

Prioritizing embryos with lower homozygosity is expected to reduce disease risk in children of related individuals undergoing preimplantation genetic testing

Tobias Wolfram^{1,*}, Mohammad Ahangari^{1,*}, Ivan Davidson¹, Laura Wartschinski¹, Jeremiah H. Li¹, Mathew Eyre¹, David Stern¹, Justin Schleede¹, Alireza Haghighi^{2,3,4}, Shai Carmi⁵, Michael Christensen¹

¹Herasight Research, USA

²Harvard Medical School, Boston, MA, USA

³Department of Medicine, Brigham and Women's Hospital, Boston, MA, USA

⁴The Broad Institute, Cambridge, MA, USA

⁵Braun School of Public Health and Community Medicine, The Hebrew University of Jerusalem, Israel

*These authors contributed equally.

Abstract

Consanguinity is a reproductive union between individuals who share a recent common ancestor. These unions are common in many regions of the world and increase the burden of rare recessive disorders by elevating autozygosity in offspring. Current reproductive genetic screening focuses on a limited set of known pathogenic variants, leaving most recessive risk unaddressed. Here we argue that embryo-level autozygosity, quantified as the fraction of the genome in long runs of homozygosity (F_{ROH}), is a potentially actionable genomic biomarker that can be integrated into routine preimplantation genetic testing as a homozygosity-informed embryo-prioritization framework (PGT-H) that can be layered onto existing embryo biopsy workflows when couples are already undergoing IVF with PGT-A or PGT-M. Using forward simulations of first-cousin and double-first-cousin couples, we show that siblings conceived by the same couple span a wide range of F_{ROH} ; selecting the lowest- F_{ROH} candidate from a cohort of five embryos reduces F_{ROH} by approximately 40% on average. Combining these reductions with empirical effect-size estimates, we estimate that for first-cousin couples this strategy could reduce risk of intellectual disability by roughly 35–45% (corresponding to an absolute risk reduction of about 1.8–2.2%) and potentially reduce excess recessive disease burden, while also modestly reducing risk of common diseases such as type 2 diabetes. We outline how existing PGT-A and PGT-M workflows could potentially be extended to report embryo-level F_{ROH} and discuss ethical and counseling considerations. Autozygosity-based embryo prioritization offers a principled way to address a component of recessive risk that current variant-centric approaches miss.

Keywords: autozygosity; runs of homozygosity (ROH); consanguinity; PGT-A; PGT-M; embryo selection; recessive disorders; congenital anomalies; neurodevelopmental disorders.

Introduction

Consanguinity is a reproductive union between individuals who share a recent common ancestor (Figure 1A). Consanguineous marriage is common in many parts of the world (Figure 1B) and can confer social and economic benefits¹. Accordingly, reproductive counseling in this context should remain culturally sensitive and non-directive. However, it is well established that offspring of related parents are at elevated risk for various poor health outcomes²⁻⁷. When parents share recent ancestors, they may transmit long genomic segments that are identical by descent. This results in elevated genome-wide autozygosity in their children, that is, genomic regions inherited identically by descent from both parents. Elevated autozygosity increases the chance that rare recessive variants are homozygous in the offspring, thereby conferring elevated disease risk. This mechanism underlies the long-recognized excess of recessive disease in consanguineous families.

Autozygosity can be measured from runs of homozygosity (ROH), which represent long homozygous genomic segments inherited identically by descent from both parents (Figure 2). In outbred populations, the fraction of the genome in ROH, or F_{ROH} , has values usually below 1%, whereas offspring of first-cousin unions have an average ROH of 6.25%. Offspring of double first-cousin unions have an average ROH of 12.5%. Individual F_{ROH} values, however, can vary substantially even among full siblings due to Mendelian transmission and the stochastic nature of meiotic recombination. This within-family variation creates a clinical opportunity to preferentially select embryos with lower autozygosity even when parents have a fixed degree of relatedness.

For reproductive medicine clinicians, the practical question is how to reduce the genetic risks (including both Mendelian and complex polygenic architectures) posed by elevated autozygosity in consanguineous couples. Long-standing ethical analyses stress that it is inconsistent with genetic counseling principles to discourage consanguineous marriage^{8,9}; rather, the emphasis should be on communicating evidence-based facts for informed decision-making and providing access to medical resources such as genetic testing, when applicable. In this article, we synthesize clinically relevant evidence, explain why offspring F_{ROH} varies widely within families and describe the utility of prioritizing embryos with lower realized autozygosity through preimplantation genetic testing for homozygosity (PGT-H), as well as how this framework can be integrated into standard IVF workflows. We emphasize what is necessary for responsible clinical adoption and realistic counseling about risk reduction rather than risk elimination.

Elevated autozygosity is associated with increased risk of rare disorders and common disease-related outcomes

The best-established consequence of elevated autozygosity is an increased burden of autosomal recessive disease. This has been demonstrated through population-genetic modeling, autozygosity studies, and sequencing-based analyses¹⁰, and several large-scale studies in populations with high rates of consanguinity have shown that a meaningful fraction of pediatric severe intellectual disability, multiple congenital anomalies, and other rare disorders can be attributed to autosomal recessive variants that are homozygous in regions of ROH¹¹⁻¹⁴.

Beyond classic Mendelian conditions, genome-wide autozygosity has been associated with a broad range of common traits and diseases. Analyses in diverse cohorts have linked higher F_{ROH} to increased risk of cardiometabolic disease, lower height, reduced cognitive performance, ad-

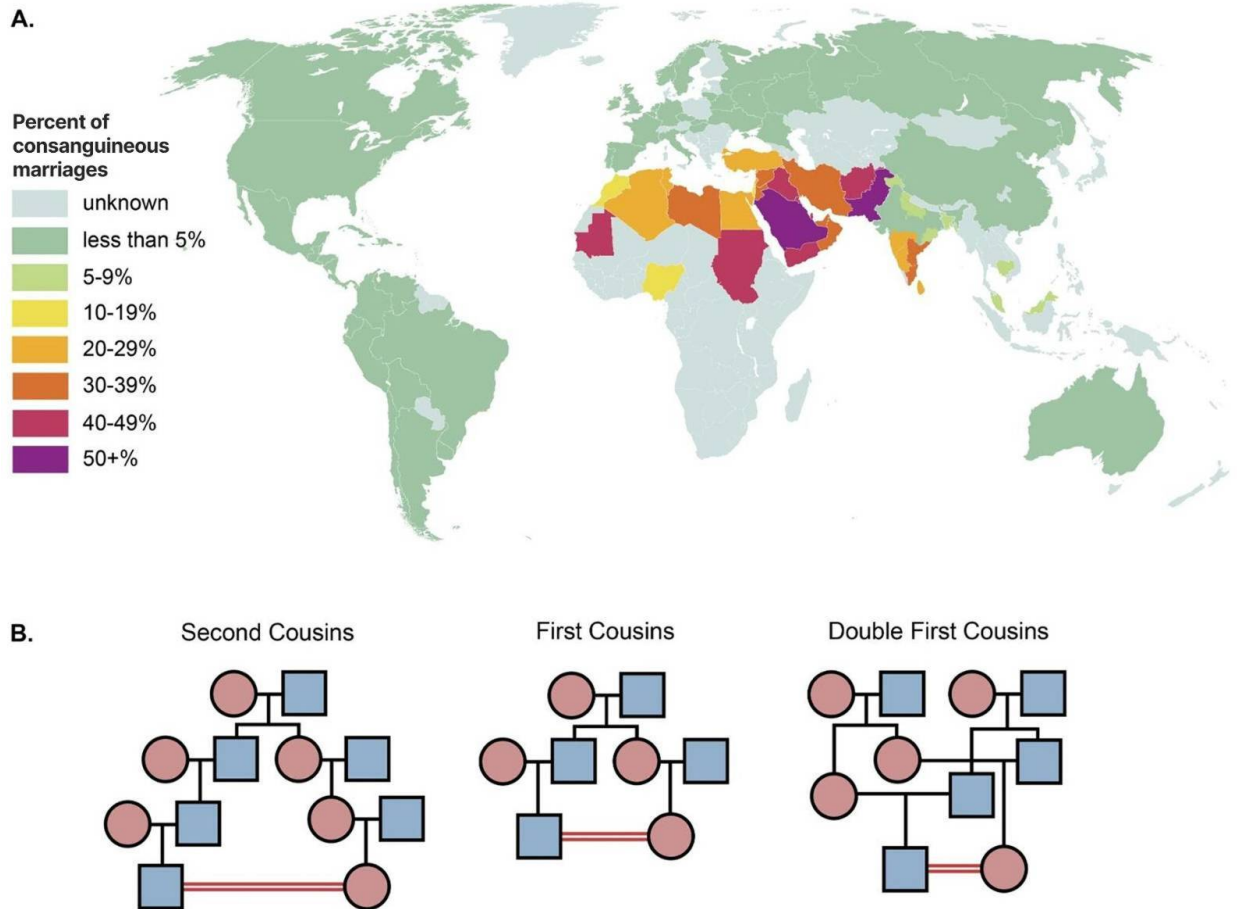


Figure 1: Consanguinity and its global distribution.

(A) Estimated proportion of marriages between second cousins or closer by country, adapted from Hamamy et al.⁹

(B) Pedigrees illustrating second-cousin, first-cousin, and double-first-cousin unions (red double lines indicate the consanguineous partnership). These relationships yield progressively higher expected offspring autozygosity (expected $F_{ROH} = 1.5625\%$, 6.25% , and 12.5% , respectively, in the absence of background homozygosity).

verse reproductive outcomes, higher BMI, increased risk of type 2 diabetes, and reduced fertility or increased risk of childlessness²⁻⁵. These associations are typically modest at the individual level but can translate into substantial population-level burdens in areas with high consanguinity.

Epidemiologic studies of consanguineous unions provide convergent evidence. Relative to offspring of unrelated parents, children of first-cousin or more closely related couples have increased risk of stillbirth, infant death, congenital anomalies, and developmental delay^{7,15-17}. In a Norwegian registry study spanning over two decades, consanguinity was associated with elevated rates of stillbirth and infant mortality even after adjusting for maternal education and other covariates¹⁵. In the Born in Bradford cohort, children of first-cousin parents had elevated risks of mortality, higher healthcare utilization, and increased rates of speech, language and learning difficulties, with corresponding impacts on educational outcomes¹⁶.

Taken together, these findings demonstrate that elevated autozygosity is a biologically interpretable and clinically meaningful risk factor for a mixture of monogenic and polygenic outcomes. Importantly, F_{ROH} captures the genome-wide burden of recessive variants, including

those at loci not yet known to be associated with disease. This motivates embryo selection strategies that directly target autozygosity rather than only known variant-level risk.

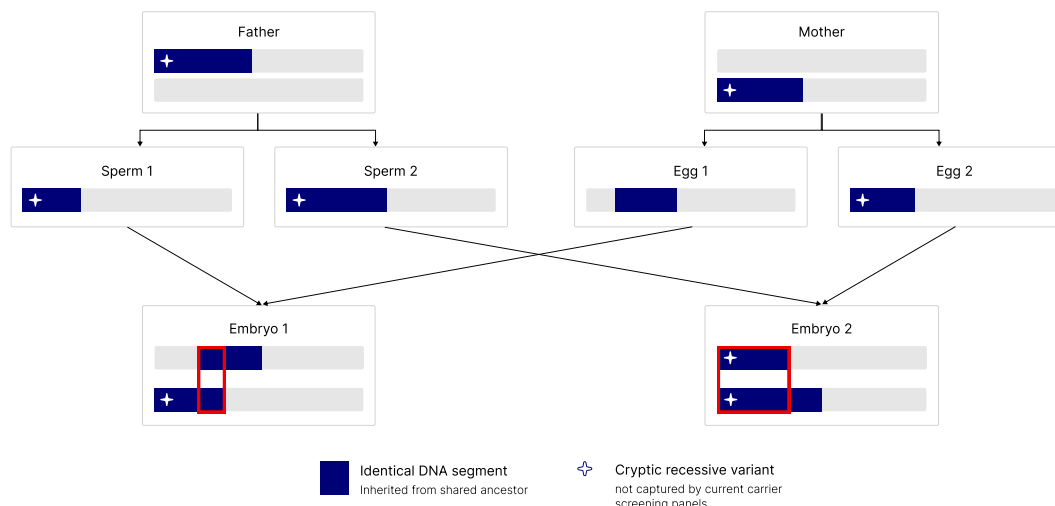


Figure 2: Autozygosity and health outcomes.

Schematic showing how having shared recent ancestors generate long identical-by-descent segments (runs of homozygosity, ROH) in offspring; recombination and meiotic division produces embryo-to-embryo variation in the location and extent of ROH and can render recessive variants homozygous (crosses), including previously unidentified or unscreened variants not currently captured by standard carrier panels.

Large within-family variation in autozygosity among embryos enables substantial disease risk reduction

Although parental relatedness determines the expected level of autozygosity in offspring, the realized F_{ROH} for any given child depends on the stochastic details of recombination and Mendelian segregation. Even among full siblings from the same consanguineous couple, F_{ROH} can vary widely. Embryo screening can leverage this within-family variation by prioritizing the embryos with lowest F_{ROH} .

To quantify this variation, we performed forward simulations of consanguineous pedigrees using realistic sex-specific recombination maps¹⁸. For each relationship type (first cousins and double first cousins), we simulated 1,000 independent families. For each simulated couple, we generated 1,000 embryos by sampling sex-specific recombination events along the parental haplotypes and tracking the inheritance of founder chromosomes through the pedigree. ROH segments were defined as autosomal tracts where the two homologous chromosomes were inherited from the same founder haplotype, and F_{ROH} was computed as total autosomal ROH length divided by total autosomal genome length.

Within each family, we characterized the distribution of embryo F_{ROH} by computing the family mean, the standard deviation, and the range across 1,000 embryos. As expected, mean F_{ROH} clustered near the theoretical values of 0.0625 for first-cousin unions and 0.125 for double-first-cousin unions, but individual embryos spanned a broad range around these means. Figure 3A illustrates the distribution of mean F_{ROH} across simulated first-cousin families, and Figure 3B shows the distribution across embryos for a representative family, highlighting substantial

within-family variance.

We then asked how much autozygosity reduction could be achieved by selecting the embryo with lowest F_{ROH} from clinically realistic cohort sizes. For each family and each embryo count ranging from 2 to 20, we repeatedly sampled a given number of embryos without replacement from the pool of 1,000 embryos and recorded the minimum, mean, and maximum F_{ROH} across 10,000 replicates. Importantly, our simulations assumed no background F_{ROH} in the population, which may underestimate the absolute reduction achievable, especially in populations with elevated background autozygosity. Because background autozygosity may increase within-family variance in some populations, our simulations may underestimate absolute selection benefit in those settings; however, the magnitude of this effect requires population-specific validation. To corroborate the simulation results, we also developed a mathematical model to approximate the reduction in F_{ROH} as a function of the number of embryos (see Supplement).

For first-cousin couples, selecting the lowest- F_{ROH} embryo from five embryos reduced realized F_{ROH} by an average of roughly 40% relative to a randomly selected embryo, with larger gains as the number of embryos increased (Figure 3C, Figure 3D, Supplementary Figure 1). For double-first-cousin couples, absolute F_{ROH} reductions were higher but the proportional reductions from embryo selection were smaller (average reduction of 3.8 F_{ROH} % points, meaning a relative reduction of approximately 30%).

Finally, comparing expected F_{ROH} with and without screening demonstrated consistent improvement across all families, with larger embryo cohorts yielding lower F_{ROH} (Figure 3E).

We emphasize that the above analysis assumes that all considered embryos have birth potential. Thus, the relevant parameter is the number of potential *births* per IVF cycle, rather than the number of embryos¹⁹.

Translating reductions in autozygosity into expected health benefits

Translating reductions in autozygosity into the full extent of health-related benefits is challenging because autozygosity has been associated with dozens of diseases and traits, and the extent to which each of these associations is biologically causal remains unclear in the absence of within-sibling replications. Rather than attempting to model all such outcomes, we focus on a small set of well-established and high-impact outcomes with evidence supporting plausible recessive contributions: infant mortality, intellectual disability, congenital anomalies, type 2 diabetes, and childlessness. The first three are likely driven predominantly by rare, large-effect recessive variants, whereas the latter two likely reflect diffuse recessive contributions from many loci.

For outcomes where the increased risk is likely due to single large-effect recessive variants, we assume that the excess risk above the outbred baseline scales approximately linearly with F_{ROH} . For an individual with $F_{\text{ROH}} = F$ and a deleterious recessive variant of frequency p , the probability of homozygosity is approximately pF for small p , so doubling F roughly doubles the contribution of such variants to disease risk. In a two-step screening framework, PGT-M first removes a fraction of that risk by targeting known pathogenic variants. For common, polygenic outcomes such as type 2 diabetes and childlessness, we instead model the log-odds of disease as proportional to F_{ROH} , as has been empirically observed in within-family analyses that help minimize environmental confounding³⁻⁵. In both cases, the key intuition is that lowering F_{ROH} reduces the chance that recessive-acting alleles are homozygous across the genome, so that each percentage-point reduction in autozygous genome content translates into a proportional reduc-

Expected FROH Reduction from Embryo Screening
(First Cousin Unions)

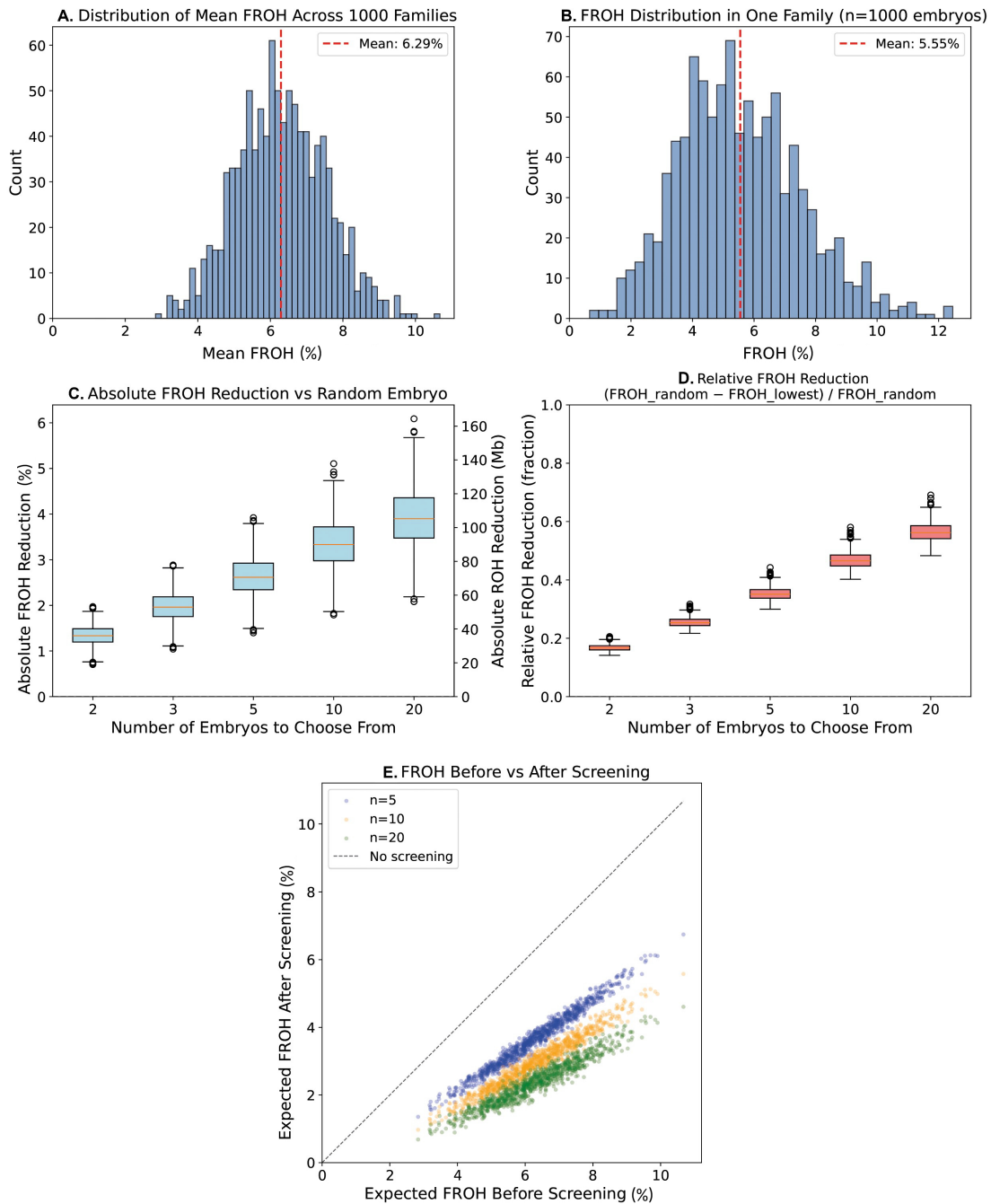


Figure 3: Simulated expected reduction in embryo F_{ROH} from screening in first-cousin unions. (A) Distribution of mean F_{ROH} per family across 1,000 simulated first-cousin families. (B) Distribution of F_{ROH} among 1,000 embryos from a representative family. (C) Expected absolute reduction in F_{ROH} comparing random embryo to lowest- F_{ROH} embryo from sets of $n = 2, 3, 5, 10, 20$ embryos. (D) Same as (C) showing relative risk reduction. (E) Expected embryo F_{ROH} with versus without screening. Each point represents a simulated family. For each family, the expected F_{ROH} is calculated as the mean embryo F_{ROH} across 1,000 simulations when embryos are selected at random versus when embryos are prioritized by F_{ROH} from a cohort of n embryos.

tion in excess recessive disease risk.

We combined published estimates of baseline risks in outbred populations with relative risks as a function of parental relatedness (via F_{ROH}) to estimate absolute risks for first-cousin and double-first-cousin couples with and without embryo screening (see Methods for details). For intellectual disability and congenital anomalies, we further incorporated the expected diagnostic yield of exome-based testing to model the fraction of recessive risk that could be removed by variant-level PGT-M alone,^{12,20} versus the additional reduction achieved by autozygosity-based embryo selection. We note that this likely provides an overly optimistic estimate for the effectiveness of PGT-M because PGT-M is typically restricted to pathogenic or likely pathogenic variants, whereas clinical diagnoses in affected individuals include variants that would not meet this classification threshold. For intellectual disability and congenital anomalies, we assume the baseline attributable fraction from recessive-acting variants is 4%, since empirical estimates suggest biallelic variation accounts for around 4% of neurodevelopmental disorders in outbred populations²¹, and assume the same value for congenital conditions since no precise estimates of the attributable fraction are known to our knowledge. Although we also modeled infant death as having a primarily monogenic genetic architecture, we could not identify sufficiently relevant estimates for the expected PGT-M risk reduction. Finally, we attribute the excess risk associated with increased F_{ROH} in these conditions entirely to the elevated burden of monogenic-acting homozygous recessive genotypes.

These estimates depend on simplifying assumptions regarding the linearity of recessive burden, diagnostic yield, and the relationship between F_{ROH} and disease risk. Table 1 summarizes the resulting expected risk reductions for the two relationship types. For double first-cousin couples, the combination of PGT-M and F_{ROH} -informed embryo screening removes about half of the excess risk for the primarily Mendelian outcomes we considered. For intellectual disability in double first-cousin couples, the absolute risk decreases from 8% to 5–6.2% with PGT-M alone (assuming 30–50% diagnostic yield^{13,20,22}) and to 4.1–4.9% with the combined strategy.

We estimated intellectual disability risk of about 5% without screening for first-cousin couples, compared with around 2% in the outbred population^{12,13}. Under the same assumptions, PGT-M alone reduced this to roughly 3.5–4.1%, and PGT-M plus F_{ROH} -based embryo screening reduced it further to 2.8–3.2% when five embryos remain after PGT-M. Infant death risk is reduced from approximately 1.3% to 1.0% when embryos are prioritized on F_{ROH} , and type 2 diabetes risk decreases from about 28% to 24.5%. These estimates illustrate that PGT-M and F_{ROH} -based embryo screening act in a complementary fashion: PGT-M removes risk mediated by known variants, whereas F_{ROH} -based prioritization reduces the remaining genome-wide recessive burden in proportion to the achievable reduction in embryo-level autozygosity. It should be emphasized that these risk reductions are achieved simultaneously for all of these and other conditions. This differs from polygenic embryo screening approaches, where prioritizing embryos specifically for low risk for one condition may not reduce risk for other genetically distinct conditions.

Outcome	Assumed baseline risk in US outbred population	Risk ratio for offspring with specified parental relatedness relative to $F_{ROH} = 0$	Average absolute risk for offspring with specified parental relatedness without embryo screening (product of previous two columns)	Average absolute risk for offspring with specified parental relatedness with PGT-M	Average absolute risk for offspring with specified parental relatedness with PGT-M + autozygosity embryo screening ($n = 5$ remaining embryos)
Double first cousin; $F_{ROH} = 0.125$					
Infant death	~ 0.6% ²³	~ 3.4 ¹⁵	~ 2.0%	—	~ 1.6%
Intellectual disability*	~ 2% ²²	~ 4.0 ¹²	~ 8%	~ 5–6.2%	~ 4.1–4.9%
Congenital anomalies [†]	~ 3% ²⁴	~ 3.4 ¹⁷	~ 10.2%	~ 6.5–8.3%	~ 5.5–6.7%
Type 2 diabetes	~ 20% ²⁵	~ 1.9 ⁵	~ 38%	—	~ 32%
Childlessness	~ 20% ^{26,27}	~ 2.6 ³	~ 51%	—	~ 41%
First cousin; $F_{ROH} = 0.0625$					
Infant death	~ 0.6%	~ 2.2	~ 1.32%	—	~ 1.02%
Intellectual disability*	~ 2%	~ 2.5	~ 5%	~ 3.5–4.1%	~ 2.8–3.2%
Congenital anomalies [†]	~ 3%	~ 2.2	~ 6.6%	~ 4.7–5.6%	~ 4.0–4.5%
Type 2 diabetes	~ 20%	~ 1.4	~ 28%	—	~ 24.5%
Childlessness	~ 20%	~ 1.7	~ 34%	—	~ 28%

Table 1: Expected risk reductions after screening for autozygosity.

For infant death, congenital anomalies, and intellectual disability, where the increased risk is likely due to large-effect recessive variants, we assume the risk ratio varies linearly with F_{ROH} . For type 2 diabetes and childlessness, we assume log-odds vary linearly with F_{ROH} . See Methods for details. *For intellectual disability rows, PGT-M values assume a 30–50% diagnostic yield^{13,20,22}. †For congenital anomaly rows, PGT-M values assume a 26–50% diagnostic yield^{20,28}

Clinical implementation and counseling considerations

Consanguinity arises in diverse cultural, religious and socio-economic contexts, and a consanguineous marriage often provides social and economic benefits^{1,9}. Ethical analyses and professional guidelines emphasize that discouraging consanguineous marriage per se is inconsistent with core principles of genetic counseling, which prioritize autonomy, non-directiveness, and respect for cultural values^{8,29–33}. Instead, the focus should be on communicating evidence-based information about risks and options, and on ensuring equitable access to testing and reproductive technologies. Ethical arguments developed around polygenic embryo screening provide a useful parallel framework for thinking through disclosure, counseling, and permissible uses of embryo autozygosity screening³⁴.

Within this framework, F_{ROH} -based embryo prioritization must be presented as a risk-reduction tool, rather than a guarantee against disease. Counseling should avoid framing PGT-H as a recommendation against consanguinity and should instead present it as an optional risk-reduction tool within a non-directive reproductive counseling framework. Counseling should clearly distinguish between: (i) risks from known pathogenic variants that can be targeted by PGT-M; (ii) the genome-wide, largely uncharacterized recessive risk reflected in F_{ROH} ; and (iii) other genetic and non-genetic contributors to adverse outcomes. Counseling should also emphasize that autozygosity-informed selection typically reduces relative risk by tens of percent, but leaves a substantial residual risk due to remaining ROH, non-recessive etiologies, and unscreened genetic mechanisms. It should also be emphasized that the reduction in ROH can vary substantially across families and embryo cohorts for a given family. Finally, it is important to communicate that risk reductions are likely to be smaller than what may be estimated for the given number

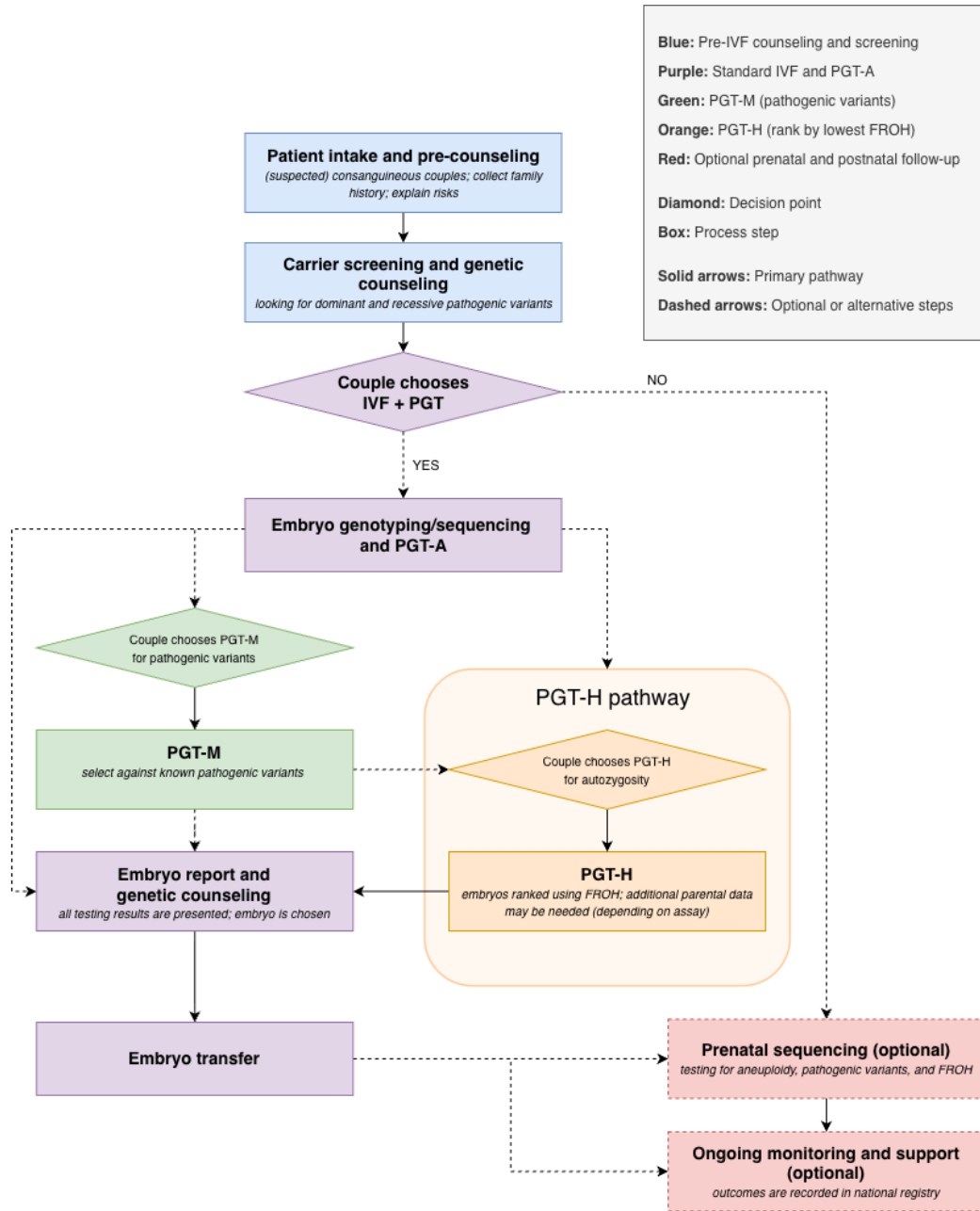


Figure 4: Workflow for integrating F_{ROH} into routine PGT-A / PGT-M.

Parental sequencing or genotyping enables construction of chromosome-scale haplotypes. Ultra-low-pass sequencing of embryo biopsies is already performed for PGT-A, from which embryo-level F_{ROH} can be computed and reported alongside aneuploidy and monogenic findings. The workflow illustrates how F_{ROH} -informed embryo prioritization (PGT-H) can be layered onto existing pre-IVF counseling and screening, carrier screening and genetic counseling, standard IVF with PGT-A and/or PGT-M, embryo transfer, and optional prenatal sequencing and follow-up for couples already undergoing IVF with embryo biopsy.

of embryos, as not all embryos will result in a live birth, and some embryos may be carriers of monogenic pathogenic variants. More formal counseling aids that account for embryo attrition and cohort size can help translate embryo-level ranking into realistic expectations of benefit at the live-birth level³⁵.

In practice, preconception or early-pregnancy genetic counseling for consanguineous couples should include discussion of the couple's genealogical relatedness, expected F_{ROH} range in offspring, and how PGT-A, PGT-M, and PGT-H intersect. For couples already undergoing IVF with PGT-M or PGT-A, embryo genotyping or sequencing data generated in standard workflows can support estimation of embryo-level F_{ROH} , particularly when combined with parental genotype data or haplotype information (Figure 4). Recent work has shown that embryo genomes can be reconstructed from ultra-low-pass PGT-A sequencing using parental genomic data, supporting the feasibility of deriving additional embryo-level genomic measures from existing PGT data³⁶.

When parental genotype data are available, as is typical when carrier screening is performed, ROH segments identified in embryos can be interpreted in light of the stated pedigree and parental haplotype structure, helping distinguish expected consanguinity-derived autozygosity from atypical homozygosity patterns.¹

In cycles yielding multiple euploid embryos, counseling may need to aid couples' interpretation and decision-making when multiple distinct metrics are presented, such as expected F_{ROH} reduction and relative and absolute risk reduction for early and late onset conditions.

Importantly, expanding access to genomic testing and PGT must not exacerbate health inequities. Consanguinity is more prevalent in populations that already face structural barriers in healthcare and research representation^{9,16}. Ensuring fair access to PGT-H will require efforts in insurance coverage, culturally sensitive outreach, and development of infrastructure in regions where consanguinity is common.

Conclusion

Genome-wide autozygosity is a robust, biologically interpretable risk factor that links the long-known epidemiology of consanguinity with modern genomic data. Our synthesis of the literature supports the conclusion that elevated F_{ROH} increases risk for a spectrum of rare and common outcomes, and our simulations show that siblings conceived by the same highly consanguineous couple can differ markedly in F_{ROH} . Selecting embryos with lower autozygosity therefore offers a principled way to reduce recessive disease burden beyond what is achievable by variant-centric PGT-M alone.

F_{ROH} -based embryo prioritization can be implemented with data already generated in routine PGT-A and PGT-M workflows. We estimate that for first-cousin and double-first-cousin couples, PGT-H in combination with PGT-M can reduce absolute risk of severe outcomes such as intellectual disability and major congenital anomalies by several percentage points, and reduce excess risks (compared to unrelated couples) by roughly one-half to two-thirds.

These risk reductions are meaningful but they do not prevent disease, which must be communicated as such. PGT-H should be framed as one tool among many for consanguineous families,

¹For example, large terminal or chromosome-scale ROH segments that are inconsistent with the expected pedigree structure may warrant clinical review for technical artifacts, uniparental disomy, or other non-Mendelian mechanisms. This quality-control use is distinct from the primary purpose of PGT-H, which is embryo prioritization based on genome-wide realized autozygosity.

complementing carrier screening, variant-level PGT, and prenatal diagnosis. Several limitations should be considered. Our modeling assumptions push our estimates in a conservative direction: Background autozygosity of 1-3%, common in populations with historically elevated consanguinity,³ would widen within-family variance and increase the absolute benefit of selection. Diagnostic yields in affected individuals may likely overstate what PGT-M can prevent in practice, since variants identified as diagnostic in affected probands would not necessarily meet clinical thresholds for PGT-M. We anticipate that autozygosity-based embryo selection may eventually become integrated into reproductive genomics in settings where consanguinity is common.

Methods

Simulation of embryo-level autozygosity in consanguineous families

We simulated consanguineous pedigrees to model the distribution of embryo-level F_{ROH} for first cousin and double-first-cousin couples. Founder haplotypes were generated on an idealized autosomal genome. Crossovers were modeled as a Poisson process along each chromosome, without interference, and recombination rates were determined by the sex-specific map for each parent¹⁸.

For each relationship type, we simulated 1,000 families. In each family, we drew founder haplotypes, propagated them forward through the consanguineous pedigree by sampling crossovers and segregating chromosomes at each meiosis, and produced 1,000 embryos for the final couple. Rather than simulating full genotypes, we traced the boundaries of founder haplotypes as they were transmitted across generations. For every embryo we recorded, for each autosomal chromosome, the piecewise-constant haplotype identity inherited on each chromosome. We then constructed the union of all breakpoints from the two homologous chromosomes of each embryo, partitioning each chromosome into minimal contiguous intervals over which haplotype identity was constant. An interval was classified as ROH (homozygous-by-descent) if the two haplotype IDs were identical, indicating that both copies traced back to the same founder haplotype. The physical lengths of all autosomal ROH segments were summed to obtain total ROH length per embryo, and F_{ROH} , which is expressed as a percentage, was computed as total autosomal ROH length divided by total autosomal genome length.

Within each family, we used the 1,000 simulated embryos to characterize the underlying F_{ROH} distribution. We estimated the family-specific mean F_{ROH} , which corresponds to random embryo choice, and the variance and range across embryos. To model embryo selection, we considered embryo cohort sizes $n \in \{2, 3, 5, 10, 20\}$. For each family and each n , we drew 10,000 independent sets of n embryos without replacement, recorded the minimum F_{ROH} in each set, and computed the expected minimum for that cohort size as the average across replicates.

From these quantities we derived the expected reduction in F_{ROH} relative to random embryo choice, defined as the difference between family mean F_{ROH} and expected minimum F_{ROH} . Simulations for first-cousin and double-first-cousin couples were analyzed separately and then summarized across the 1,000 families in each group.

Founders were assumed to have no background ROH, so simulated F_{ROH} reflects only autozygosity arising from the explicit consanguineous pedigree. We restricted attention to autosomes and did not simulate sex chromosomes. Within these constraints, the simulation captures stochastic segregation of ancestral haplotypes and the resulting variability in F_{ROH} among em-

bryos of consanguineous unions. All simulations and analyses were performed in Python using custom scripts.

Estimating risk reductions for monogenic and polygenic outcomes

We modeled risk reductions achieved by combining PGT-M and F_{ROH} -based embryo selection for outcomes with different genetic architectures. For largely monogenic outcomes (infant death, intellectual disability, congenital anomalies), we assumed that the excess risk above the outbred baseline is entirely attributable to increased recessive burden in consanguineous families. For polygenic outcomes (type 2 diabetes and childlessness) we assumed that log-odds vary approximately linearly with F_{ROH} and used effect sizes estimated in within-sibling analyses to minimize confounding from shared environment³⁻⁵.

Let p_{baseline} denote the lifetime risk in an outbred population (e.g. for type 2 diabetes we used published estimates for the U.S. population²⁵). For a parental relatedness corresponding to $F_{\text{ROH}} = F$, we let RR_F denote the risk ratio for offspring relative to $F = 0$, estimated from cohort studies of consanguinity and autozygosity^{2-5,7,15,16}. The absolute risk without screening is then

$$p_{\text{no screen}} = p_{\text{baseline}} \times RR_F.$$

To approximate the impact of PGT-M, we used the diagnostic yield in patient cohorts with the specified condition as a proxy for the fraction of recessive risk attributable to known pathogenic variants. For intellectual disability and congenital anomalies we drew on recent exome sequencing studies in diverse populations^{10,12-14,28}. If D denotes the diagnostic yield (e.g. 30–50% for severe intellectual disability; 26–50% for major congenital anomalies), we modeled PGT-M as removing a fraction D of the excess risk above p_{baseline} as well as the risk due to biallelic causes in a baseline, B , outbred population (assumed to be 4%), yielding

$$p_{\text{PGT-M}} = p_{\text{baseline}} [(1 - B) + B(1 - D) + (RR_F - 1)(1 - D)]$$

We neglected the slight reduction in F_{ROH} that is expected to result from PGT-M, given that PGT-M excludes embryos that are homozygous at the risk locus and thus have on average a slightly higher F_{ROH} compared to the other embryos. We note this is likely an overestimate of the achievable PGT-M benefit, because diagnostic yield in affected individuals may include variants that would not meet strict criteria for preimplantation testing. We also assume the PGT-M risk reduction of the baseline risk of biallelic monogenic causes is equal to that of the excess risk due to consanguinity.

Autozygosity-informed embryo selection further reduces risk by lowering F_{ROH} from its baseline family value F to a post-screening value F_{post} when selecting the lowest- F_{ROH} embryo among n candidates. For largely monogenic outcomes, we assumed that the residual excess risk after PGT-M scales linearly with F_{ROH} , so that

$$p_{\text{PGT-M+PGT-H}} = p_{\text{baseline}} \left[(1 - B) + B(1 - D) + (RR_F - 1)(1 - D) \frac{F_{\text{post}}}{F} \right]$$

For polygenic outcomes we instead modeled log-odds as proportional to F_{ROH} and applied the same F_{post}/F scaling on the logit scale based on within-sibling effect estimates. In both cases, we assumed that recessive contributions to p_{baseline} are negligible in outbred populations, consistent with empirical estimates in cohorts with low autozygosity.

Baseline risks for infant death, intellectual disability, congenital anomalies, and type 2 diabetes were drawn from U.S. vital statistics and large epidemiological studies²²⁻²⁵. We used fertility data from the National Center for Family and Marriage Research and the IPUMS Current Population Survey to approximate lifetime childlessness^{26,27}. When multiple sources were available, we averaged lifetime risk estimates across recent, high-quality studies.

References

- [1] Bittles Alan H.. *Consanguinity in Context*. Cambridge: Cambridge University Press 2012.
- [2] Ceballos Francisco C., Joshi Peter K., Clark David W., Ramsay Michèle, Wilson James F.. Runs of Homozygosity: Windows into Population History and Trait Architecture *Nature Reviews Genetics*. 2018;19:220–234.
- [3] Clark David W., Okada Yukinori, Moore Kristjan H. S., et al. Associations of Autozygosity with a Broad Range of Human Phenotypes *Nature Communications*. 2019;10:4957.
- [4] Ceballos Francisco C., Hazelhurst Scott, Clark David W., et al. Autozygosity Influences Cardiometabolic Disease-Associated Traits in the AWI-Gen Sub-Saharan African Study *Nature Communications*. 2020;11:5754.
- [5] Malawsky Daniel S., Van Walree Eva, Jacobs Benjamin M., et al. Influence of Autozygosity on Common Disease Risk across the Phenotypic Spectrum *Cell*. 2023;186:4514-4527.e14.
- [6] Gamsiz Ece D., Viscidi Emma W., Frederick Abbie M., et al. Intellectual Disability Is Associated with Increased Runs of Homozygosity in Simplex Autism *The American Journal of Human Genetics*. 2013;93:103–109.
- [7] Zlotogora Joël, Shalev Stavit A.. The Consequences of Consanguinity on the Rates of Malformations and Major Medical Conditions at Birth and in Early Childhood in Inbred Populations *American Journal of Medical Genetics Part A*. 2010;152A:2023–2028.
- [8] Modell Bernadette, Darr Aamra. Genetic Counselling and Customary Consanguineous Marriage *Nature Reviews Genetics*. 2002;3:225–229.
- [9] Hamamy Hanan, Antonarakis Stylianos E., Cavalli-Sforza Luigi Luca, et al. Consanguineous Marriages, Pearls and Perils: Geneva International Consanguinity Workshop Report *Genetics in Medicine*. 2011;13:841–847.
- [10] Fridman Hila, Yntema Helger G., Mägi Reedik, et al. The Landscape of Autosomal-Recessive Pathogenic Variants in European Populations Reveals Phenotype-Specific Effects *The American Journal of Human Genetics*. 2021;108:608–619.
- [11] Alkuraya Fowzan S.. Autozygome Decoded *Genetics in Medicine*. 2010;12:765–771.
- [12] Jamra Rami. Genetics of Autosomal Recessive Intellectual Disability *Medizinische Genetik*. 2018;30:323–327.
- [13] Al-Kasbi Ghalia, Al-Murshedi Fathiya, Al-Kindi Adila, et al. The Diagnostic Yield, Candidate Genes, and Pitfalls for a Genetic Study of Intellectual Disability in 118 Middle Eastern Families *Scientific Reports*. 2022;12:18862.
- [14] Wright Caroline F., Campbell Patrick, Eberhardt Ruth Y., et al. Genomic Diagnosis of Rare Pediatric Disease in the United Kingdom and Ireland *New England Journal of Medicine*. 2023;388:1559–1571.

- [15] Stoltenberg C., Magnus P., Lie R. T., Daltveit A. K., Irgens L. M.. Influence of Consanguinity and Maternal Education on Risk of Stillbirth and Infant Death in Norway, 1967-1993 *American Journal of Epidemiology*. 1998;148:452–459.
- [16] Small Neil, Kelly Brian, Malawsky Daniel S., Lodh Rajib, Oddie Sam, Wright John. Mortality, Morbidity and Educational Outcomes in Children of Consanguineous Parents in the Born in Bradford Cohort *Wellcome Open Research*. 2024;9:319.
- [17] Sheridan Eamonn, Wright John, Small Neil, et al. Risk Factors for Congenital Anomaly in a Multiethnic Birth Cohort: An Analysis of the Born in Bradford Study *The Lancet*. 2013;382:1350–1359.
- [18] Bhérer Claude, Campbell Christopher L., Auton Adam. Refined Genetic Maps Reveal Sexual Dimorphism in Human Meiotic Recombination at Multiple Scales *Nature Communications*. 2017;8:14994.
- [19] Klausner Liraz, Revital Ateret, Lencz Todd, Carmi Shai. PEstimate: Predicting Offspring Disease Risk after Polygenic Embryo Screening 2025. medRxiv preprint.
- [20] Al-Dewik Nader, Mohd Howaida, Al-Mureikhi Mariam, et al. Clinical Exome Sequencing in 509 Middle Eastern Families with Suspected Mendelian Diseases: The Qatari Experience *American Journal of Medical Genetics Part A*. 2019;179:927–935.
- [21] Chundru V Kartik, Zhang Zhancheng, Wigdor Emilie M, et al. Federated analysis of autosomal recessive coding variants in 29,745 developmental disorder patients from diverse populations *Nature Genetics*. 2024;56:2046–2053.
- [22] Jansen Sandra, Vissers Lisenka E. L. M., De Vries Bert B. A.. The Genetics of Intellectual Disability *Brain Sciences*. 2023;13:231.
- [23] Ely Danielle, Driscoll Anne. Infant Mortality in the United States, 2023: Data From the Period Linked Birth/Infant Death File tech. rep.National Center for Health Statistics (U.S.) 2025.
- [24] Centers for Disease Control and Prevention . Update on Overall Prevalence of Major Birth Defects — Atlanta, Georgia, 1978–2005 *MMWR Morb Mortal Wkly Rep*. 2008;57:1–5.
- [25] Narayan K.M.V., Boyle James P., Thompson Theodore J., Gregg Edward W., Williamson David F.. Effect of BMI on Lifetime Risk for Diabetes in the U.S. *Diabetes Care*. 2007;30:1562–1566.
- [26] Guzzo Karen, Loo Jaden. Number of Children to Women Aged 40-44, 1980-2022 tech. rep.National Center for Family and Marriage Research 2023.
- [27] Flood Sarah, King Miriam, Rodgers Renae, et al. IPUMS, Current Population Survey: Version 11.0 2023.
- [28] Slavotinek Anne, Rego Shannon, Sahin-Hodoglugil Nuriye, et al. Diagnostic Yield of Pediatric and Prenatal Exome Sequencing in a Diverse Population *npj Genomic Medicine*. 2023;8:10.

- [29] Bennett Robin L., Motulsky Arno G., Bittles Alan, et al. Genetic Counseling and Screening of Consanguineous Couples and Their Offspring: Recommendations of the National Society of Genetic Counselors *Journal of Genetic Counseling*. 2002;11:97–119.
- [30] Hamamy Hanan. Consanguineous Marriages: Preconception Consultation in Primary Health Care Settings *Journal of Community Genetics*. 2012;3:185–192.
- [31] Clarke Angus J., Wallgren-Pettersson Carina. Ethics in Genetic Counselling *Journal of Community Genetics*. 2019;10:3–33.
- [32] Thain Emily, Shuman Cheryl, Miller Kristen, et al. Prenatal and Preconception Genetic Counseling for Consanguinity: Consanguineous Couples’ Expectations, Experiences, and Perspectives *Journal of Genetic Counseling*. 2019;28:982–992.
- [33] Resta Robert, Biesecker Barbara Bowles, Bennett Robin L., et al. A New Definition of Genetic Counseling: National Society of Genetic Counselors’ Task Force Report *Journal of Genetic Counseling*. 2006;15:77–83.
- [34] Anomaly Jonathan, Davidson Ivan, Moore Spencer, Wolfram Tobias, Young Alexander Strudwick. The Ethics of Embryo Screening *PhilPapers*. 2025.
- [35] Craig Alexander, Wartschinski Laura, Eyre Mathew, Davidson Ivan, Christensen Michael, Wolfram Tobias. Stage-Structured, Distributional Prediction of IVF Outcomes with Conditional Updating 2025.
- [36] Li Jeremiah H., Ahangari Mohammad, Wolfram Tobias, Christensen Michael. ImputePGTA: Imputing Embryo Genomes from Low-pass Sequencing and Whole-genome Parental Data *medRxiv*. 2025.
- [37] Mason-Otey Akailah, Seifer David B.. United States Racial/Ethnic Disparities in PGT-A Use: An Analysis of 2014–2020 SART CORS Database *Reproductive Biology and Endocrinology*. 2025;23:66.
- [38] Thomas Alun, Skolnick Mark H., Lewis Cathryn M.. Genomic Mismatch Scanning in Pedigrees *Mathematical Medicine and Biology*. 1994;11:1–16.
- [39] Edge Michael (Doc), Coop Graham. How lucky was the genetic investigation in the Golden State Killer case? *bioRxiv*. 2019.
- [40] Erlich Yaniv, Shor Tal, Pe’er Itsik, Carmi Shai. Identity Inference of Genomic Data Using Long-Range Familial Searches *Science*. 2018;362:690–694.
- [41] Karavani Ehud, Zuk Or, Zeevi Danny, et al. Screening Human Embryos for Polygenic Traits Has Limited Utility *Cell*. 2019;179:1424–1435.e8.
- [42] Campbell Christopher L., Furlotte Nicholas A., Eriksson Nick, Hinds David, Auton Adam. Escape from Crossover Interference Increases with Maternal Age *Nature Communications*. 2015;6:6260.

Supplement: A model for ROH reduction

We follow the classic model of Thomas et al.³⁸ and its recent applications^{39,40}. Denote by m the number of meioses from the common ancestor to a hypothetical child of the related couple. Following Figure 1, $m = 6$ when the parents are first cousins or double first cousins, and $m = 8$ for second cousins. The model assumes that recombination chops each chromosome into equal-length blocks. Denote by L the total autosomal genome length in Morgan (equal to the number of autosomal recombination events per meiosis) and by $c = 22$ the number of autosomes; the number of blocks is $mL + c$. The two copies of each block of the hypothetical child are identical by descent with probability F , which is the inbreeding coefficient. We have $F = 1/16$ for first cousins, $F = 1/8$ for double first cousins, and $F = 1/64$ for second cousins. Thus, the number of blocks that would be homozygous (by descent from the recent common ancestor) in the child is approximately a binomial variable, $\text{Bin}(mL + c, F)$. We can approximate F_{ROH} as f_H , the proportion of homozygous blocks. The variance of f_H is

$$\text{Var}(f_H) = \frac{F(1-F)}{mL+c}.$$

The above expression has two limitations. First, it is an underestimate of the variance of F_{ROH} , because, in reality, block lengths are variable due to the randomness of recombination. On the other hand, the above expression for $\text{Var}(f_H)$ includes variability both between-family and within-family, whereas only the within-family variability is of interest for embryo screening. It is thus an overestimate of the within-family variance of F_{ROH} . Overall, these two deviations partly cancel out, so we use the above expression as an approximation for the within-family variance $\text{Var}_w(F_{\text{ROH}})$.

In previous work (Karavani et al.⁴¹), embryo selection for traits was considered, and the mean change in a polygenic trait relative to random selection was derived. Suppose n embryos are available. Figure 1 shows that the distribution of F_{ROH} within family is approximately normal, similarly to the distribution of a polygenic score. We thus use the result from Karavani et al. Denote by R the value of the maximum of n standard normal variables. Asymptotically, $E(R) \approx 1.09 \sqrt{\ln(n)}$. The mean difference in F_{ROH} between a randomly selected embryo and the embryo with the lowest F_{ROH} is given by

$$\text{Mean ROH reduction} \approx \text{SD}_w(f_H) E(R) \approx 1.09 \sqrt{\frac{F(1-F)}{mL+c}} \sqrt{\ln(n)}.$$

Plugging $L = 35 \text{ M}^{42}$, we have, for first cousins ($F = \frac{1}{16}$, $m = 6$):

$$\text{Mean ROH reduction} \approx 1.09 \sqrt{\frac{(\frac{1}{16})(\frac{15}{16})}{6 \cdot 35 + 22}} \sqrt{\ln(n)} \approx 1.73\% \sqrt{\ln(n)}.$$

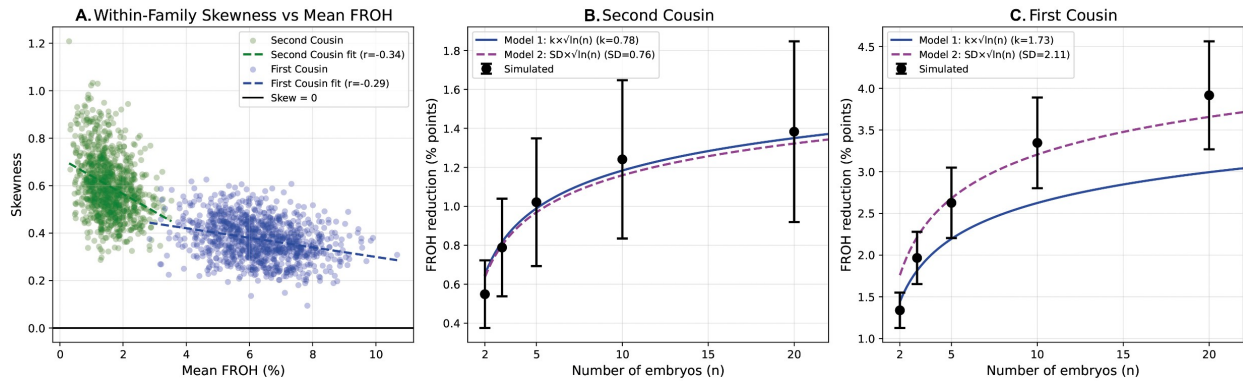
For second cousins ($F = \frac{1}{64}$, $m = 8$):

$$\text{Mean ROH reduction} \approx 1.09 \sqrt{\frac{(\frac{1}{64})(\frac{63}{64})}{8 \cdot 35 + 22}} \sqrt{\ln(n)} \approx 0.78\% \sqrt{\ln(n)}.$$

We also consider a second model of within-family F_{ROH} variance estimation derived from the simulation framework. Since the distribution has positive skew (Supplementary Figure 1A) and

we are interested in approximating the lower half of the distribution as normal, we estimated within-family variance by reflecting the values below the median and estimating variance using this augmented, symmetric distribution.

We then compared the simulated reductions from Figure 3 to the theoretical expectations derived here. Both approaches work well in the case of second cousin parental relatedness (Supplementary Figure 1B). However, in the case of first cousin parental relatedness, the initial model appeared to underestimate the simulated reduction in F_{ROH} , while the second model provided a relatively good fit (Supplementary Figure 1C). Thus, it appears that modelling risk reduction based on a normal approximation of the F_{ROH} distribution was appropriate, though our derivation of within-family variance in F_{ROH} is miscalibrated in the case of first cousin parental relatedness.



Supplementary Figure 1: Theoretical expected reduction in embryo ROH from screening in second- and first-cousin unions.

(A) Within-family skewness of F_{ROH} distribution versus mean F_{ROH} for 1,000 simulated families from second-cousin (green) and first-cousin (blue) parental matings. Dashed lines show linear fits; horizontal black line indicates zero skewness.

(B) Expected absolute reduction in F_{ROH} when selecting the lowest- F_{ROH} embryo from sets of $n = 2, 3, 5, 10, 20$ embryos for second-cousin parental relatedness. Black points show simulated means \pm SD. Model 1 (solid blue): $k\sqrt{\ln(n)}$ with $k = 0.78$; Model 2 (dashed magenta): $SD\sqrt{\ln(n)}$ where $SD = 0.76$ is the symmetric within-family standard deviation.

(C) As in (B), for first-cousin parental relatedness, with $k = 1.73$ and $SD = 2.11$.