse's sibling (abbreviated "law-sib-law", or *LsibL*), with the correlation of phenotypes described by the pathway diagram below:

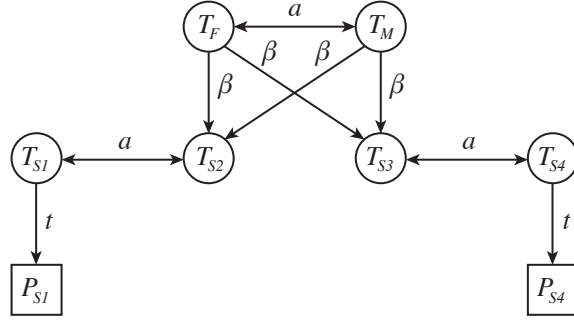


Diagram S13

This relationship again includes all of the terms of siblings-in-law, as described in equation 22, but this time the path between spouses' latent states is traversed twice, so its correlation

term (a) is squared:

$$\rho(P_{S1}, T_{S2}) = \rho(T_{S3}, P_{S4}) = at \quad (36)$$

$$\rho(T_{S2}, T_{S3}) = 2\beta^2(1+a) \quad (37)$$

$$\begin{aligned} \rho_{LsibL} &= \rho(P_{S1}, T_{S2}) \times \rho(T_{S2}, P_{S3}) \times \rho(T_{S3}, P_{S4}) \\ &= at \times 2\beta^2(1+a) \times at \\ &= 2\beta^2(1+a)a^2t^2 \end{aligned} \quad (38)$$

3.2 Definition of a useful meta-term: A_{sib}

To simplify the discussion and equations below, we will define a meta-term, A_{sib} , meant to represent the additive-genetics (or more properly, transference) portion of the sibling correlation. In equation 17, the phenotypic correlation between two siblings was defined as:

$$\rho_{sib} = 2\beta^2(1+a)t^2 + c_{sib}$$

We define A_{sib} to be the full description of the transferrable portion of that correlation:

$$A_{sib} = 2\beta^2(1+a)t^2 \quad (39)$$

The value of A_{sib} can be calculated by using the square of the correlation of siblings-in-law ($sibL$; equation 22) divided by the correlation of co-siblings-in-law of the $LsibL$ variety (equation 38) to cancel out the extra a in those equations:

$$\begin{aligned} \frac{\rho_{sibL}^2}{\rho_{LsibL}} &= \frac{[2\beta^2(1+a) \times at^2]^2}{2\beta^2(1+a) \times a^2 \times t^2} \\ &= \frac{[2\beta^2(1+a)]^2 \times a^2 \times t^4}{2\beta^2(1+a) \times a^2 \times t^2} \\ &= 2\beta^2(1+a) \times t^2 \\ &= A_{sib} \end{aligned} \quad (40)$$

3.3 Calculation of the transferable variance term t^2

In a similar manner to the calculation of A_{sib} in equation 40 above, we can calculate the equivalent term but without the t^2 component by dividing the correlation of co-siblings-in-law of the $sibLsib$ variety (equation 35) by the correlation of siblings-in-law ($sibL$; equation

22). We'll label this meta-term F_{sib} :

$$\begin{aligned}\frac{\rho_{sibLsib}}{\rho_{sibL}} &= \frac{[2\beta^2(1+a)]^2 \times at^2}{2\beta^2(1+a) \times at^2} \\ &= 2\beta^2(1+a) \\ &= F_{sib}\end{aligned}\tag{41}$$

If we divide A_{sib} by F_{sib} , all components of the two meta-terms cancel out except for t^2 , allowing us to calculate its value. Substituting their formulas for calculation from observed correlations, we arrive at an equation for calculating t^2 from data:

$$\begin{aligned}t^2 &= \frac{2\beta^2(1+a) \times t^2}{2\beta^2(1+a)} = \frac{A_{sib}}{F_{sib}} \\ &= \frac{\rho_{sibL}^3}{\rho_{LsibL} \times \rho_{sibLsib}}\end{aligned}\tag{42}$$

3.4 Calculation of the assortative mating term a (two methods)

This section will describe two methods for calculating the assortative mating term a , each using a distinct set of correlation values as their source data. We will refer to the two estimates given by these methods as a'_1 and a'_2 . The first uses a combination of sibling-in-law and co-sibling-in-law data; the second is more general, using any blood-relative type in conjunction with its in-law equivalent.

The first estimate, a'_1 , uses sibling-in-law $sibL$ correlation in conjunction with the co-sibling-in-law correlation of the law-sib-law type, $LsibL$. Through division of the latter (eq. 38) by the former (eq. 22), all variables other than an a coming from one of the exterior spousal pairs are cancelled:

$$a'_1 = \frac{\rho_{LsibL}}{\rho_{sibL}} = \frac{2\beta^2(1+a) \times a^2 \times t^2}{2\beta^2(1+a) \times a \times t^2} = a\tag{43}$$

The second estimate, a'_2 , exploits the structural similarity between the equations for remote relatives (ρ_{ii} ; eq. 30) and their in-law equivalents (ρ_{iiL} ; eq. 32). If the latter is divided by the former for any consistent values of i and i' , all terms cancel except for a :

$$a'_2 = \frac{\rho_{ii'L}}{\rho_{ii}} = \frac{2\beta^{n+1}(1+a)^n t^2 a}{2\beta^{n+1}(1+a)^n t^2} = a\tag{44}$$

3.5 Calculation of the inheritance term β

With t^2 and a solved, there are many equations that can be used for the solving of the path coefficient for the sharing of transferable factors between parent and child, β . We present

a solution here that is once again reliant on siblings-in-law ($sibL$; eq. 22) and the two varieties of co-sibling-in-law ($LsibL$ from eq. 38 and $sibLsib$ from eq. 35). This derivation begins by eliminating, through division, the final spousal edge's a term, in equation 45. Then, in equation 46, we replace the remaining a from the parental $(1 + a)$ term with the solution for a presented above, in equation 43. We then solve for β , arriving at the solution in equation 47:

$$\frac{2 \times \rho_{sibLsib}}{\rho_{sibL}} = \frac{2 \times [2\beta^2(1 + a)]^2 \times a \times t^2}{2\beta^2(1 + a) \times a \times t^2} \quad (45)$$

$$= 2 \times 2\beta^2(1 + a)$$

$$= 4 \times \beta^2(1 + a)$$

$$4 \times \beta^2(1 + a'_1) = \frac{2 \times \rho_{sibLsib}}{\rho_{sibL}} \quad (46)$$

$$4 \times \beta^2\left(1 + \frac{\rho_{LsibL}}{\rho_{sibL}}\right) = \frac{2 \times \rho_{sibLsib}}{\rho_{sibL}}$$

$$4 \times \beta^2\left(\frac{\rho_{sibL} + \rho_{LsibL}}{\rho_{sibL}}\right) = \frac{2 \times \rho_{sibLsib}}{\rho_{sibL}}$$

$$4 \times \beta^2(\rho_{sibL} + \rho_{LsibL}) = 2 \times \rho_{sibLsib}$$

$$2 \times \beta \sqrt{\rho_{sibL} + \rho_{LsibL}} = \sqrt{2 \times \rho_{sibLsib}}$$

$$\beta = \frac{1}{2} \sqrt{\frac{2 \times \rho_{sibLsib}}{\rho_{sibL} + \rho_{LsibL}}} \quad (47)$$

3.6 Shared non-transferrable environment: calculation of c_{sib} and c_{sp}

With A_{sib} defined above as the additive-transference portion of the sibling correlation (eq. 39) and solved using ρ_{sibL} and $\rho_{sibLsib}$ (eq. 40), we can return to equation 17 for siblings and use the difference between that observed value and the calculated value for A_{sib} to solve the shared childhood environment component of sibling correlation (as discussed above, this term also encapsulates dominance and epistasis variance):

$$\begin{aligned} \rho_{sib} &= 2\beta^2(1 + a)t^2 + c_{sib} \\ c_{sib} &= \rho_{sib} - 2\beta^2(1 + a)t^2 \\ &= \rho_{sib} - A_{sib} \\ c_{sib} &= \rho_{sib} - \frac{\rho_{sibL}^2}{\rho_{LsibL}} \end{aligned} \quad (48)$$

Similarly, we can use the three (co-)sibling-in-law relationships to identify the magnitude of spousal phenotypic correlation (ρ_{sp}) that is due to shared adulthood environment rather

than assortative mating based on transferable factors. In equation 19, the correlation between spouses is defined. Rearranging that to solve for the shared-environment term (c_{sp}), we get:

$$c_{sp} = \rho_{sp} - at^2 \quad (49)$$

The sum on the right has two components. The first is a directly-observable correlation value (ρ_{sp}). The second is a combination of variables that were solved above (at^2). Using the structural equations for siblings-in-law ($sibL$; eq. 22) and co-siblings-in-law of the $sibLsib$ variety (eq. 35), we can calculate at^2 as follows:

$$\frac{\rho_{sibL}^2}{\rho_{sibLsib}} = \frac{[2\beta^2(1+a)at^2]^2}{[2\beta^2(1+a)]^2 at^2} = at^2 \quad (50)$$

Substituting that solution for at^2 into equation 49, we arrive at a means of calculating c_{sp} from observable data:

$$c_{sp} = \rho_{sp} - \frac{\rho_{sibL}^2}{\rho_{sibLsib}} \quad (51)$$

3.7 Estimation of t^2 by the assortment-correction method

Above, we described methods for estimating all of the core parameters of our structural equation model (t^2 , a , and β) using three similar and informative relationship-types: siblings-in-law ($sibL$) and two types of co-sibling-in-law ($LsibL$ and $sibLsib$). For remote relatives, such trios of similar relationship-types are lacking, being either not similar in the same manner as the three classes of (co-)siblings-in-law or impractical to efficiently measure, computationally, across a massive pedigree. However, for one parameter, measurement is always straightforward: a , using similar in-law relatives and equation 44.

We have seen the major impact that assortative mating can have on the calculation of inaccurate heritability estimates (Diagram S11,S12). With knowledge of the a parameter's value, those effects can be mitigated by cancelling a out of the equations. Starting with the equation for phenotypic correlation between remote relatives (eq. 32) where $n = i + i' + 1$ for the i th and i' th descendants of a pair of siblings, we establish the relationship between correlation values and the non-assortative terms:

$$\begin{aligned} 2\beta^{n+1}(1+a)^n t^2 &= \rho_{ii'} \\ 2\beta^{n+1} t^2 &= \frac{\rho_{ii'}}{(1+a)^n} \\ 2\beta^{n+1} t^2 &= \frac{\rho_{ii'}}{[1 + (\rho_{ii'L}/\rho_{ii'})]^n} \end{aligned} \quad (52)$$

Equation 52 is equivalent to the traditional equation for calculating heritability, with $2\beta^{n+1}$ being the additive relatedness term, and with the denominator on the right side cancelling

out the effects of assortative mating. Traditionally, for genetics, β is not treated as a variable, but rather is given its meiotically-imposed value of $\frac{1}{2}$. We do the same here, to arrive at our “assortment-corrected estimate of transferable variance” ($t_{ac}^{2'}$; eq. 53):

$$\begin{aligned}
t^2 &= \frac{1}{2\beta^{n+1}} \frac{\rho_{ii'}}{[1 + (\rho_{ii'L}/\rho_{ii'})]^n} \\
t_{ac}^{2'} &= \frac{1}{2(\frac{1}{2})^{n+1}} \frac{\rho_{ii'}}{[1 + (\rho_{ii'L}/\rho_{ii'})]^n} \\
&= 2^n \frac{\rho_{ii'}}{[1 + (\rho_{ii'L}/\rho_{ii'})]^n} \\
t_{ac}^{2'} &= \left[\frac{2}{1 + (\rho_{ii'L}/\rho_{ii'})} \right]^n \rho_{ii'}
\end{aligned} \tag{53}$$

4 Identity-by-descent analysis of SAP accuracy

4.1 Genotyping and IBD-calculation methods

Data utilized for this research project was de-identified and anonymized prior to its use. Samples are collected as follows: a customer orders an AncestryDNA kit; upon receiving the kit, the customer collects saliva into a sample stabilizing solution and mails the sample to the laboratory; DNA from the saliva is extracted; finally, genotypes are called using an Illumina genotyping array (details below). In order to be included in this study, customers must have completed two steps: activate a DNA sample by providing basic personal information - including year of birth, name, gender and consent to research. Next, the customer must have associated their DNA sample to a user generated pedigree and made this pedigree public. This analysis is based on a June 19, 2016 snapshot of 698812 genotyped individuals linked to the SAP. SNP variants were called by technicians at the Illumina FastTrack Microarray Services lab using the GenomeStudio platform. Genotype data was generated using an Illumina genotyping array with approximately 730,000 SNPs.

IBD was measured with a custom algorithm, J-GERMLINE (Ball et al., 2016)[6] based upon the GERMLINE algorithm (Gusev et al. 2009)[7]. An additional down-weight of IBD matches less than 90 cM was applied in the form ‘Timber’ to filter uninformative matches (Ball et al., 2016)[6].

4.2 IBD between in-law relatives

We examined IBD matches between spouses and in-law relatives (Fig. S1A,B). The mean value of IBD matches between spouses was 14.0 cM, or 0.2% of the genome. This was the

equivalent of the IBD-sharing expected for 4th cousins (Fig. S1A). Siblings-in-law and 1st cousins-in-law shared similar values, with IBD sharing equivalent to that of 4th cousins and 4th cousins once removed, respectively. Moving out to more distant relatives, we observed that the sharing of IBD between in-law relatives dropped in parallel with the two-fold dilution per relative-degree that was both expected and observed for blood relatives (Fig. S1B). In-law relatives thereby maintained an approximately 2^7 -fold (128-fold) dilution of genetics versus their blood-relative equivalents. Given this observed less-than-one-percent contribution of genetic similarity through IBD for in-law versus genetic relatives and between spouses, we considered IBD-sharing (i.e. inbreeding) to be a negligible contributor to the assortment that was observed in later analyses.

5 Evaluation of standard-error accuracies within the network structure of the SAP

The method described in the Methods of the Main Text for estimating S.E. is standard, and it is specifically recommended for heritability error estimation (“The precision of an estimate of heritability... is easily obtained from the standard error of the regression or correlation from which the heritability is estimated.” Falconer & Mackay, 1996[8], p.177). However, in the context of large pedigrees, it is possible for relative-pairs to be less-than-independent of one another (e.g. for two siblings A and B who are 1st -cousins of siblings C and D, the 1st-cousin datum (A,C) is only partially independent of the (B,D) datum). This complication, and the complexity of comprehensively accounting for it, has led some authors to conclude that it “is not possible to place standard errors on the heritability estimate of the phenotypes due to the complex relatedness structure of individuals” (Zaitlen et al, 2013[9], p.10).

5.1 Method: empirical estimation of correlation errors

To address and quantify the potential under-estimation of error by the standard technique, we performed a series of analyses to measure the degree to which error was under-estimated by the sample-size statistic described in the Methods. For three series of heritability estimates across decade-long birth cohorts, we performed parallel analyses using independent sets of probands from the same birth-decade: those with years-of-birth that were either even- or odd-numbered. We presumed these data sets to be independent, but to nonetheless be unlikely to differ substantially due to historical or demographic trends.

For a subset of the heritabilities calculated in this manuscript using probands from decade-long birth cohorts, each birth cohort was split by probands born in even versus odd years.

For each split cohort, the heritability and its nominal S.E. were estimated for each sub-cohort. Standard errors of the mean are assumed to be normally-distributed, and their values can therefore be considered similarly to the standard deviation of a normal distribution. Generally, the variance of the differences between random draws x and y from two independent distributions x and y is given by the equation:

$$\sigma_{x-y}^2 = \sigma_x^2 + \sigma_y^2 \quad (54)$$

In our case, the *expected* standard deviation for differences within the pairs of assumedly-equivalent cohorts is given by the standard errors of the respective estimates. So if ρ_{even} and ρ_{odd} are the correlations from the even-year and odd-year sub-cohorts, respectively, and σ_{even} and σ_{odd} are the standard errors of those estimates, respectively (i.e. the standard deviations of the error distributions), then the difference between those two correlations ($\Delta_{even-odd}$) will have an expected distribution with a standard deviation ($\sigma_{\Delta_{even-odd}}$) given by:

$$\sigma_{\Delta_{even-odd}} = \sqrt{\sigma_{even}^2 + \sigma_{odd}^2} \quad (55)$$

In order to evaluate the relationship between traditionally-calculated versus observed estimate-error, we normalized each observed estimate-difference between equivalent even-year versus odd-year cohorts (Δ_{norm}), dividing it by the expected difference given by the respective S.E.'s:

$$\Delta_{norm} = \frac{\rho_{even} - \rho_{odd}}{\sqrt{\sigma_{even}^2 + \sigma_{odd}^2}} \quad (56)$$

The standard deviation of the distribution of Δ_{norm} values ($\sigma_{\Delta_{norm}}$) was therefore expected to be 1. Greater or lesser values would therefore indicate decreased or increased precision of the estimation method versus the reported standard errors, respectively.

5.2 Result: observed errors were approximately as-expected

Looking across three relative-types (siblings, 1st cousins, and 1st cousins in-law), the observed differences between even-year versus odd-year sub-cohorts were in one case slightly greater than expectation (siblings, at 105.4% the expected deviation) and in two cases slightly less than expectation (1st cousins, at 82.5% the expected deviation; and 1st cousins in-law, at 87.4% the expected deviation). In all three cases, the deviation from expectation was consistent across the full range of powers, supported the generality of these under-estimations. These results indicated that the datum-weighting strategy we employed for calculation of standard errors (see Main Text, Methods) has adequately tempered our reporting of errors, and that the error bars presented throughout the paper can be taken at approximately face value.

6 Gompertzian mortality hazard through recent human history

6.1 Method: measurement of age-specific mortality hazard

For each analyzed birth cohort, all individuals with appropriate years of birth and valid lifespans (up to 120), and also satisfying other specified criteria (gender, parenthood, etc.; see Main Text, Methods), were pooled into a population. Age-specific survival was calculated by iteratively subtracting the number of people who died at each age from the population total, moving from young to old. Age-specific death rates (d_a) were calculated as the number of people dying at age a divided by the number of people surviving up to age a . Those were used to calculate age-specific hazard (hz_a) using a simplified Sacher formula (Sacher, 1956 [10]; Gavrilov and Gavrilova, 2011[11]):

$$hz_a = -\ln(1 - d_a) \quad (57)$$

6.2 Lifespan differences between genders varied through history

As discussed above, the reduced-heritability observed between relatives from distinct birth cohorts implied the alleles that affect lifespan to be non-overlapping between historical eras. Historical variation in the factors affecting human lifespan was evident when viewing the profound increase in average adult lifespans (i.e. for individuals who had recorded children in the SAP) across the birth cohorts from this study (Figure S3A). The profundity of change to mortality hazard was further illustrated by the inconsistency of gender-advantage over the 120 years pictured. In the 20th century, female lifespans are well-documented to have exceeded those of males, with much debate as to the relative contributions of sex-specific biology versus sociological factors (Gjonça, 2005)[12]. While female lifespan advantage was substantial post-1900, it diminished moving backwards in time, and flipped to a disadvantage pre-1870 (Figure S3A).

Age-specific mortality hazard provided more insight into the gender-specific dynamics of lifespans through history (Figure S3B). Human lifespan is known to be distributed as described by Gompertz: in mid-to-late life, mortality is distributed according to a Gompertz distribution, whose defining feature is exponentially-increasing hazard (i.e. linearly-increasing log-hazard; Gompertz, 1825)[13]. In early life, as noted by Makeham (1860)[14], age-independent extrinsic hazards can overwhelm the Gompertzian trend, resulting in an approximately stable hazard. Both of these features were repeatedly observable across the birth cohorts of our study, for both genders (Figure S3B).

Several gender-specific deviations from Gompertzian expectations were apparent in the age-specific mortality hazard plots. For males, several transient “bumps” appeared in

young-adulthood; in each case, the age-spans of elevated hazard matched with birth-cohort to corresponded to military enlistment ages for major wars. Members of the 1830-1839 cohort were 22-35 years old during the U.S. Civil War, 1861-1865 (bump around 25-35); members of the 1890-1899 cohort were 18-28-year-olds during U.S. involvement in World War I, 1917-18 (bump around 25); and members of the 1910-1919 cohort were 22-35 years old during U.S. involvement in World War II, 1941-1945 (bump around 25-35y; Figure S3B).

In early cohorts, female mortality hazard was elevated across an early-life age-window that corresponded well to the interval of female fertility (Fig. S3B): beginning with the onset of reproductive maturity in the mid-to-late teens (Apter, 1980; Anderson et al, 2003)[15, 16] and ending menopause in the mid-40s (Gold et al, 2001)[17]. This interval was consistent with birthing-age data in the SAP, shown for the 1900-1909 birth cohort in Fig. S3C. Historically, the danger of childbirth is known to have dropped dramatically across the 1800's (Cutler et al, 2006)[18], consistent with the loss of elevated female-mortality hazard in that interval across the birth cohorts shown in Figure S3B.

6.3 Gompertzian hazard was consistent through history

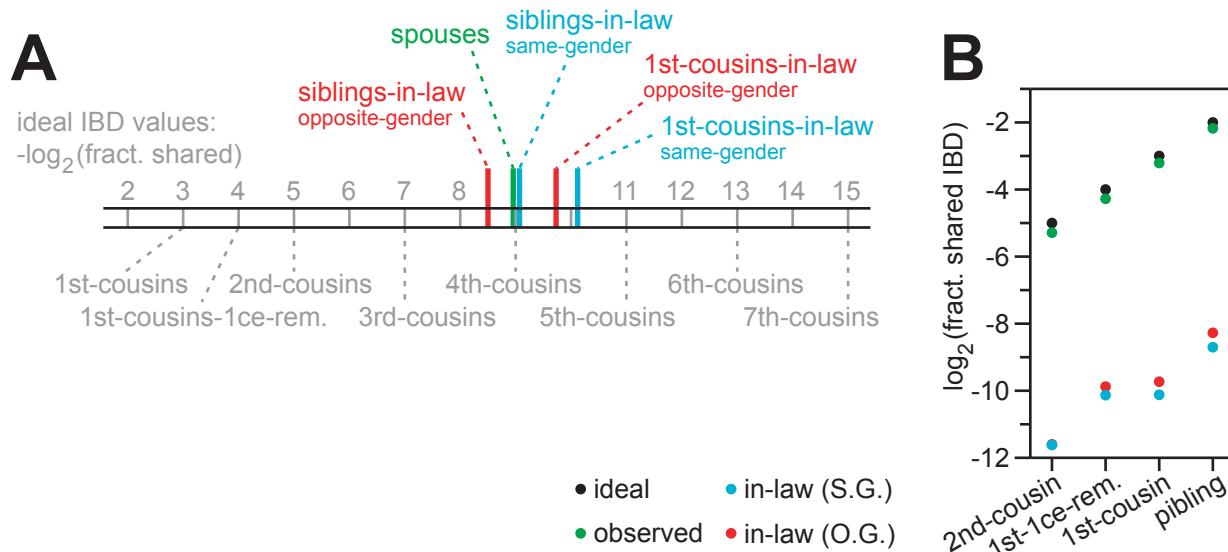
A second trend in hazard distinguished males from females: across the 19th century, while the Gompertzian trend in late-life mortality hazard was consistent for both men and women, the y-axis placement of the hazard slope decreased rapidly for females but slowly for males (Figure S3B). This across-the-board lowering of hazard likely reflected the multitude of progress made by humankind in medicine, health, and safety over the 19th and 20th centuries. Notably, however, the doubling-rate of mortal hazard with age remained approximately constant across that historical span (Figure S3D).

References

- [1] Rao, DC, NE Morton, CR Cloninger (1979) *Path analysis under generalized assortative mating. I. Theory*. Genet Res Camb, 33:175-188.
- [2] Rice, J, CR Cloninger, T Reich (1978) *Multifactorial Inheritance with Cultural Transmission and Assortative Mating. I. Description and Basic Properties of the Unitary Models*. Am J Hum Genet, 30:618-643.
- [3] Cloninger, CR, J Rice, T Reich (1979) *Multifactorial Inheritance with Cultural Transmission and Assortative Mating. II. A General Model of Combined Polygenic and Cultural Inheritance*. Am J Hum Genet, 31:176-198.
- [4] Cloninger, CR, J Rice, T Reich (1979) *Multifactorial Inheritance with Cultural Transmission and Assortative Mating. III. Family Structure and the Analysis of Separation Experiments*. Am J Hum Genet, 31:366-388.
- [5] Rao, DC, NE Morton, S Yee (1976) *Resolution of Cultural and Biological Inheritance by Path Analysis* Am J Hum Genet, 28:228-242.
- [6] Ball, CA, MJ Barber, J Byrnes, Pr Carbonetto, KG Chahine, RE Curtis, JM Granka, E Han, EL Hong, AR Kermany, NM Myres, K Noto, J Qi, K Rand, Y Wang, L Willmore (2016). *Ancestry DNA Matching White Paper*. Retrieved from: <https://www.ancestry.com/corporate/sites/default/files/AncestryDNA-Matching-White-Paper.pdf>
- [7] Gusev, A, JK Lowe, M Stoffel, MJ Daly, D Altshuler, JL Breslow, JM Friedman, I Pe'er (2009). *Whole population, genome-wide mapping of hidden relatedness*. Genome Research, 19, 318–26.
- [8] Falconer, DS, and TFC Mackay. (1996) *Introduction to Quantitative Genetics (4th Edition)*. Pearson Education Limited, England.
- [9] Zaitlen N, P Kraft, N Patterson, B Pasaniuc, G Bhatia, S Pollack, AL Price (2013) *Using Extended Genealogy to Estimate Components of Heritability for 23 Quantitative and Dichotomous Traits*. PLoS Genet 9:e1003520.
- [10] Sacher, GA. (1956) *On the Statistical Nature of Mortality, with Especial Reference to Chronic Radiation Mortality*. Radiology, 67:250–257.
- [11] Gavrilov, LA, and NS Gavrilova (2011) *Mortality measurement at advanced ages: a study of the Social Security Administration Death Master File*. N Am Actuar J, 15:432-447.

- [12] Gjonça, A, C Tomassini, B Toson, S Smallwood. (2005). *Sex Differences in Mortality, A Comparison of the United Kingdom and Other Developed Countries*. Health Statistics Quarterly, 26:6–16.
- [13] Gompertz, B (1825) *On the Nature of the Function Expressive of the Law of Human Mortality and on a New Mode of Determining Life Contingencies*. Philosophical Transactions of the Royal Society of London A, 115: 513–585.
- [14] Makeham, WM (1860) *On the law of mortality and the construction of annuity tables*. J. Inst. Actuaries and Assur. Mag., 8: 301–310.
- [15] Apter, D (1980) *Serum steroids and pituitary hormones in female puberty: a partly longitudinal study.* Clin. Endocrinol., 12: 107–20.
- [16] Anderson, SE, GE Dallal, A Must (2003) *Relative weight and race influence average age at menarche: results from two nationally representative surveys of us girls studied 25 years apart*. Pediatrics, 111:844–50.
- [17] Gold EB, J Bromberger, S Crawford, S Samuels, GA Greendale, SD Harlow, J Skurnick (2001) *Factors associated with age at natural menopause in a multiethnic sample of midlife women*. Am J Epidemiol., 153:865–874.
- [18] Cutler, D, A Deaton, A Lleras-Muney (2006) *The Determinants of Mortality*. Journal of Economic Perspectives, 20:97–120.

7 Supplemental Figure S1: Identity-by-descent evaluation of SAP relationships.

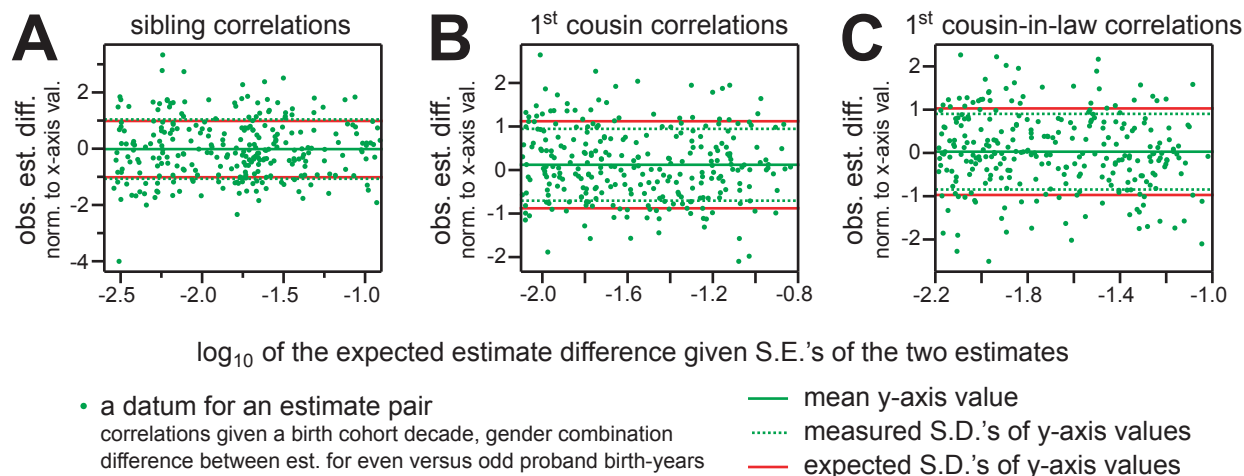


7.1 Figure S1 Legend

Identity-by-descent evaluation of SAP relationships.

- **Panel A:** for select non-genetic relationships, the base-2 log mean fraction of the genome shared-IBD is plotted on a linear representation of the expected value, given additive relatedness, for other relative-types. Observed values for non-genetic relatives are shown in green (spouses), cyan (same-gendered in-laws), and red (opposite-gendered in-laws). Ideal reference values are in grey.
- **Panel B:** for relative-types of increasing additive relatedness (x-axis), the base-2 log values of four fractional-shared-IBD values (y-axis) are plotted: the ideal value based on additive relatedness (black), the observed value in the SAP (green), the observed value for in-law equivalents of the same gender (cyan), and the observed value for in-law equivalents of opposite genders (red).

8 Supplemental Figure S2: Evaluation of standard-error accuracies

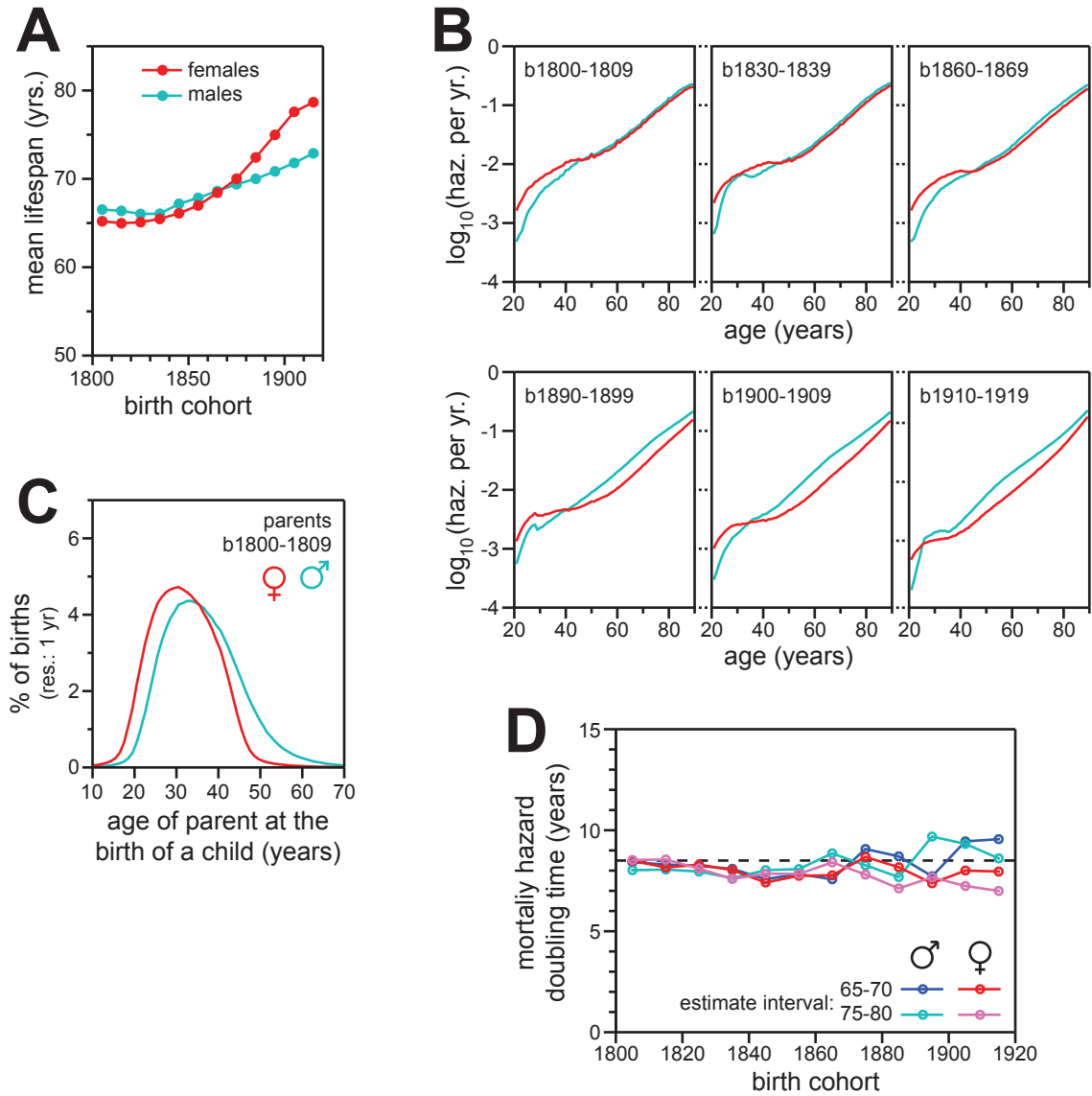


8.1 Figure S2 Legend

Evaluation of standard-error accuracies

- **Panel A:** Empirical versus N-based estimate errors, using sibling correlations. Pairs of correlation estimates taken across particular gender combinations (MM, FF, FM, or MF) and proband-relative birth-year offsets (1-10y, 11-20y, 21-30y) from particular decade-long birth cohorts; each pair is the estimate taken using probands born in the even- versus odd-numbered years of the decade. X-axis: expected difference between two estimates based on their respective standard errors. Y-axis: observed difference, divided by the expected difference. Green dots: values for one estimate-pair. Green solid line: the mean of y-axis values. Red lines: expected standard deviations of y-axis values (1 unit from the mean). Green dashed lines: observed standard deviations of y-axis values.
- **Panel B:** Empirical versus N-based estimate errors, as in Panel A, using first-cousin correlations.
- **Panel C:** Empirical versus N-based estimate errors, as in Panel A, using first-cousin-in-law correlations.

9 Supplemental Figure S3: Consistent Gompertzian mortality hazard through history



9.1 Figure S3 Legend

Consistent Gompertzian mortality hazard through history

- **Panel A:** Mean lifespan for males (cyan) and females (red), by decade birth cohorts.
- **Panel B:** Age-specific base-10 log(mortality hazard per year) for males (cyan) and females (red), for assorted birth-decade cohorts.
- **Panel C:** The distribution of ages at which parents (born 1900-1909) had children (mothers: red; fathers: cyan).
- **Panel D:** The doubling rate of human mortality hazard, by birth decade cohort, measured using the increase of age-specific hazard between the ages of 65 and 70 (blue and red), or 75 and 80 (cyan and pink), for males (blue and cyan) and females (red and pink).