Cell, Volume *179*

Supplemental Information

Screening Human Embryos for Polygenic Traits Has Limited Utility

Ehud Karavani, Or Zuk, Danny Zeevi, Nir Barzilai, Nikos C. Stefanis, Alex Hatzimanolis, Nikolaos Smyrnis, Dimitrios Avramopoulos, Leonid Kruglyak, Gil Atzmon, Max Lam, Todd Lencz, and Shai Carmi

Methods S1

1 Background and model

We assume a couple has generated n embryos, and we would like to select the top-scoring embryo with respect to a given polygenic trait. We assume that the genetic architecture of the trait is infinitesimal, namely that there are numerous causal variants, uniformly distributed along the genome. Denote the value of the trait as z, the number of variants as N, the variance of the trait as σ_z^2, and the heritability as h^2, and assume that the trait has zero mean.

Mathematically, we assume an additive model, where for a given individual,

$$z = g + e = \sum_{i=1}^{N} \beta_i (\tilde{\kappa}_{i,p} + \tilde{\kappa}_{i,m}) + e. \tag{1}$$

In the above equation, $\tilde{\kappa}_{i,p} = \kappa_{i,p} - f_i$, where $\kappa_{i,p} \in \{0,1\}$ is the number of minor alleles at site i on the paternal chromosome and f_i is the minor allele frequency. $\tilde{\kappa}_{i,m}$ is similarly defined for the maternal chromosome. β_i is the additive effect size per allele.

The polygenic score for the trait is defined as

$$\mathrm{PS} = \sum_{i=1}^{N} \hat{\beta}_i (\tilde{\kappa}_{i,p} + \tilde{\kappa}_{i,m}), \tag{2}$$

where the $\hat{\beta}_i$s are the estimated effect sizes. We further assume that the trait can be modeled as

$$z = \mathrm{PS} + \epsilon. \tag{3}$$

The error term now represents both the environmental component as well as unaccounted-for genetic components. The proportion of variance of z explained by the polygenic score PS is denoted

$$r_{\mathrm{ps}}^2 = \frac{\mathrm{Var}\,(\mathrm{PS})}{\sigma_z^2}. \tag{4}$$

r_{ps} is also the correlation coefficient between the polygenic score and the trait. Eqs. (3) and (4) imply $\mathrm{Var}\,(\epsilon) = \sigma_z^2(1 - r_{\mathrm{ps}}^2)$.

Next, we make the following assumptions. First, we assume that there is no assortative mating. This implies that beyond linkage disequilibrium, there is no correlation between the contributions to the polygenic score from *(i)* the two homologous chromosomes of an individual, at the same locus; *(ii)* two chromosomes of spouses, at the same locus; *(iii)* two distinct loci, coming from the same chromosome; and *(iv)* two distinct loci, coming from either two homologous chromosomes or from chromosomes of spouses. While assortative mating was demonstrated for several polygenic traits [1, 2, 3], our empirical data shows that the correlation between polygenic scores of spouses is relatively small. Specifically, we found that the correlation in the polygenic scores for height between actual spouses was relatively low and did not reach statistical significance ($r = 0.12$, $P = 0.25$). The correlation for the polygenic scores for IQ was similarly low ($r = -0.03$, $P = 0.76$). Either way, since assortative mating is usually positive, our results may represent an upper bound for the utility of embryo selection.

Second, to avoid correlation due to linkage disequilibrium (LD), we write the polygenic score as a sum of K elements, where each element is the score in a single LD block,

$$\mathrm{PS} = \sum_{i=1}^{K} (\mathrm{PS}_{i,p} + \mathrm{PS}_{i,m}). \tag{5}$$

Above, $\mathrm{PS}_{i,p} = \sum_{k \in B_i} \hat{\beta}_k \tilde{\kappa}_{k,p}$, where B_i is the set of variants in block i, and similarly for $\mathrm{PS}_{i,m}$. Under the above assumption of no assortative mating, and assuming no correlation across LD blocks, this implies that for all $i \neq j$, the random variables $\mathrm{PS}_{i,p}$, $\mathrm{PS}_{i,m}$, $\mathrm{PS}_{j,p}$, $\mathrm{PS}_{j,m}$ are all uncorrelated. Moreover, $\mathrm{PS}_{i,p}, \mathrm{PS}_{i,m}$ for any one individual are uncorrelated with $\mathrm{PS}_{i,p}$ and $\mathrm{PS}_{i,m}$ in the spouse of that individual, for any block i. The LD blocks can be identified, e.g., as in [4].

We further assume that all blocks contribute equally to the variance (although this can be easily relaxed, leading to the same result). Thus, under the above model, we have

$$\mathrm{Var}\left(\mathrm{PS}_{i,p}\right) = \mathrm{Var}\left(\mathrm{PS}_{i,m}\right) = \sigma_z^2 \frac{r_{\mathrm{ps}}^2}{2K}, \tag{6}$$

as well as

$$\mathrm{E}\left(\mathrm{PS}\right) = \mathrm{E}\left(\mathrm{PS}_{i,p}\right) = \mathrm{E}\left(\mathrm{PS}_{i,m}\right) = 0. \tag{7}$$

Next, we consider the vector $\mathbf{PS} = (\mathrm{PS}^1, \ldots, \mathrm{PS}^n)$ of polygenic scores for n embryos. We assume that the distribution of the polygenic scores, PS, is normal in each embryo (due to the polygenic nature of most complex traits [5]), and that the joint distribution of the polygenic scores over n embryos is multivariate normal,

$$\mathbf{PS} = \left(\mathrm{PS}^1, \ldots, \mathrm{PS}^n\right) \sim \mathrm{MVN}(\boldsymbol{\mu}, \boldsymbol{\Sigma}), \tag{8}$$

with $\boldsymbol{\mu} = \mathbf{0}_n$ (a column vector of zeros of length n). The diagonal elements of the covariance matrix $\boldsymbol{\Sigma}$ are $\mathrm{Var}\left(\mathrm{PS}^i\right) = \sigma_z^2 r_{\mathrm{ps}}^2$ for all $i = 1, \ldots, n$. We will compute the off-diagonal covariances below (Section 2).

We define the *gain* G due to embryo selection as the difference between the polygenic score of the top-scoring embryo and the average scores of all embryos.

Mathematically,

$$G = \max\left(\mathrm{PS}^1, \ldots, \mathrm{PS}^n\right) - \frac{\mathrm{PS}^1 + \cdots + \mathrm{PS}^n}{n}. \tag{9}$$

The gain G is a random variable, with a sample space over all theoretical sets of n siblings. In the following, we will examine the statistical properties (e.g., mean and variance) of the gain as a function of n, σ_z^2, and r_{ps}^2.

For the mean gain, using Eq. (7),

$$\mathrm{E}(G) = \mathrm{E}\left(\max\left(\mathrm{PS}^1, \ldots, \mathrm{PS}^n\right)\right). \tag{10}$$

We derive approximate formulas for the mean gain in Section 3 and for the variance of the gain in Section 4. In Section 5 we derive the mean gain conditional on the parental scores, and in Section 6 we investigate prediction intervals for the actual trait of the selected embryo. We consider additional properties of the gain in Section 7 and selection for multiple traits in Section 8.

2 The covariance of the scores of siblings

In order to obtain the joint distribution of $\left(\mathrm{PS}^1, \ldots, \mathrm{PS}^n\right)$, we need to compute $\mathrm{Cov}\left(\mathrm{PS}^A, \mathrm{PS}^B\right)$, the covariance between the polygenic scores of two distinct embryos (or siblings), which we name A and B. For two individuals A and B with kinship coefficient Θ, standard quantitative genetics theory gives the covariance $\mathrm{Cov}(z_A, z_B) = 2\sigma_z^2\Theta h^2$ for a quantitative additive trait z with heritability h^2 under the infinitesimal model [6]. For full siblings, $\Theta = 1/4$, and thus $\mathrm{Cov}(z_A, z_B) = \sigma_z^2 h^2/2$. For completeness, we derive the corresponding result here for the polygenic scores, PS^A and PS^B.

Recall that we modeled the polygenic score as $\mathrm{PS} = \sum_{i=1}^{K}(\mathrm{PS}_{i,p} + \mathrm{PS}_{i,m})$, where $\mathrm{PS}_{i,p}$ is the score of the i^{th} LD block in the paternal chromosome and $\mathrm{PS}_{i,m}$ is the score from the maternal chromosome. For a pair of siblings and for a given LD block, their scores come from the same parental chromosome with probability 1/2, or from different parental chromosomes with probability 1/2. (We neglect the possibility of a recombination event taking place in the middle of an LD block, because, first, by definition, recombination is depleted within LD blocks, and second, the distance between crossovers is much greater than the distance between LD blocks [7].)

Consider the two homologous chromosomes of the father at block i. Denote the polygenic score of the first chromosome (say, grandpaternal) as $u_{i,1}$ and the score of the second chromosome (say, grandmaternal) as $u_{i,2}$. Similarly, denote the polygenic scores of the two maternal chromosomes as $v_{i,1}$ and $v_{i,2}$. For embryo A, denote by $p_{A,i}$ the choice of the paternal chromosome it has inherited at block i: $p_{A,i} = 1, 2$ with equal probability. Similarly, $m_{A,i} = 1, 2$ denotes the identity of the maternal chromosome transmitted to embryo A at block i. With the above notation, the polygenic score of embryo A can be written as:

$$\mathrm{PS}^A = \sum_{i=1}^{K} \left(u_{i,p_{A,i}} + v_{i,m_{A,i}} \right). \tag{11}$$

Similarly,

$$\mathrm{PS}^B = \sum_{i=1}^{K} \left(u_{i,p_{B,i}} + v_{i,m_{B,i}} \right). \tag{12}$$

The covariance between the scores of the two embryos is

$$\mathrm{Cov}\left(\mathrm{PS}^A, \mathrm{PS}^B\right) = \mathrm{Cov}\left(\sum_{i=1}^{K} \left(u_{i,p_{A,i}} + v_{i,m_{A,i}} \right), \sum_{i=1}^{K} \left(u_{i,p_{B,i}} + v_{i,m_{B,i}} \right) \right). \tag{13}$$

According to the assumptions of Section 1, there is no correlation between the scores of any two blocks on two chromosomes of spouses, or between distinct blocks on the same chromosome. Thus,

$$\mathrm{Cov}\left(\mathrm{PS}^A, \mathrm{PS}^B\right) = K \left[\mathrm{Cov}\left(u_{p_A}, u_{p_B}\right) + \mathrm{Cov}\left(v_{m_A}, v_{m_B}\right) \right], \tag{14}$$

where p_A, p_B, m_A, m_B are the identities of the chromosomes transmitted by the father/mother to embryos A and B at a representative block, and u_1, u_2, v_1, v_2 are the scores of the four parental chromosomes in that block. p_A, p_B, m_A, m_B are independent random variables taking the values 1 or 2 with equal probabilities. To compute the remaining terms, we invoke the law of total covariance, by conditioning on p_A, p_B or on m_A, m_B. For example,

$$\begin{aligned} \mathrm{Cov}\left(u_{p_A}, u_{p_B}\right) = {} & \mathrm{E}\left(\mathrm{Cov}\left(u_{p_A}, u_{p_B} | p_A, p_B\right)\right) \\ & + \mathrm{Cov}\left(\mathrm{E}\left(u_{p_A} | p_A, p_B\right), \mathrm{E}\left(u_{p_B} | p_A, p_B\right)\right). \end{aligned} \tag{15}$$

However, $\mathrm{E}\left(u_{p_A} | p_A, p_B\right) = \mathrm{E}\left(u_{p_B} | p_A, p_B\right) = 0$, and are both in general independent of p_A or p_B. Thus, the second term (covariance of expectations) vanishes. We can expand the first term as follows,

$$\begin{aligned} \mathrm{E}\left(\mathrm{Cov}\left(u_{p_A}, u_{p_B} | p_A, p_B\right)\right) = {} & \frac{1}{4}\mathrm{Cov}\left(u_1, u_1\right) + \frac{1}{4}\mathrm{Cov}\left(u_2, u_2\right) \\ & + \frac{1}{4}\mathrm{Cov}\left(u_1, u_2\right) + \frac{1}{4}\mathrm{Cov}\left(u_2, u_1\right). \end{aligned} \tag{16}$$

Again according to the assumptions of Section 1, there is no correlation between the scores of blocks from homologous chromosomes. Thus, the two terms in the second line vanish. Finally, using Eq. (6),

$$\mathrm{E}\left(\mathrm{Cov}\left(u_{p_A}, u_{p_B} | p_A, p_B\right)\right) = \frac{1}{4}\mathrm{Var}\left(u_1\right) + \frac{1}{4}\mathrm{Var}\left(u_2\right) = \sigma_z^2 \frac{r_{\mathrm{ps}}^2}{4K}. \tag{17}$$

A similar result holds for the maternal scores. Using Eq. (14),

$$\mathrm{Cov}\left(\mathrm{PS}^A, \mathrm{PS}^B\right) = \frac{1}{2}\sigma_z^2 r_{\mathrm{ps}}^2. \tag{18}$$

We have thus specified the distribution of the polygenic scores of the n embryos,

$$\mathbf{PS} = \left(\mathrm{PS}^1, \ldots, \mathrm{PS}^n\right) \sim \mathrm{MVN}(\boldsymbol{\mu}, \boldsymbol{\Sigma}), \tag{19}$$

where $\boldsymbol{\mu} = \mathbf{0}_n$ and $\boldsymbol{\Sigma}$ is an $n \times n$ covariance matrix with elements

$$\boldsymbol{\Sigma} = \sigma_z^2 r_{\mathrm{ps}}^2 \begin{pmatrix} 1 & \frac{1}{2} & \ldots & \frac{1}{2} \\ \frac{1}{2} & 1 & \ldots & \frac{1}{2} \\ \ldots & \ldots & \ldots & \ldots \\ \frac{1}{2} & \frac{1}{2} & \ldots & 1 \end{pmatrix}. \tag{20}$$

3 The mean score of the top-scoring embryo

Define $\mathrm{PS}_{\max} = \max\left(\mathrm{PS}^1, \ldots, \mathrm{PS}^n\right)$. The mean gain (as defined in Section 1) is the mean of the score of the top-scoring embryo, $\mathrm{E}\,(G) = \mathrm{E}\,(\mathrm{PS}_{\max})$ (Eq. (10)).

Written more generally, we would like to compute the mean of the maximum of n multivariate normal variables, denoted $\mathbf{PS} = \left(\mathrm{PS}^1, \ldots, \mathrm{PS}^n\right) \sim \mathrm{MVN}(\mathbf{0}_n, \boldsymbol{\Sigma})$, where the covariance matrix $\boldsymbol{\Sigma}$ is defined according to Eq. (20). We can write the covariance matrix also as $\boldsymbol{\Sigma} = \boldsymbol{\Sigma}_{\mathrm{ind}} + \boldsymbol{\Sigma}_{\mathrm{same}}$, where

$$\boldsymbol{\Sigma}_{\mathrm{ind}} = \frac{1}{2}\sigma_z^2 r_{\mathrm{ps}}^2 \begin{pmatrix} 1 & 0 & \ldots & 0 \\ 0 & 1 & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & \ldots & 1 \end{pmatrix} \tag{21}$$

and

$$\boldsymbol{\Sigma}_{\mathrm{same}} = \frac{1}{2}\sigma_z^2 r_{\mathrm{ps}}^2 \begin{pmatrix} 1 & 1 & \ldots & 1 \\ 1 & 1 & \ldots & 1 \\ \ldots & \ldots & \ldots & \ldots \\ 1 & 1 & \ldots & 1 \end{pmatrix} \tag{22}$$

Given this decomposition, we can write the distribution of the polygenic scores as a sum of two independent multivariate normal variables, $\mathbf{PS} = \boldsymbol{Y} + \boldsymbol{C}$, where

$$\boldsymbol{Y} = (y_1, \ldots, y_n) \sim \mathrm{MVN}(\mathbf{0}_n, \boldsymbol{\Sigma}_{\mathrm{ind}}) \tag{23}$$

and

$$\boldsymbol{C} = (c_1, \ldots, c_n) \sim \mathrm{MVN}(\mathbf{0}_n, \boldsymbol{\Sigma}_{\mathrm{same}}). \tag{24}$$

The covariance matrix $\boldsymbol{\Sigma}_{\mathrm{ind}}$ of $\boldsymbol{Y}$ is diagonal, and hence the variables in $\boldsymbol{Y}$ are independent. $\boldsymbol{C}$ has a constant covariance matrix $\boldsymbol{\Sigma}_{\mathrm{same}}$, which means that the correlation between all variables is 1. Thus, all elements of $\boldsymbol{C}$ are equal to the same normal variable,

$$c_1 \sim N\left(0, \frac{1}{2}\sigma_z^2 r_{\mathrm{ps}}^2\right) \text{ and } c_2 = c_3 = \cdots = c_n = c_1. \tag{25}$$

Since $\mathbf{PS} = \boldsymbol{Y} + \boldsymbol{C}$, we have

$$\begin{aligned} \mathrm{PS}_{\max} &= \max(y_1 + c_1, \ldots, y_n + c_n) \\ &= \max(y_1 + c_1, \ldots, y_n + c_1) \\ &= \max(y_1, \ldots, y_n) + c_1 \\ &= y_{\max} + c_1, \end{aligned} \tag{26}$$

where we defined $y_{\max} = \max(y_1, \ldots, y_n)$. The expectation of $\mathrm{PS}_{\max}$ is

$$\begin{aligned} \mathrm{E}\left(\mathrm{PS}_{\max}\right) &= \mathrm{E}\left(y_{\max}\right) + \mathrm{E}\left(c_1\right) \\ &= \mathrm{E}\left(y_{\max}\right). \end{aligned} \tag{27}$$

Therefore, the mean of the maximum of $\left(\mathrm{PS}^1, \ldots, \mathrm{PS}^n\right)$ is the same as the mean of the maximum of n *independent* normal variables with variance $\frac{1}{2}\sigma_z^2 r_{\mathrm{ps}}^2$ each.

For independent normal variables, we can calculate numerically the expectation of the maximum. Consider the maximum $R = \max(x_1, \ldots, x_n)$ of independent standard normals random variables $x_i \sim N(0,1)$. The cumulative distribution function of R is $\Phi_n(x) = [\Phi(x)]^n$ and the density of R is $\phi_n(x) = n\phi(x)\left[\Phi(x)\right]^{n-1}$, where Φ is the standard normal cumulative distribution function and ϕ is the standard normal density. The expectation of R is given by

$$\mathrm{E}\left(R\right) = \int_{-\infty}^{\infty} x\phi_n(x)dx. \tag{28}$$

While Eq. (28) does not result in a closed form expression, approximate analytical results are available. For example, extreme value theory shows that for large n [8, 9], R has an approximate *Gumbel* distribution, with parameters (μ_n, β_n) and CDF

$$\Phi_n(x) \approx \exp\left(-\exp\left(-\frac{x-\mu_n}{\beta_n}\right)\right), \tag{29}$$

where

$$\mu_n = \Phi^{-1}\left(1 - \frac{1}{n}\right), \tag{30}$$

$$\beta_n = \frac{1}{n\phi\left(\Phi^{-1}\left(1 - \frac{1}{n}\right)\right)}, \tag{31}$$

and Φ^{-1} is the inverse CDF of the standard normal variable. The mean of a Gumbel random variable is $\mu_n + \beta_n\gamma$, where γ is the Euler-Mascheroni constant ($\gamma \approx 0.577$). Thus,

$$\mathrm{E}\left(R\right) \approx \Phi^{-1}\left(1 - \frac{1}{n}\right) + \frac{\gamma}{n\phi\left(\Phi^{-1}(1 - \frac{1}{n})\right)}. \tag{32}$$

Finally, as we have $G = \sigma R$ with $\sigma^2 = \frac{1}{2}\sigma_z^2 r_{\mathrm{ps}}^2$,

$$\mathrm{E}\left(G\right) \approx \frac{\sigma_z r_{\mathrm{ps}}}{\sqrt{2}}\left[\Phi^{-1}\left(1 - \frac{1}{n}\right) + \frac{\gamma}{n\phi\left(\Phi^{-1}(1 - \frac{1}{n})\right)}\right]. \tag{33}$$

We found this equation to fit the simulations reasonably well (Supplementary Figure 1).

To gain more insight into the behavior of the gain for large n, we use the result that for $n \to \infty$ [8, 9]

$$\mu_n = \mathcal{O}\left(\sqrt{\log n}\right),\ \beta_n = \mathcal{O}\left(\frac{1}{\sqrt{\log n}}\right). \tag{34}$$

Thus, to leading order in n,

$$\mathrm{E}\left(R\right) \propto \sqrt{\log n}. \tag{35}$$

Numerically, we found the best fit to Eq. (35) (over n from 1 to 50) when the coefficient of proportion was ≈ 1.09. Thus, the mean gain can be approximated as

$$\mathrm{E}\left(G\right) \approx 1.09\frac{\sigma_z r_{\mathrm{ps}}}{\sqrt{2}}\sqrt{\log n} = 0.77\sigma_z r_{\mathrm{ps}}\sqrt{\log n}. \tag{36}$$

Due to its simple functional form, we reported Eq. (36) as Eq. (1) of the main text and used it to produce main Figures 1 and 2.

4 The variance of the gain

To compute the variance of the gain, we start from its definition (Eq. (9)) and use the decomposition of the score (Eq. (26)),

$$\begin{aligned}
\mathrm{Var}\left(G\right) &= \mathrm{Var}\left(\mathrm{PS}_{\max} - \frac{\mathrm{PS}^1 + \cdots + \mathrm{PS}^n}{n}\right) \\
&= \mathrm{Var}\left(y_{\max} + c_1 - \frac{1}{n}\sum_{i=1}^{n}(y_i + c_1)\right) \\
&= \mathrm{Var}\left(y_{\max} - \frac{1}{n}\sum_{i=1}^{n} y_i\right) \\
&= \mathrm{Var}\left(y_{\max}\right) + \frac{1}{n}\mathrm{Var}\left(y_1\right) - 2\mathrm{Cov}\left(y_{\max}, \frac{1}{n}\sum_{i=1}^{n} y_i\right) \\
&= \mathrm{Var}\left(y_{\max}\right) + \frac{1}{2n}\sigma_z^2 r_{\mathrm{ps}}^2 - 2\mathrm{E}\left(y_1 y_{\max}\right).
\end{aligned} \tag{37}$$

The variance of $y_{\max}$ can be calculated numerically by

$$\mathrm{Var}\left(y_{\max}\right) = \frac{1}{2}\sigma_z^2 r_{\mathrm{ps}}^2 \left[\int_{-\infty}^{\infty} x^2 \phi_n(x)dx - \left(\int_{-\infty}^{\infty} x\phi_n(x)dx\right)^2\right]. \tag{38}$$

Extreme value theory can also provide an expression for the variance of $y_{\max}$. The variance of a *Gumbel*(μ, β) random variable is $\frac{\pi^2\beta^2}{6}$. Thus,

$$\mathrm{Var}\left(y_{\max}\right) \approx \frac{1}{2}\sigma_z^2 r_{\mathrm{ps}}^2 \times \frac{\pi^2}{6\left[n\phi\left(\Phi^{-1}(1-\frac{1}{n})\right)\right]^2}. \tag{39}$$

The last term in Eq. (37) can be computed as follows,

$$\begin{aligned}\mathrm{E}\left(y_1 y_{\max}\right) = {} & P(y_1 < y_{\max})\mathrm{E}\left(y_1 y_{\max} | y_1 < y_{\max}\right) \\ & + P(y_1 = y_{\max})\mathrm{E}\left(y_1^2 | y_1 = y_{\max}\right). \end{aligned} \tag{40}$$

Due to symmetry, we have $P(y_1 = y_{\max}) = \frac{1}{n}$ and $P(y_1 < y_{\max}) = 1 - \frac{1}{n}$. Thus,

$$P(y_1 = y_{\max})\mathrm{E}\left(y_1^2 | y_1 = y_{\max}\right) = \frac{1}{n}\mathrm{E}\left(y_{\max}^2\right). \tag{41}$$

Assume next that all y_j's have unit variance and are thus standard normals (we will bring back the variance later). As in the previous section, denote the cumulative and density distribution functions of the maximum of n standard normal variables by Φ_n and ϕ_n, respectively.

To compute $\mathrm{E}\left(y_1 y_{\max} | y_1 < y_{\max}\right)$, we define $\delta_{\{y_1 < y_{\max}\}}$ as the indicator for the event $y_1 < y_{\max}$. In the regime $y_1 < y_{\max}$, $y_{\max}$ is the maximum of $(y_2, \ldots, y_n)$, and can thus be redefined as a random variable with density ϕ_{n-1} that is independent of y_1. Thus,

$$\begin{aligned}\mathrm{E}\left(y_1 y_{\max} | y_1 < y_{\max}\right) &= \mathrm{E}\left(y_1 y_{\max}\delta_{\{y_1 < y_{\max}\}}\right)/P(y_1 < y_{\max}) \\ &= \frac{1}{1-\frac{1}{n}}\int_{-\infty}^{\infty} dy_{\max}\int_{-\infty}^{y_{\max}} dy_1 [y_{\max} y_1 \phi_{n-1}(y_{\max})\phi(y_1)]. \end{aligned} \tag{42}$$

The term $\left(1 - \frac{1}{n}\right)$ cancels when multiplying by $P(y_1 < y_{\max})$,

$$\begin{aligned} & P(y_1 < y_{\max})\mathrm{E}\left(y_1 y_{\max} | y_1 < y_{\max}\right) \\ &= \int_{-\infty}^{\infty} dy_{\max}\int_{-\infty}^{y_{\max}} dy_1 \left[y_{\max} y_1 \phi_{n-1}(y_{\max})\phi(y_1)\right] \\ &= \int_{-\infty}^{\infty} y_{\max}\phi_{n-1}(y_{\max})\left[\int_{-\infty}^{y_{\max}} y_1\phi(y_1)dy_1\right]dy_{\max} \\ &= -\int_{-\infty}^{\infty} y_{\max}\phi_{n-1}(y_{\max})\phi(y_{\max})dy_{\max}. \end{aligned} \tag{43}$$

The last equality results from

$$\int t\phi(t)dt = \int \frac{t}{\sqrt{2\pi}}e^{-t^2/2}dt = -\frac{e^{-t^2/2}}{\sqrt{2\pi}} = -\phi(t). \tag{44}$$

Plugging Eqs. (41) and (43) into Eq. (40) we have

$$\begin{aligned}\mathrm{E}\left(y_1 y_{\max}\right) &= \int_{-\infty}^{\infty}\left[\frac{1}{n}y_{\max}^2 \cdot \phi_n(y_{\max}) - y_{\max}\phi_{n-1}(y_{\max})\phi(y_{\max})\right]dy_{\max} \\ &= \frac{1}{n}\int_{-\infty}^{\infty}\left[\Phi_n(y_{\max}) - \phi_n(y_{\max})y_{\max}\right]' dy_{\max} \\ &= \frac{1}{n}\left[\Phi_n(y_{\max}) - \phi_n(y_{\max})y_{\max}\right]_{y_{\max}=-\infty}^{\infty} \\ &= \frac{1}{n}(1-0) = \frac{1}{n}. \end{aligned} \tag{45}$$

In the second line above, we used

$$\begin{aligned}\frac{d[\Phi_n(t)-\phi_n(t)t]}{dt} &= \phi_n(t) - \left[\phi_n(t) + t\frac{d\phi_n(t)}{dt}\right] \\ &= -t\frac{d}{dt}\left[n\phi(t)\left[\Phi(t)\right]^{n-1}\right] \\ &= -nt\frac{d\phi(t)}{dt}\left[\Phi(t)\right]^{n-1} - nt\phi(t)\phi_{n-1}(t) \\ &= -nt\frac{d}{dt}\left[\frac{e^{-t^2/2}}{\sqrt{2\pi}}\right]\left[\Phi(t)\right]^{n-1} - nt\phi(t)\phi_{n-1}(t) \\ &= t^2 n\phi(t)\left[\Phi(t)\right]^{n-1} - nt\phi(t)\phi_{n-1}(t) \\ &= t^2\phi_n(t) - nt\phi_{n-1}(t)\phi(t). \end{aligned} \tag{46}$$

Eq. (45) applies to standard normal random variables. In our case, the y_i's have variance $\frac{1}{2}\sigma_z^2 r_{\text{ps}}^2$, and thus,

$$\mathrm{E}\left(y_1 y_{\max}\right) = \frac{1}{2n}\sigma_z^2 r_{\text{ps}}^2. \tag{47}$$

Using Eqs. (37) and (38), the variance of the gain is

$$\begin{aligned}\mathrm{Var}\left(G\right) &= \mathrm{Var}\left(y_{\max}\right) + \frac{1}{2n}\sigma_z^2 r_{\text{ps}}^2 - 2\mathrm{E}\left(y_1 y_{\max}\right) \\ &= \mathrm{Var}\left(y_{\max}\right) + \frac{1}{2n}\sigma_z^2 r_{\text{ps}}^2 - \frac{1}{n}\sigma_z^2 r_{\text{ps}}^2 \\ &= \mathrm{Var}\left(y_{\max}\right) - \frac{1}{2n}\sigma_z^2 r_{\text{ps}}^2 \\ &= \frac{1}{2}\sigma_z^2 r_{\text{ps}}^2\left[\int_{-\infty}^{\infty} x^2\phi_n(x)dx - \left(\int_{-\infty}^{\infty} x\phi_n(x)dx\right)^2 - \frac{1}{n}\right]. \end{aligned} \tag{48}$$

Or, using the Gumbel approximation (Eq. (39)),

$$\mathrm{Var}\left(G\right) \approx \frac{1}{2}\sigma_z^2 r_{\text{ps}}^2\left[\frac{\pi^2}{6\left[n\phi\left(\Phi^{-1}(1-\frac{1}{n})\right)\right]^2} - \frac{1}{n}\right]. \tag{49}$$

In our simulations with random couples for height, we found that the variance has a maximum around $n \approx 10-15$ (Figure 1 below). The exact integral based on Eq. (48) has a maximum around $n \approx 5$, and then it underestimates the variance until $n \gtrsim 40$. The approximate expression of Eq. (49) shows no maximum, but becomes close to the simulations around $n \approx 15$.

5 The mean gain conditional on the parental scores or phenotypes

The actual gain realized for a specific embryo selection procedure has two sources of variation: first, variation between families (i.e. the genetics of the parents),

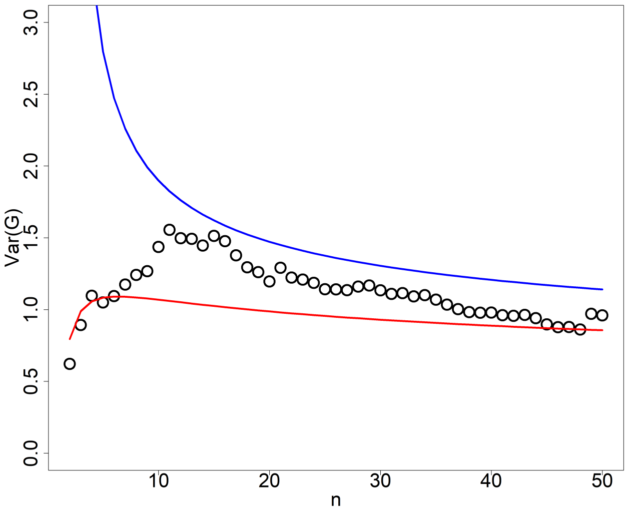

Figure 1: The variance in the gain vs the number of embryos n. Circles show simulations for height (the Longevity cohort; see the main text), where for each n, the variance was computed over 100 random couples. The red line is the theoretical curve, Eq. (48), evaluated numerically, where we used $\sigma_z = 6$cm and $r^2_{\text{ps}} = 0.248$. The blue line is the approximate expression based on the Gumbel distribution, Eq. (49).

and then variation between random embryos that can be generated within the family. In Section 3, we have calculated the mean gain over the *entire population*, while averaging over these two sources of variation. It is of interest to consider the expected gain for a *specific family*, i.e., exclude the first source of randomness, and determine how the mean gain depends on the parental polygenic scores or phenotypes.

We start by assuming that the polygenic scores of the parents are known for each block of the genome (recall Section 1). This scenario may be realized when the parents are genotyped and phased. These calculations will be used later when only the total polygenic scores (or phenotypes) are available.

Recall our notation from Section 2: denote the polygenic score of the father on his paternal chromosome at block i as $u_{i,1}$, and the score of the father on his maternal chromosome as $u_{i,2}$. Similarly, denote the polygenic scores of the two maternal chromosomes as $v_{i,1}$ and $v_{i,2}$. Denote by p_i the choice of the paternal chromosome transmitted to an embryo at block i ($p_i = 1, 2$ with equal probability), and define $m_i = 1, 2$ similarly. Finally, let $\text{PS}_i \equiv u_{i,p_i} + v_{i,m_i}$ be the score of the embryo at block i.

With this notation, the polygenic score of the embryo can be written as

$$\mathrm{PS} = \sum_{i=1}^{K} \mathrm{PS}_i = \sum_{i=1}^{K} \left(u_{i,p_i} + v_{i,m_i}\right). \tag{50}$$

Define $\boldsymbol{u} = (u_{1,1}, \ldots, x_{K,1}, u_{1,2}, \ldots, u_{K,2})$ and $\boldsymbol{v} = (v_{1,1}, \ldots, v_{K,1}, v_{1,2}, \ldots, v_{K,2})$. The mean score of the embryo, given the scores of the parents, is

$$\mathrm{E}\left(\mathrm{PS}|\boldsymbol{u}, \boldsymbol{v}\right) = \frac{1}{2} \sum_{i=1}^{K} (u_{i,1} + u_{i,2} + v_{i,1} + v_{i,2}). \tag{51}$$

To calculate the variance of PS (given the scores of the parents), we consider first the variance at each block,

$$\begin{aligned}
\mathrm{Var}\left(\mathrm{PS}_i|\boldsymbol{u}, \boldsymbol{v}\right) &= \mathrm{E}\left(\mathrm{PS}_i^2|u_{i,1}, u_{i,2}, v_{i,1}, v_{i,2}\right) - \mathrm{E}\left(\mathrm{PS}_i|u_{i,1}, u_{i,2}, v_{i,1}, v_{i,2}\right)^2 \\
&= \frac{1}{4}(u_{i,1} + v_{i,1})^2 + \frac{1}{4}(u_{i,1} + v_{i,2})^2 + \frac{1}{4}(u_{i,2} + v_{i,1})^2 + \frac{1}{4}(u_{i,2} + v_{i,2})^2 \\
&- \frac{1}{4}(u_{i,1} + u_{i,2} + v_{i,1} + v_{i,2})^2 \\
&= \frac{1}{4}\left(u_{i,1}^2 - 2u_{i,1}u_{i,2} + u_{i,2}^2 + v_{i,1}^2 - 2v_{i,1}v_{i,2} + v_{i,2}^2\right) \\
&= \frac{1}{4}(u_{i,1} - u_{i,2})^2 + \frac{1}{4}(v_{i,1} - v_{i,2})^2 \\
&= \frac{1}{4}\tilde{u}_i^2 + \frac{1}{4}\tilde{v}_i^2,
\end{aligned} \tag{52}$$

where $\tilde{u}_i \equiv u_{i,1} - u_{i,2}$ is the difference between the two *grandparental* polygenic scores that were transmitted to the father. $\tilde{v}_i \equiv v_{i,1} - v_{i,2}$ is similarly defined for the mother.

Next, as opposed to Section 2, here the polygenic scores per block are fixed, and thus we need to take into account the covariances between blocks.

$$\mathrm{Cov}\left(\mathrm{PS}_i, \mathrm{PS}_j|\boldsymbol{u}, \boldsymbol{v}\right) = \mathrm{Cov}\left(u_{i,p_i} + v_{i,m_i}, u_{j,p_j} + v_{j,m_j}|\boldsymbol{u}, \boldsymbol{v}\right). \tag{53}$$

The scores transmitted from the father and those transmitted from the mother are independent and their covariance vanishes. We next compute the covariance for scores transmitted from the same parent in different blocks. To compute the covariance for the father, we condition on p_i and p_j, the identity of the paternal chromosomes transmitted from the father at blocks i and j.

$$\begin{aligned}
\mathrm{Cov}\left(u_{i,p_i}, u_{j,p_j}|\boldsymbol{u}\right) &= \mathrm{E}\left(\mathrm{Cov}\left(u_{i,p_i}, u_{j,p_j}|p_i, p_j, \boldsymbol{u}\right)\right) \\
&+ \mathrm{Cov}\left(\mathrm{E}\left(u_{i,p_i}|p_i, p_j, \boldsymbol{u}\right), \mathrm{E}\left(u_{j,p_j}|p_i, p_j, \boldsymbol{u}\right)\right) \\
&= \mathrm{Cov}\left(\mathrm{E}\left(u_{i,p_i}|p_i, \boldsymbol{u}\right), \mathrm{E}\left(u_{j,p_j}|p_j, \boldsymbol{u}\right)\right),
\end{aligned} \tag{54}$$

because for given p_i and p_j, u_{i,p_i} and u_{j,p_j} are constants with zero covariance. To proceed, we denote by c_{ij} the probability that $p_i \neq p_j$, i.e., the probability of an odd number of crossovers between the two blocks ($c_{ij} = 1/2$ for unlinked blocks and $c_{ii} = 0$).

$$\begin{aligned}
&\mathrm{Cov}\left(\mathrm{E}\left(u_{i,p_i}|p_i,\boldsymbol{u}\right),\mathrm{E}\left(u_{j,p_j}|p_j,\boldsymbol{u}\right)\right)\\
&= \mathrm{Cov}\left(u_{i,p_i},u_{j,p_j}|p_i,p_j,\boldsymbol{u}\right)\\
&= \mathrm{E}\left(u_{i,p_i}u_{j,p_j}|p_i,p_j,\boldsymbol{u}\right)-\mathrm{E}\left(u_{i,p_i}|p_i,\boldsymbol{u}\right)\mathrm{E}\left(u_{j,p_j}|p_j,\boldsymbol{u}\right)\\
&= \frac{1}{2}(1-c_{ij})u_{i,1}u_{j,1}+\frac{1}{2}c_{ij}u_{i,1}u_{j,2}+\frac{1}{2}(1-c_{ij})u_{i,2}u_{j,2}+\frac{1}{2}c_{ij}u_{i,2}u_{j,1}\\
&- \frac{1}{4}(u_{i,1}+u_{i,2})(u_{j,1}+u_{j,2}). \qquad (55)
\end{aligned}$$

After some algebra, we obtain

$$\mathrm{Cov}\left(\mathrm{E}\left(u_{i,p_i}|p_i,p_j,\boldsymbol{u}\right),\mathrm{E}\left(u_{j,p_j}|p_i,p_j,\boldsymbol{u}\right)\right)=\frac{1}{4}(1-2c_{ij})\tilde{u}_i\tilde{u}_j. \qquad (56)$$

Thus (using Eq. (53)),

$$\mathrm{Cov}\left(\mathrm{PS}_i,\mathrm{PS}_j|\boldsymbol{u},\boldsymbol{v}\right)=\frac{1}{4}(1-2c_{ij})\left(\tilde{u}_i\tilde{u}_j+\tilde{v}_i\tilde{v}_j\right). \qquad (57)$$

Finally, we can write the variance of the entire polygenic score (using Eqs. (50), (52), and (57)),

$$\begin{aligned}
\mathrm{Var}\left(\mathrm{PS}|\boldsymbol{u},\boldsymbol{v}\right) &= \sum_{i=1}^{K}\mathrm{Var}\left(\mathrm{PS}_i|\boldsymbol{u},\boldsymbol{v}\right)+\sum_{i=1}^{K}\sum_{j=1,j\neq i}^{K}\mathrm{Cov}\left(\mathrm{PS}_i,\mathrm{PS}_j|\boldsymbol{u},\boldsymbol{v}\right)\\
&= \frac{1}{4}\sum_{i=1}^{K}\left[\tilde{u}_i^2+\tilde{v}_i^2+\sum_{j=1,j\neq i}^{K}(1-2c_{ij})\left(\tilde{u}_i\tilde{u}_j+\tilde{v}_i\tilde{v}_j\right)\right]\\
&= \frac{1}{4}\sum_{i,j=1}^{K}(1-2c_{ij})\left(\tilde{u}_i\tilde{u}_j+\tilde{v}_i\tilde{v}_j\right). \qquad (58)
\end{aligned}$$

We thus conclude that the variance of the score depends only on the grandparental differences, which we denote as vectors $\tilde{\boldsymbol{u}} = (\tilde{u}_1,\ldots,\tilde{u}_K)$ and $\tilde{\boldsymbol{v}} = (\tilde{v}_1,\ldots,\tilde{v}_K)$.

Defining as before PS^i as the score of embryo i, the mean conditional gain is defined as

$$\begin{aligned}
\mathrm{E}\left(G|\boldsymbol{u},\boldsymbol{v}\right) &= \mathrm{E}\left(\mathrm{PS}_{\max}|\boldsymbol{u},\boldsymbol{v}\right)-\mathrm{E}\left(\frac{1}{n}\sum_{i=1}^{n}\mathrm{PS}^i|\boldsymbol{u},\boldsymbol{v}\right) \qquad (59)\\
&= \mathrm{E}\left(\max\left(\mathrm{PS}^1,\ldots,\mathrm{PS}^n\right)|\boldsymbol{u},\boldsymbol{v}\right)-\mathrm{E}\left(\mathrm{PS}|\boldsymbol{u},\boldsymbol{v}\right)\\
&= \mathrm{E}\left(\max\left(\widetilde{\mathrm{PS}}^1,\ldots,\widetilde{\mathrm{PS}}^n\right)|\boldsymbol{u},\boldsymbol{v}\right), \qquad (60)
\end{aligned}$$

where we defined $\widetilde{\mathrm{PS}}^i = \mathrm{PS}^i-\mathrm{E}\left(\mathrm{PS}|\boldsymbol{u},\boldsymbol{v}\right)$. Now note that $\mathrm{E}\left(\widetilde{\mathrm{PS}}^i|\boldsymbol{u},\boldsymbol{v}\right)=0$ and $\mathrm{Var}\left(\widetilde{\mathrm{PS}}^i|\boldsymbol{u},\boldsymbol{v}\right)=\mathrm{Var}\left(\mathrm{PS}|\boldsymbol{u},\boldsymbol{v}\right)$, as calculated in Eq. (58). Further, the $\widetilde{\mathrm{PS}}^i$'s are independent conditional on the parental scores $\boldsymbol{u}, \boldsymbol{v}$, because, given the pair of scores per block of each parent, knowledge of the score of one embryo does not provide any information on the score of another. Thus, we are back to the problem of the mean of the maximum of n independent normal variables with zero mean each, now with variances given by Eq. (58). From here on, we can use the same approximations for the mean of the maximum as we used in Section 3. For example, we can write

$$\mathrm{E}\left(G|\boldsymbol{u},\boldsymbol{v}\right) \approx \frac{1.09}{2}\sqrt{\left[\sum_{i,j=1}^{K}\left(1-2c_{ij}\right)\left(\tilde{u}_i\tilde{u}_j+\tilde{v}_i\tilde{v}_j\right)\right]\log n}. \tag{61}$$

Thus, the mean gain does not depend on the scores of the parents, but rather on the differences between the scores of the two grandparents from each side. Each combination of scores with the same differences, regardless of the absolute magnitude of the scores, will yield the same mean gain for that family.

5.1 Conditioning on the total parental polygenic scores

Next, we would like to compute the mean gain given not the entire vector of scores per block of each parent, but given their total scores. In other words, assume that we know the paternal total score $u=\sum_{i=1}^{K}(u_{i,1}+u_{i,2})$ and the maternal total score $v=\sum_{i=1}^{K}(v_{i,1}+v_{i,2})$. We will compute the conditional mean gain $\mathrm{E}\left(G|u,v\right)$.

First, assume that the scores per block, $u_{1,1},u_{1,2},\ldots,u_{K,1},u_{K,2}$, are normally distributed in the population. This should be approximately satisfied for highly polygenic traits in the absence of alleles of very large effects. The scores are independent by the assumptions on Section 1. Thus, $\tilde{u}_i=u_{i,1}-u_{i,2}$ and $u=\sum_{j=1}^{K}(u_{j,1}+u_{j,2})$ are jointly normal variables. Their covariance is

$$\begin{aligned}\mathrm{Cov}\left(\tilde{u}_i,u\right) &= \mathrm{Cov}\left(u_{i,1}-u_{i,2},\sum_{j=1}^{K}(u_{j,1}+u_{j,2})\right)\\ &= \mathrm{Cov}\left(u_{i,1},\sum_{j=1}^{K}(u_{j,1}+u_{j,2})\right)-\mathrm{Cov}\left(u_{i,2},\sum_{j=1}^{K}(u_{j,1}+u_{j,2})\right)\\ &= \mathrm{Var}\left(u_{i,1}\right)-\mathrm{Var}\left(u_{i,2}\right)=0.\end{aligned} \tag{62}$$

To arrive at the last line, we used the assumption of independence across loci and across chromosomes from Section 1. As $\tilde{u}_i$ and u are jointly normal and uncorrelated, they must be independent, for all i. The same is true for $\tilde{v}_i$ and v. Thus, the grandparental differences $\tilde{\boldsymbol{u}}$ and $\tilde{\boldsymbol{v}}$ are independent of u and v.

Given this result, we can calculate $\mathrm{E}\left(G|u,v\right)$ by conditioning on the difference vectors $\tilde{\boldsymbol{u}}$ and $\tilde{\boldsymbol{v}}$, and using the law of total expectation,

$$\begin{aligned} \mathrm{E}\left(G|u,v\right) &= \mathrm{E}\left(\mathrm{E}\left(G|\tilde{\boldsymbol{u}},\tilde{\boldsymbol{v}},u,v\right)|u,v\right) \\ &= \mathrm{E}\left(\mathrm{E}\left(G|\tilde{\boldsymbol{u}},\tilde{\boldsymbol{v}}\right)|u,v\right) \\ &= \mathrm{E}\left(\mathrm{E}\left(G|\tilde{\boldsymbol{u}},\tilde{\boldsymbol{v}}\right)\right) \\ &= \mathrm{E}\left(G\right). \end{aligned} \tag{63}$$

In the second line, we used the fact that the mean gain depends only on the difference vectors (Eq. (61)). In the third, we used the independence of the differences and the total scores (recall that the outer expectation is over $\tilde{\boldsymbol{u}}$ and $\tilde{\boldsymbol{v}}$). In the last line, we used again the law of total expectation. We have thus shown that the mean gain is independent of the total polygenic scores of the parents, u and v.

Intuitively, the reason for the independence is that knowledge of the sum of the parental scores does not provide information regarding the *difference* between the *grandparental* scores, and it is only these grandparental differences that can generate differences between embryos. In Figure 2, we show empirically that for our height and IQ data, the gain indeed seems independent of the parental polygenic scores.

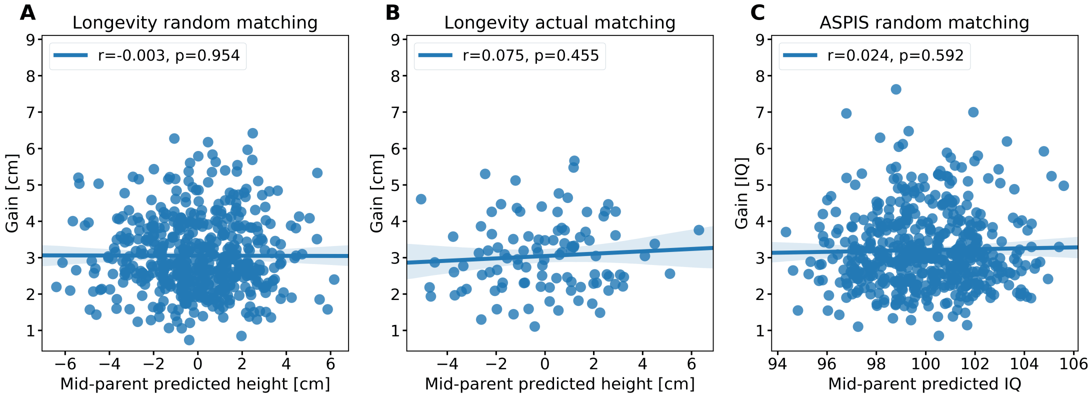

Figure 2: The gain in embryo selection vs the mid-parental *predicted* traits based on their polygenic scores. The height was corrected for sex and age. (A) Random mating for height. (B) Actual couples for height. (C) Random mating for IQ. The gain was calculated over $n = 10$ embryos. The correlation coefficient and its associated P-value are shown at the top of each panel.

5.2 Conditioning on the parental traits

We have so far studied the conditional gain given the polygenic scores (per block or total) of the parents. Suppose that the scores are not available, but that the parental phenotypes are observable (this is almost always possible, in particular for traits such as height or BMI). Denote the paternal and maternal traits as z^p, z^m, respectively, and recall that u and v are the parental polygenic scores. By Eq. (3), we have

$$\begin{aligned} z^p &= u + \epsilon^p, \\ z^m &= v + \epsilon^m, \end{aligned} \tag{64}$$

where $\mathrm{Var}\left(\epsilon^p\right) = \mathrm{Var}\left(\epsilon^m\right) = \sigma_z^2(1 - r_{\mathrm{ps}}^2)$. Conditioning on the parental scores u and v and using the law of total expectation,

$$\begin{aligned} \mathrm{E}\left(G|z^p, z^m\right) &= \mathrm{E}\left(\mathrm{E}\left(G|u, v, z^p, z^m\right)|z^p, z^m\right) \\ &= \mathrm{E}\left(\mathrm{E}\left(G|u, v\right)|z^p, z^m\right) \\ &= \mathrm{E}\left(\mathrm{E}\left(G\right)|z^p, z^m\right) \\ &= \mathrm{E}\left(G\right). \end{aligned} \tag{65}$$

In the second line above, we used the fact that $\mathrm{E}\left(G|u, v, z^p, z^m\right) = \mathrm{E}\left(G|u, v\right)$, because the gain is determined by the polygenic scores of the embryos, and thus, given the scores of the parents, there is no additional information from their phenotypes. We then used the independence of $\mathrm{E}\left(G|u, v\right)$ on u and v (Eq. (63)). Hence, conditioning on the parental phenotypes does not change the mean gain. In Figure 3, we show empirically that for our height and IQ data, the gain seems independent of the parental trait values.

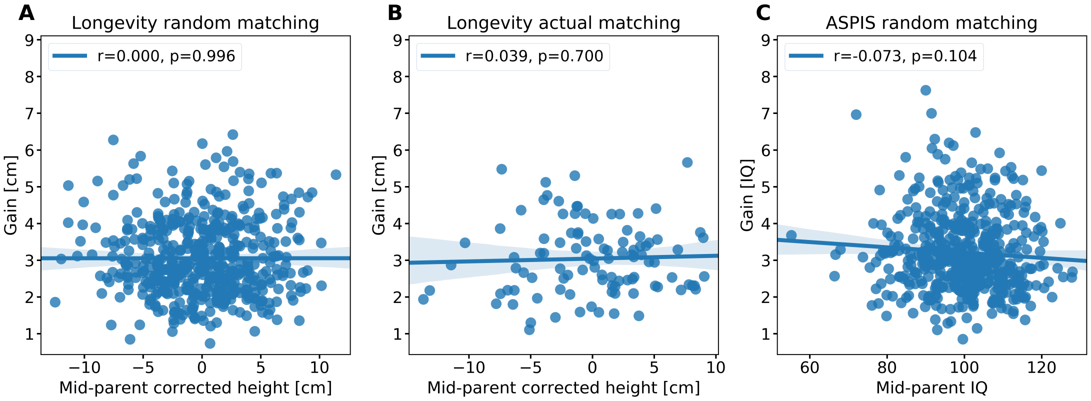

Figure 3: The gain in embryo selection vs the mid-parental trait value. The figure is the same as Figure 2, except that the x-axis here is the average of the actual traits (residual height or IQ) of the two parents.

6 A prediction interval for the phenotype of the top-scoring embryo

The actual value of the trait of the top-scoring embryo, $z_{\max}$, may differ considerably from that predicted by its polygenic score or by the mean gain. We have

$$z_{\max} = \mathrm{PS}_{\max} + \epsilon. \tag{66}$$

Following Section 1, ϵ has zero mean and variance

$$\mathrm{Var}\left(\epsilon\right) = \sigma_z^2 \left(1 - r_{\mathrm{ps}}^2\right). \tag{67}$$

Given the PS of the top embryo, and assuming a normal distribution for ϵ, a 95% prediction interval for the actual value of the trait will be approximately

$$\left[\mathrm{PS}_{\max} - 1.96\sigma_z\sqrt{1 - r_{\mathrm{ps}}^2}, \mathrm{PS}_{\max} + 1.96\sigma_z\sqrt{1 - r_{\mathrm{ps}}^2}\right]. \tag{68}$$

Eq. (68) is Eq. (2) in the main text.

In a naïve calculation for no selection, we assume no information is available regarding the embryo, and thus, the 95% prediction interval would be

$$\left[-1.96\sigma_z, 1.96\sigma_z\right], \tag{69}$$

as for any normal variable with zero mean and variance σ_z^2. However, the phenotype can be predicted based on the traits of the parents. Denote the trait of an offspring as z_o and the mid-parental trait value (i.e., the average of the (sex-adjusted) trait between the two parents) as z_{mp}. A well-known result in quantitative genetics is that the slope of the regression of z_o on z_{mp} is equal to the heritability h^2 [6]. The correlation coefficient is the product of the slope and the ratio of the standard deviations, $r = h^2 \frac{\sigma_{\mathrm{mp}}}{\sigma_o}$. But $\sigma_o^2 = \sigma_z^2$ and $\sigma_{\mathrm{mp}}^2 = \frac{\sigma_z^2}{2}$. Thus, $r = h^2 \frac{\sigma_z/\sqrt{2}}{\sigma_z} = \frac{h^2}{\sqrt{2}}$. The proportion of variance explained is $r^2 = \frac{h^4}{2}$ (see also, e.g., [10]), and the remaining variance is $\sigma_z^2\left(1 - \frac{h^4}{2}\right)$. Thus, a more realistic 95% prediction interval for the case of no selection would be

$$\left[h^2 z_{\mathrm{mp}} - 1.96\sigma_z\sqrt{1 - \frac{h^4}{2}}, h^2 z_{\mathrm{mp}} + 1.96\sigma_z\sqrt{1 - \frac{h^4}{2}}\right]. \tag{70}$$

In theory, having both the mid-parental value and the offspring's PS may lead to a more accurate prediction, with a narrower prediction interval, even for the case of selection. Prediction in this setting is in general non-trivial. However, the combination of both the PS and the mid-parental value cannot explain more variance than implicated by the heritability. Thus, the proportion of variance explained by all of the available data (PS and parents' trait value) can be anything within the range $\left[\max\left(r_{\mathrm{ps}}^2, \frac{h^4}{2}\right), h^2\right]$, i.e., it is at least the best of the two predictors, but no higher than the heritability.

7 Properties of the top-scoring embryo

7.1 The mean difference between the top-ranked trait and the trait of the top-scoring embryo

In the main text, we analyzed real large nuclear families. When reduced to $n = 7$ children per family, we found that the average height difference between the tallest child and the child with the maximal PS was 3.0cm. To determine the expectation based on our quantitative model, consider n siblings, whose polygenic scores are modeled as a multivariate normal variable, $\left(\mathrm{PS}^1, \dots, \mathrm{PS}^n\right) \sim \mathrm{MVN}\left(\mathbf{0}_n, \mathbf{\Sigma}\right)$ with $\mathbf{\Sigma}$ defined in Eq. (20).

We assume that the phenotypes, $z_1, \dots, z_n$, can be modeled as

$$\boldsymbol{z} = (z_1, \dots, z_n) = \boldsymbol{g} + \boldsymbol{e}, \tag{71}$$

where $\boldsymbol{g} \sim \mathrm{MVN}\left(\mathbf{0}_n, \mathbf{\Sigma}_g\right)$ and $\boldsymbol{e} \sim \mathrm{MVN}\left(\mathbf{0}_n, \mathbf{\Sigma}_e\right)$, with

$$\mathbf{\Sigma}_g = \sigma_z^2 h^2 \begin{pmatrix} 1 & \frac{1}{2} & \dots & \frac{1}{2} \\ \frac{1}{2} & 1 & \dots & \frac{1}{2} \\ \dots & \dots & \dots & \dots \\ \frac{1}{2} & \frac{1}{2} & \dots & 1 \end{pmatrix} \tag{72}$$

and

$$\mathbf{\Sigma}_e = \sigma_z^2 \left(1 - h^2\right) \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & 1 \end{pmatrix}. \tag{73}$$

In the matrix $\mathbf{\Sigma}_g$, the off-diagonal elements are $\frac{1}{2}\sigma_z^2 h^2$ due to the covariance between sibs, as in Section 2. We assume no covariance between the environmental components or between them and the genetic components. Thus, in total, $(z_1, \dots, z_n) \sim \mathrm{MVN}\left(\mathbf{0}_n, \mathbf{\Sigma}_z\right)$, where

$$\mathbf{\Sigma}_z = \mathbf{\Sigma}_g + \mathbf{\Sigma}_e = \sigma_z^2 \begin{pmatrix} 1 & \frac{h^2}{2} & \dots & \frac{h^2}{2} \\ \frac{h^2}{2} & 1 & \dots & \frac{h^2}{2} \\ \dots & \dots & \dots & \dots \\ \frac{h^2}{2} & \frac{h^2}{2} & \dots & 1 \end{pmatrix}. \tag{74}$$

As we showed in Section 3, because the covariance terms are all equal, the mean of the maximum of the phenotypes ($\boldsymbol{z}$) is equal to the mean of the maximum of n independent normal variables, each with zero mean and variance $\sigma_z^2(1-h^2/2)$. Denote by $\mathrm{E}\left(R\right)$ the mean of the maximum of n standard normal variables (e.g., as we calculated in Eq. (32)), and denote the maximum phenotype across the sibs as $z_{\mathrm{max,actual}}$. Since the identity of this sib is not known at the time of selection, the phenotype of the selected embryo, $z_{\mathrm{max,selected}}$, may be lower, and we have (using Eq. (66)),

$$\begin{aligned}\mathrm{E}\left(z_{\text{max,actual}}-z_{\text{max,selected}}\right) &= \mathrm{E}\left(z_{\text{max,actual}}\right)-\mathrm{E}\left(z_{\text{max,selected}}\right)\\ &= \sigma_z\sqrt{1-\frac{h^2}{2}}\mathrm{E}\left(R\right)-\mathrm{E}\left(\mathrm{PS}_{\max}\right)\\ &= \sigma_z\sqrt{1-\frac{h^2}{2}}\mathrm{E}\left(R\right)-\sigma_z\frac{r_{\text{ps}}}{\sqrt{2}}\mathrm{E}\left(R\right)\\ &= \sigma_z\mathrm{E}\left(R\right)\left(\sqrt{1-\frac{h^2}{2}}-\sqrt{\frac{r_{\text{ps}}^2}{2}}\right). \qquad (75)\end{aligned}$$

Eq. (75) gives the expected loss in the phenotype due to the imperfect predictive power of the score. To find the expected difference for the nuclear families, we calculated $\mathrm{E}\left(R\right)$ exactly based on numerical integration (Eq. (28)), substituted $h^2=0.8$, $\sigma_z=5.6\text{cm}$, and $r_{\text{ps}}^2=0.27$ (as in the real families data), and obtained $\mathrm{E}\left(z_{\text{max,actual}}-z_{\text{max,selected}}\right)=3.1\text{cm}$, very similar to the observed value.

7.2 The probability of the top-scoring embryo to have the top-ranked trait

When reduced to $n=7$ children per family, we found in the real data that on average, in $\approx 31.5\%$ of the families the child whose PS was ranked first was also ranked first in actual height. To determine the expected probability under our quantitative model, consider again n siblings. Recall that their phenotypes, $\boldsymbol{z}=(z_1,\ldots,z_n)$, are modeled as

$$\boldsymbol{z}=\mathbf{PS}+\boldsymbol{\epsilon}, \qquad (76)$$

as in Eq. (3). The polygenic scores are multivariate normal, as defined above (Eq. (8)). For the error term, we have $\boldsymbol{\epsilon}\sim\mathrm{MVN}\left(\mathbf{0}_n,\boldsymbol{\Sigma}_\epsilon\right)$, and

$$\boldsymbol{\Sigma}_\epsilon=\sigma_z^2\begin{pmatrix}1-r_{\text{ps}}^2 & \frac{h^2-r_{\text{ps}}^2}{2} & \ldots & \frac{h^2-r_{\text{ps}}^2}{2}\\ \frac{h^2-r_{\text{ps}}^2}{2} & 1-r_{\text{ps}}^2 & \ldots & \frac{h^2-r_{\text{ps}}^2}{2}\\ \ldots & \ldots & \ldots & \ldots\\ \frac{h^2-r_{\text{ps}}^2}{2} & \frac{h^2-r_{\text{ps}}^2}{2} & \ldots & 1-r_{\text{ps}}^2\end{pmatrix} \qquad (77)$$

To explain the above equation, each ϵ_i has variance $\sigma_z^2\left(1-r_{\text{ps}}^2\right)$. However, here the ϵ_i's must be correlated because they model not only the environment but also genetic factors not modeled by the PS. The off-diagonal entries in the covariance matrix of the phenotypes $\boldsymbol{z}$ are equal to $\sigma_z^2\frac{h^2}{2}$ from Eq. (74). Assuming independence between $\mathbf{PS}$ and $\boldsymbol{\epsilon}$, these entries are equal to the sum of the off-diagonal entries in the covariance matrix of $\mathbf{PS}$, i.e., $\sigma_z^2\frac{r_{\text{ps}}^2}{2}$ (Eq. (18)), and the off-diagonal entries in the covariance matrix of $\boldsymbol{\epsilon}$. Thus, the latter must be $\sigma_z^2\frac{h^2-r_{\text{ps}}^2}{2}$.

To estimate the probability that the top-scoring child is also the tallest, we simulated values for $\mathbf{PS}$ and $\boldsymbol{\epsilon}$, assuming $n = 7$, $h^2 = 0.8$, and $r^2_{\text{ps}} = 0.27$, and then calculated the phenotypes according to Eq. (76). (The value of σ_z does not change the relative ranks, and can be set to any value.) We found that in $\approx 33.4\%$ of the simulations, the sibling top-ranked for the score (PS) was also top-ranked for the phenotype (z), in a reasonable agreement with the empirical results. An integral expression for this probability can also be derived and solved numerically, giving the same result (not shown). An analytical approximation can be derived based on Eq. (14) in [11].

7.3 The probability that the realized gain is negative

Using the above simulated values of $\mathbf{PS}$ and $\boldsymbol{\epsilon}$, we calculated the proportion of simulations in which the realized gain was negative, i.e., $z_{\max} < \frac{1}{n}\sum_{i=1}^{n} z_i$. The proportion came out as 22.5%, compared to 16.4% in the real data.

8 Multiple traits

8.1 The model

Consider T traits normalized to have zero means and variances $\sigma^2_{z,1}, \ldots, \sigma^2_{z,T}$. Let $\mathbf{PS} = (\text{PS}_{(1)}, ..., \text{PS}_{(T)})^t$ be a column vector of polygenic scores for these traits (for a single individual). ($\boldsymbol{x}^t$ is the transpose of a vector $\boldsymbol{x}$.) We assume that $\mathbf{PS}$ is multivariate normal,

$$\mathbf{PS} \sim \left(\text{PS}_{(1)}, ..., \text{PS}_{(T)}\right)^t \sim N\left(\mathbf{0}_T, \mathbf{\Sigma}^{(T)}\right), \tag{78}$$

with

$$\mathbf{\Sigma}^{(T)} = \begin{pmatrix} \sigma^2_{\text{ps},1} & \sigma^2_{\text{ps},12} & \cdots & \sigma^2_{\text{ps},1T} \\ \sigma^2_{\text{ps},12} & \sigma^2_{\text{ps},2} & \cdots & \sigma^2_{\text{ps},2T} \\ \cdots & \cdots & \cdots & \cdots \\ \sigma^2_{\text{ps},1T} & \sigma^2_{\text{ps},2T} & \cdots & \sigma^2_{\text{ps},T} \end{pmatrix}. \tag{79}$$

We defined $\sigma^2_{\text{ps},i} = \text{Var}\left(\text{PS}_{(i)}\right) = \sigma^2_{z,i} r^2_{\text{ps},i}$ as the variance of the PS of trait i, and $\sigma^2_{\text{ps},ij} = \text{Cov}\left(\text{PS}_{(i)}, \text{PS}_{(j)}\right)$ as the covariance between the scores of pairs of traits, which may be non-zero due to pleiotropy.

The polygenic scores for all T traits and for all n sibling embryos can be represented as an $T \times n$ matrix with a matrix normal distribution [12]:

$$\begin{pmatrix} \text{PS}^1_{(1)} & \cdots & \text{PS}^n_{(1)} \\ \cdots & \cdots & \cdots \\ \text{PS}^1_{(T)} & \cdots & \text{PS}^n_{(T)} \end{pmatrix} \sim \text{MN}_{T,n}\left(\mathbf{0}_{T\times n}, \mathbf{\Sigma}^{(T)}, \mathbf{\Sigma}^{(s)}\right), \tag{80}$$

where $\mathbf{\Sigma}^{(s)}$ is an $n \times n$ matrix representing twice the kinship coefficient between siblings, and is given, similarly to Eq. (20), by:

$$\mathbf{\Sigma}^{(s)} = \begin{pmatrix} 1 & \frac{1}{2} & \dots & \frac{1}{2} \\ \frac{1}{2} & 1 & \dots & \frac{1}{2} \\ \dots & \dots & \dots & \dots \\ \frac{1}{2} & \frac{1}{2} & \dots & 1 \end{pmatrix}. \tag{81}$$

Eq. (80) holds because for two embryos $i \neq j$ and two traits $k \neq l$,

$$\begin{aligned} \mathrm{Var}\left(\mathrm{PS}^i_{(k)}\right) &= \sigma^2_{\mathrm{ps},k}, \\ \mathrm{Cov}\left(\mathrm{PS}^i_{(k)}, \mathrm{PS}^i_{(l)}\right) &= \sigma^2_{\mathrm{ps},kl} \\ \mathrm{Cov}\left(\mathrm{PS}^i_{(k)}, \mathrm{PS}^j_{(k)}\right) &= \frac{1}{2}\sigma^2_{\mathrm{ps},k} \\ \mathrm{Cov}\left(\mathrm{PS}^i_{(k)}, \mathrm{PS}^j_{(l)}\right) &= \frac{1}{2}\sigma^2_{\mathrm{ps},kl}. \end{aligned} \tag{82}$$

The first two equations hold by definition, and the third was shown in Section 2. The proof of the fourth equation (covariance between scores of different traits in two embryos) is analogous to that of the third equation (Section 2), and is thus omitted.

8.2 Selection for a linear combination of traits

Suppose we select for a fixed linear combination of the scores, where the weight of trait i is a real number w_i. Denote $\boldsymbol{w} = (w_1, \dots, w_T)^t$, and define the *combined score* of an individual as $\mathrm{PS}_w = \sum_{i=1}^T w_i \mathrm{PS}_{(i)} = \boldsymbol{w}^t \mathbf{PS}$. PS_w is a scalar, and is a linear transformation of a multivariate normal variable. Thus,

$$\mathrm{PS}_w \sim N(0, \boldsymbol{w}^t \mathbf{\Sigma}^{(T)} \boldsymbol{w}), \tag{83}$$

Denote $\sigma^2_w = \boldsymbol{w}^t \mathbf{\Sigma}^{(T)} \boldsymbol{w}$ (the variance of the combined score of a single individual) and PS^i_w as the combined score of embryo i. Based on the distribution of a linear transformation of a matrix normal variable [12], the combined scores of all n embryos are distributed as

$$\left(\mathrm{PS}^1_w, \dots, \mathrm{PS}^n_w\right) \sim \mathrm{MVN}(\mathbf{0}_n, \sigma^2_w \mathbf{\Sigma}^{(s)}). \tag{84}$$

Define G_w be the gain in the combined score by selecting the embryo with the top combined score, as compared to the average combined score (across embryos),

$$G_w = \max\left(\mathrm{PS}^1_w, \dots, \mathrm{PS}^n_w\right) - \frac{1}{n}\sum_{i=1}^n \mathrm{PS}^i_w. \tag{85}$$

Following the same derivation as for a single trait, we have

$$\mathrm{E}\left(G_w\right) = \mathrm{E}\left(\max\left(\mathrm{PS}^1_w, .., \mathrm{PS}^n_w\right)\right) = \frac{\sigma_w}{\sqrt{2}}\mathrm{E}\left(R\right) \approx 0.77\sigma_w\sqrt{\log n}. \tag{86}$$

As in previous sections, R is the maximum of n independent standard normals. The approximation based on extreme value theory can also be used (Eq. (32)).

8.3 The mean gain per trait

Above, we calculated the mean gain of the combined score PS_w. In practice, we are interested in the gain *per trait*. Denote the combined score of the top-scoring embryo as PS_w^*, and the scores of the individual traits (for the top-scoring embryo) as $\mathrm{PS}_{(1)}^*, \ldots \mathrm{PS}_{(T)}^*$. Denote the gain of trait i as $G_{w,i}$,

$$G_{w,i} = \mathrm{PS}_{(i)}^* - \frac{1}{n}\sum_{j=1}^{n} \mathrm{PS}_{(i)}^j \tag{87}$$

Based on the law of total expectation,

$$\mathrm{E}\left(G_{w,i}\right) = \mathrm{E}\left(\mathrm{PS}_{(i)}^*\right) = \mathrm{E}\left(\mathrm{E}\left(\mathrm{PS}_{(i)}^* | \mathrm{PS}_w^*\right)\right). \tag{88}$$

Let us first compute the inner expectation. As shown in Eq. (83), for a given embryo,

$$\mathrm{PS}_w = \sum_{i=1}^{T} w_i \mathrm{PS}_{(i)} \sim N(0, \sigma_w^2). \tag{89}$$

The joint distribution of $\mathrm{PS}_{(i)}, \mathrm{PS}_w$ is bivariate normal with zero means and covariance

$$\mathrm{Cov}\left(\mathrm{PS}_{(i)}, \mathrm{PS}_w\right) = \mathrm{Cov}\left(\mathrm{PS}_{(i)}, \sum_{j=1}^{T} w_i \mathrm{PS}_{(j)}\right) = w_i \sigma_{\mathrm{ps},i}^2 + \sum_{j=1, j\neq i}^{T} w_j \sigma_{\mathrm{ps},ij}^2. \tag{90}$$

The conditional mean for a bivariate normal variable is

$$\begin{aligned}\mathrm{E}\left(\mathrm{PS}_{(i)} | \mathrm{PS}_w\right) &= \mathrm{PS}_w \frac{\mathrm{Cov}\left(\mathrm{PS}_{(i)}, \mathrm{PS}_w\right)}{\mathrm{Var}\left(\mathrm{PS}_w\right)} \\ &= \mathrm{PS}_w \frac{w_i \sigma_{\mathrm{ps},i}^2 + \sum_{j=1, j\neq i}^{T} w_j \sigma_{\mathrm{ps},ij}^2}{\sigma_w^2},\end{aligned} \tag{91}$$

which gives the inner expectation. Using Eq. (86), we can compute the outer expectation,

$$\begin{aligned}
\mathrm{E}\left(G_{w,i}\right) &= \mathrm{E}\left(\mathrm{PS}^*_{(i)}\right) = \mathrm{E}\left(\mathrm{E}\left(\mathrm{PS}^*_{(i)} | \mathrm{PS}^*_w\right)\right) \\
&= \mathrm{E}\left(\mathrm{PS}^*_w \frac{w_i\sigma^2_{\mathrm{ps},i} + \sum_{j=1,j\neq i}^T w_j\sigma^2_{\mathrm{ps},ij}}{\sigma^2_w}\right) \\
&= \mathrm{E}\left(\mathrm{PS}^*_w\right) \frac{w_i\sigma^2_{\mathrm{ps},i} + \sum_{j=1,j\neq i}^T w_j\sigma^2_{\mathrm{ps},ij}}{\sigma^2_w} \\
&= \frac{\sigma_w}{\sqrt{2}} \mathrm{E}\left(R\right) \frac{w_i\sigma^2_{\mathrm{ps},i} + \sum_{j=1,j\neq i}^T w_j\sigma^2_{\mathrm{ps},ij}}{\sigma^2_w} \\
&= \mathrm{E}\left(R\right) \frac{w_i\sigma^2_{\mathrm{ps},i} + \sum_{j=1,j\neq i}^T w_j\sigma^2_{\mathrm{ps},ij}}{\sqrt{2}\sigma_w} \\
&= \mathrm{E}\left(R\right) \frac{w_i\sigma^2_{\mathrm{ps},i} + \sum_{j=1,j\neq i}^T w_j\sigma^2_{\mathrm{ps},ij}}{\sqrt{2}\sqrt{\sum_{j=1}^T w_j^2\sigma^2_{\mathrm{ps},j} + \sum_{j=1}^T\sum_{k=1,k\neq j}^T w_j w_k \sigma^2_{\mathrm{ps},jk}}}. \qquad (92)
\end{aligned}$$

8.4 Specific cases

More insight can be gained under *natural* weights, $w_i = \sigma^{-1}_{\mathrm{ps},i}$, under which all traits are equally weighted after scaling by the standard deviations of their scores. Substituting in Eq. (92), and using the correlation coefficient $\rho_{\mathrm{ps},ij} = \frac{\sigma^2_{\mathrm{ps},ij}}{\sigma_{\mathrm{ps},i}\sigma_{\mathrm{ps},j}}$, we obtain

$$\begin{aligned}
\mathrm{E}\left(G_{i,w}\right) &= \mathrm{E}\left(R\right) \frac{\sigma_{\mathrm{ps},i} + \sum_{j=1,j\neq i}^T \sigma_{\mathrm{ps},i}\rho_{\mathrm{ps},ij}}{\sqrt{2}\sqrt{T + \sum_{j=1}^T\sum_{k=1,k\neq j}^T \rho_{\mathrm{ps},jk}}} \\
&= \sigma_{\mathrm{ps},i}\mathrm{E}\left(R\right) \frac{1 + \sum_{j=1,j\neq i}^T \rho_{\mathrm{ps},ij}}{\sqrt{2}\sqrt{T + \sum_{j=1}^T\sum_{k=1,k\neq j}^T \rho_{\mathrm{ps},jk}}}. \qquad (93)
\end{aligned}$$

We can gain further insight by assuming all correlations across traits to be equal, $\rho_{\mathrm{ps},ij} = \rho$ for all $i \neq j$. Substituting in Eq. (93), we obtain

$$\begin{aligned}
\mathrm{E}\left(G_{w,i}\right) &= \sigma_{\mathrm{ps},i}\mathrm{E}\left(R\right) \frac{1 + (T-1)\rho}{\sqrt{2}\sqrt{T + T(T-1)\rho}} \\
&= \frac{\sigma_{\mathrm{ps},i}\mathrm{E}\left(R\right)\sqrt{1 + (T-1)\rho}}{\sqrt{2T}}. \qquad (94)
\end{aligned}$$

Let us analyze Eq. (94) in more detail. First, consider the case of no pleiotropy, or $\rho = 0$. In this limiting case, we obtain

$$\mathrm{E}\left(G_{w,i}\right) = \frac{\sigma_{\mathrm{ps},i}\mathrm{E}\left(R\right)}{\sqrt{2T}} = \frac{\mathrm{E}\left(G\right)}{\sqrt{T}}. \qquad (95)$$

In other words, when selecting for the sum of T independent traits, the gain per trait decreases by a factor $\sqrt{T}$ compared to selecting for each individual trait. When $\rho = 1$, we obtain

$$\mathrm{E}\left(G_{w,i}\right) = \frac{\sigma_{\mathrm{ps},i}\mathrm{E}\left(R\right)\sqrt{T}}{\sqrt{2T}} = \mathrm{E}\left(G\right). \tag{96}$$

That is, for $\rho = 1$, all traits are equal (after scaling), and selecting for one trait is equivalent to selecting for all others. The most negative value of ρ possible (allowing a valid joint distribution of the scores per trait) is $\rho = -1/(T-1)$. This corresponds to all traits being maximally inversely correlated with one another. In that case, Eq. (94) gives $\mathrm{E}\left(G_{i,w}\right) = 0$, and selection based on the sum of the scores does not lead to any gain in any individual trait.

Finally, assume we wish to retain the natural scaling of the weights, but assign unequal importance to each trait. We can use $w_i = \lambda_i \sigma_{\mathrm{ps},i}^{-1}$. (We can also impose $\sum_{i=1}^{T} \lambda_i = 1$ or $\sum_{i=1}^{T} |\lambda_i| = 1$ for an easier interpretation; however, only the relative value of each λ_i matters.) From Eq. (92), we have

$$\mathrm{E}\left(G_{w,i}\right) = \sigma_{\mathrm{ps},i}\mathrm{E}\left(R\right)\frac{\lambda_i + \sum_{j=1,j\neq i}^{T} \lambda_j \rho_{\mathrm{ps},ij}}{\sqrt{2}\sqrt{\sum_{j=1}^{T} \lambda_j^2 + \sum_{j=1}^{T}\sum_{k=1,k\neq j}^{T} \lambda_j \lambda_k \rho_{\mathrm{ps},jk}}}. \tag{97}$$

For the case of two traits analyzed in the main text, the coefficients can be written as $\lambda_1 = \lambda$ and $\lambda_2 = 1 - \lambda$. This gives (denoting $\rho \equiv \rho_{\mathrm{ps},12}$)

$$\begin{aligned}
\mathrm{E}\left(G_{w,1}\right) &= \sigma_{\mathrm{ps},1}\mathrm{E}\left(R\right)\frac{\lambda + (1-\lambda)\rho}{\sqrt{2}\sqrt{\lambda^2 + (1-\lambda)^2 + 2\lambda(1-\lambda)\rho}},\\
\mathrm{E}\left(G_{w,2}\right) &= \sigma_{\mathrm{ps},2}\mathrm{E}\left(R\right)\frac{1 - \lambda + \lambda\rho}{\sqrt{2}\sqrt{\lambda^2 + (1-\lambda)^2 + 2\lambda(1-\lambda)\rho}}.
\end{aligned} \tag{98}$$

For the case of height and BMI, we select for smaller values of the second trait by setting $0 < \lambda < 1$ and $\lambda_2 = -(1-\lambda)$. This gives

$$\begin{aligned}
\mathrm{E}\left(G_{w,\mathrm{height}}\right) &= \sigma_{\mathrm{ps,height}}\mathrm{E}\left(R\right)\frac{\lambda - (1-\lambda)\rho}{\sqrt{2}\sqrt{\lambda^2 + (1-\lambda)^2 - 2\lambda(1-\lambda)\rho}},\\
\mathrm{E}\left(G_{w,\mathrm{BMI}}\right) &= \sigma_{\mathrm{ps,BMI}}\mathrm{E}\left(R\right)\frac{-(1-\lambda) + \lambda\rho}{\sqrt{2}\sqrt{\lambda^2 + (1-\lambda)^2 - 2\lambda(1-\lambda)\rho}}.
\end{aligned} \tag{99}$$

For height and BMI, $\rho < 0$. Thus, $\mathrm{E}\left(G_{w,\mathrm{height}}\right) > 0$, while $\mathrm{E}\left(G_{w,\mathrm{BMI}}\right) < 0$.

9 Code availability

R code that implements some of the calculations in this document can be found at `https://github.com/orzuk/EmbryoSelectionCalculator`.

References

[1] W. J. Peyrot, M. R. Robinson, B. W. Penninx, and N. R. Wray. Exploring boundaries for the genetic consequences of assortative mating for psychiatric traits. *JAMA Psychiatry*, 73:1189, 2016.

[2] M. R. Robinson, A. Kleinman, M. Graff, A. A. E. Vinkhuyzen, D. Couper, M. B. Miller, W. J. Peyrot, A. Abdellaoui, B. P. Zietsch, I. M. Nolte, J. V. van Vliet-Ostaptchouk, H. Snieder, The LifeLines Cohort Study, Genetic Investigation of Anthropometric Traits (GIANT) consortium, S. E. Medland, N. G. Martin, P. K. E. Magnusson, W. G. Iacono, M. McGue, K. E. North, J. Yang, and P. M. Visscher. Genetic evidence of assortative mating in humans. *Nat Hum Behav*, 1:0016, 2017.

[3] L. Yengo, M. R. Robinson, M. C. Keller, K. E. Kemper, Y. Yang, M. Trzaskowski, J. Gratten, P. Turley, D. Cesarini, D. J. Benjamin, N. R. Wray, J. Yang, M. E. Goddard, and P. M. Visscher. Imprint of assortative mating on the human genome. *Nat Hum Behav*, 2:948–954, 2018.

[4] T. Berisa and J. K. Pickrell. Approximately independent linkage disequilibrium blocks in human populations. *Bioinformatics*, 32:283, 2016.

[5] P. M. Visscher, N. R. Wray, Q. Zhang, P. Sklar, M. I. McCarthy, M. A. Brown, and J. Yang. 10 years of GWAS discovery: Biology, function, and translation. *Am J Hum Genet*, 101:5, 2017.

[6] M. Lynch and B. Walsh. *Genetics and Analysis of Quantitative Traits*. Sinauer Associates, 1998.

[7] G. A. T. McVean, S. R. Myers, S. Hunt, P. Deloukas, D. R. Bentley, and P. Donnelly. The fine-scale structure of recombination rate variation in the human genome. *Science*, 304:581, 2004.

[8] H. A. David and H. N. Nagaraja. *Order Statistics*. Wiley-Interscience, third edition, 2003.

[9] L. de Haan. Sample extremes: an elementary introduction. *Statistica Neerlandica*, 30:161, 1976.

[10] P. M. Visscher, B. McEvoy, and J. Yang. From Galton to GWAS: quantitative genetics of human height. *Genet Res (Camb)*, 92:371, 2010.

[11] O. Zuk, L. Ein-Dor, and E. Domany. Ranking under uncertainty. In *Uncertainty in Artificial Intelligence*, pages 466–473, 2007.

[12] A. K. Gupta and D. K. Nagar. *Matrix variate distributions*. Chapman and Hall/CRC, 1999.