What is the utility of screening human embryos for polygenic traits?

Shai Carmi

Public Health

The Hebrew University of Jerusalem

scarmilab.org

@ShaiCarmi

HUJI

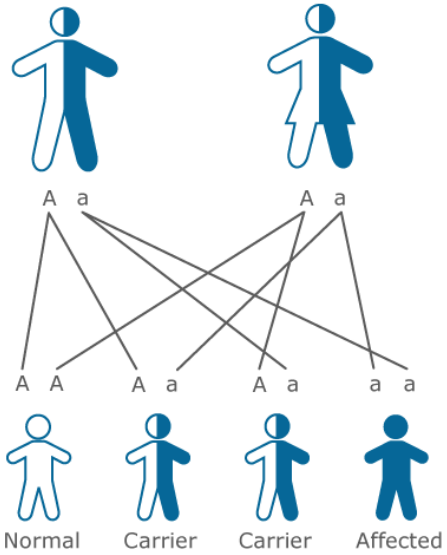
December 2019

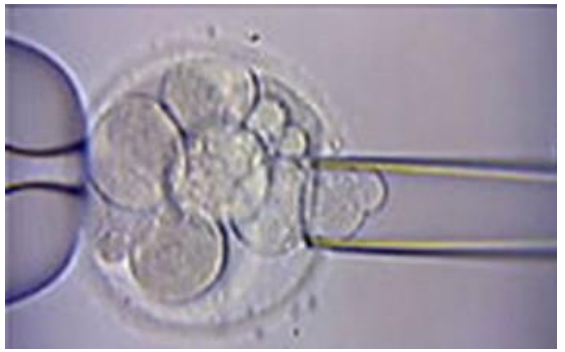
בית הספר לבריאות הציבור ורפואה קהילתית
של האוניברסיטה העברית והדסה ע"ש בראון
Braun School of Public Health and Community Medicine
Hebrew University-Hadassah

הפקולטה לרפואה
FACULTY OF MEDICINE
האוניברסיטה העברית בירושלים
THE HEBREW UNIVERSITY OF JERUSALEM

Genetic screening of embryos

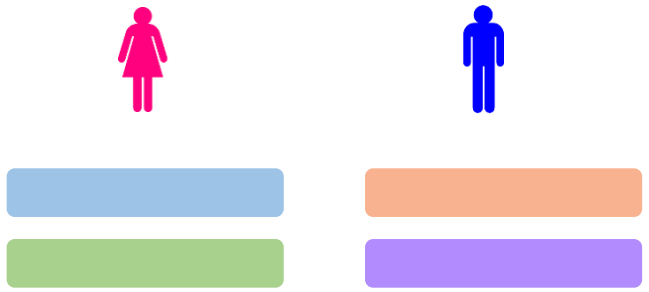
- Why?
  - Mendelian disease mutations
  - Recurrent pregnancy loss
- How?
  - Grow IVF embryos for 3-5 days
  - Amplify DNA from a single cell
- What?
  - Traditionally: single mutations, aneuploidy
  - Now: whole-genome haplotypes, CNVs
  - Universal, fast, accurate, low cost





How could it be possible?

Parents (microarray)



- Embryos are a mosaic of the parents
- Only need to infer crossover locations
- (Up to *de-novo* mutations)

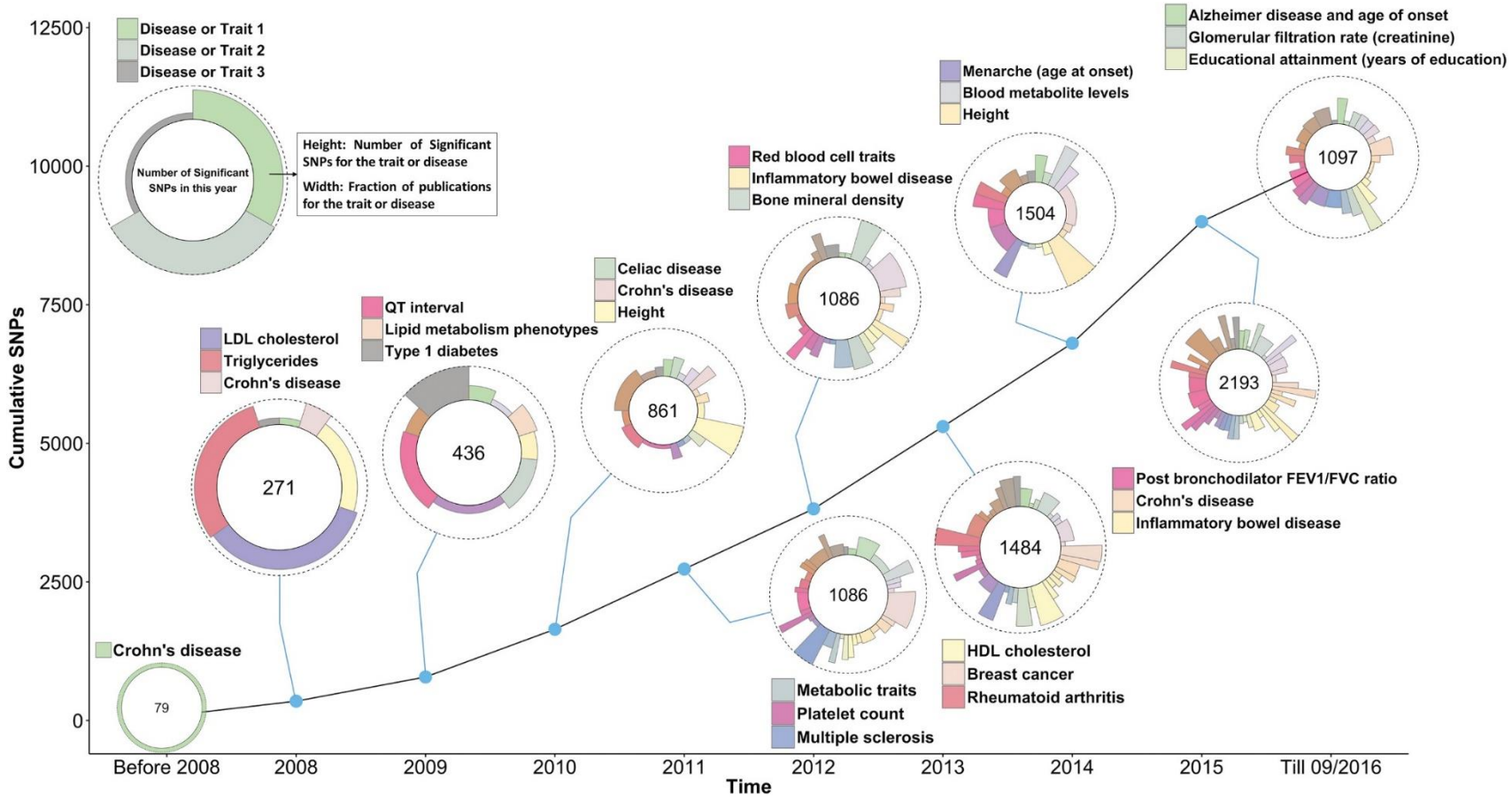
Embryo 1

Embryo 2

Embryo 3

Array/low-coverage sequencing

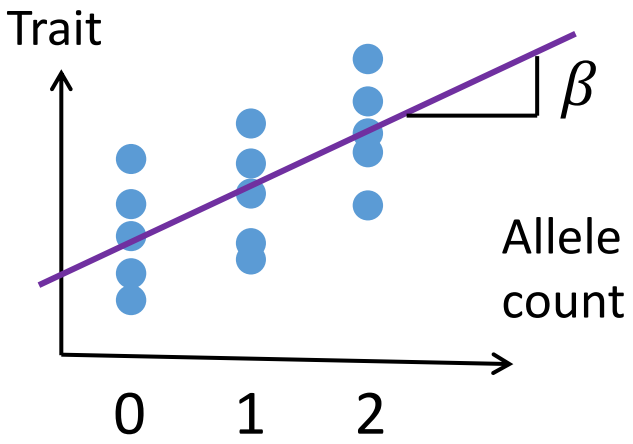
In parallel, progress in complex trait genetics



Visscher et al., 10 Years of GWAS Discovery: Biology, Function, and Translation, AJHG, 2017

Discovery ➔ prediction

- Summary statistics are publicly/freely available



- Effect size:
- Quantitative traits: increase in trait with each allele (regression slope)
- Disease (binary) traits: log odds-ratio

SNP	Chr	Position	Effect allele	P-value	Effect size $\hat{\beta}$
rs1234	1	134346223	A	$2 \cdot 10^{-5}$	0.001
rs2345	3	124572521	G	$4 \cdot 10^{-3}$	-0.0006
rs3456	6	73422152	A	$2 \cdot 10^{-8}$	0.02
rs4567	14	66452342	C	$7 \cdot 10^{-4}$	-0.003

Polygenic scores (PS)

- Using summary statistics, we can predict the trait of a new individual

- $PS = \sum_{i=1}^{M} \hat{\beta}_i g_i$

- M: number of SNPs
- g_i: number of effect alleles at SNP i (0,1,2)
- $\hat{\beta}_i$: estimated effect size at SNP i

- Statistical methods refine the set of SNPs and the weights

Performance of polygenic scores

- h^2: proportion of variance explained by genetics (heritability)
- h^2_{snp}: proportion of variance explained by chip SNPs (SNP heritability)
- r^2_{ps}: proportion of variance explained by score

$$r^2_{ps} < h^2_{snp} < h^2 < 1$$

Trait	h^2	h^2_{snp}	r^2_{ps}	References (PMID)
Height	70-80%	46%	25%	19818695,30124842
BMI	40-70%	25%	10%	22645519, 30124842
Educational attainment	63%	22%	12%	25985137, 23722424, 30038396
Cognitive function	50%	20%	5%	25985137, 29942086
LDL cholesterol	40-50%	13%	3.1-4.7%	17903299, 23766260, 30127800
Blood pressure	47%	14%	4%	25985137, 30224653

Implications

- Screening embryos for complex traits now feasible

- At least one company is already offering the test

Screening embryos for complex traits now feasible

- From GP website:
  - Type 1 and Type 2 Diabetes
  - Coronary Artery Disease, Heart Attack Risk, Hypercholesterolemia, Hypertension
  - Breast Cancer, Testicular Cancer, Prostate Cancer, Malignant Melanoma, Basal Cell Carcinoma
  - Intellectual Disability
  - Idiopathic Short Stature

European Journal of Medical Genetics

Validation of concurrent preimplantation genetic testing for polygenic and monogenic disorders, structural rearrangements, and whole and segmental chromosome aneuploidy with a single universal platform

Nathan R. Treff[a,*], Raymond Zimmerman[a], Elan Bechor[a], Jeff Hsu[a], Bhavini Rana[a], Jens Jensen[a], Jeremy Li[a], Artem Samoilenko[a], William Mowrey[a], James Van Alstine[a], Mark Leondires[b], Kathy Miller[b], Erica Paganetti[b], Louis Lello[c], Steven Avery[c], Stephen Hsu[c], Laurent C.A. Melchior Tellier[a]

April 2019

frontiers in Endocrinology

December 2019

Utility and First Clinical Application of Screening Embryos for Polygenic Disease Risk Reduction

Nathan R. Treff[1,2], Jennifer Eccles[1,2], Lou Lello[1,3], Elan Bechor[1], Jeffrey Hsu[1], Kathryn Plunkett[1,2], Raymond Zimmerman[1,2], Bhavini Rana[1,2], Artem Samoilenko[1], Steven Hsu[3] and Laurent C. A. M. Tellier[1,2,3]*

[1] Genomic Prediction Inc., North Brunswick, NJ, United States, [2] Genomic Prediction Clinical Laboratory, North Brunswick, NJ, United States, [3] Department of Physics and Astronomy, Michigan State University, East Lansing, MI, United States

Obviously, ethical concerns

Biotechnology / DNA Testing

Eugenics 2.0: We're at the Dawn of Choosing Embryos by Health, Height, and More

Will you be among the first to pick your kids' IQ? As machine learning unlocks predictions from DNA databases, scientists say parents could have choices never before possible.

MIT Technology Review
November 2017
Antonio Regalado

Baby steps

A slippery slope towards designer babies?

The Economist, November 2018

A new genetic-screening technique lets parents choose embryos most likely to grow into healthy adults

New Scientist, November 2018

Exclusive: A new test can predict IVF embryos' risk of having a low IQ

A new genetic test that enables people having IVF to screen out embryos likely to have a low IQ or high disease risk could soon become available in the US

The Times, November 2018

New test can predict intelligence in embryos

Obviously, ethical concerns

Embryo editing for higher IQ is a fantasy. Embryo profiling for it is almost here

STAT, February 2019

By ERIK PARENS, PAUL APPELBAUM, *and* WENDY CHUNG / FEBRUARY 12, 2019

Opinion | THE PRIVACY PROJECT

NY Times, April 2019

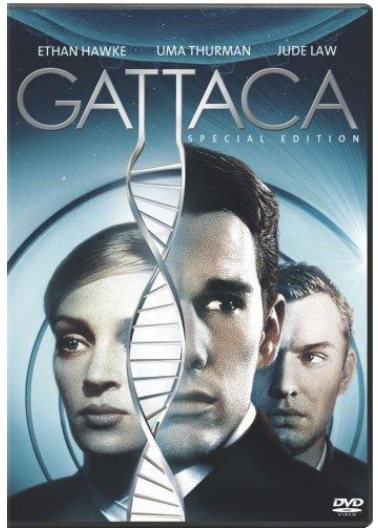
Making Babies in the Year 2045

Huge pools of health data collected over the past generation allow you to pick many of your child's genetic traits. Are you comfortable with that?

IVF couples could be able to choose the 'smartest' embryo

The Guardian, May 2019

US scientist says it will be possible to rank embryos by 'potential IQ' within 10 years



BUT…

Bioethics ISSN 0269-9702
Volume 15 Number 5/6 2001

PROCREATIVE BENEFICENCE: WHY WE SHOULD SELECT THE BEST CHILDREN

JULIAN SAVULESCU

No data!

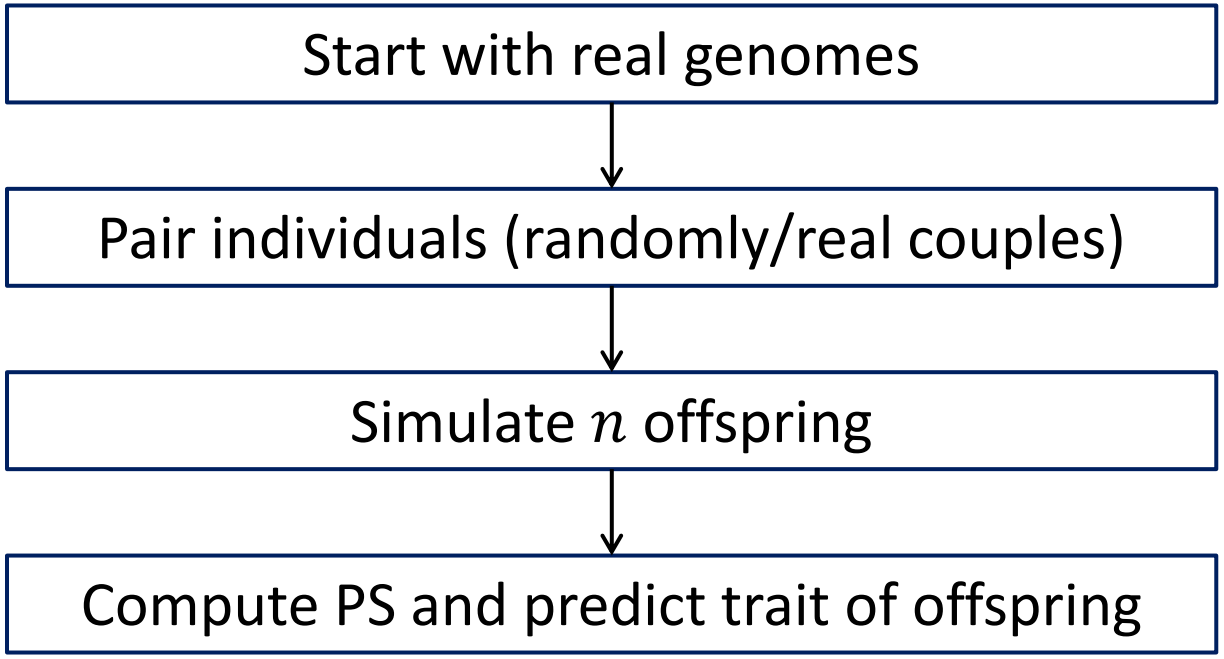
Does it work? What are the expected outcomes?

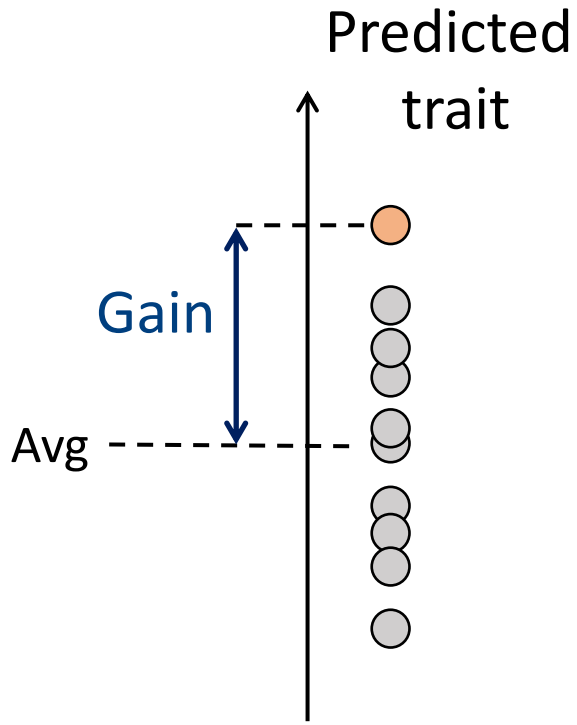
So far: economic analysis, no empirical data (Shulman and Bostrom, 2014; gwern blog)

Our approach:

1. Simulations based on real data
2. Quantitative genetic model
3. Large nuclear families

Simulations overview





Gain = (prediction of top-scoring embryo) — (average prediction)

Traits/cohorts

Gil Atzmon, Nir Barzilai
Einstein College of Medicine

- **Height**
- 102 couples, 700k SNPs
- Ashkenazi Jews, a longevity study (Sathyan et al., 2018)

- **Cognitive ability (IQ)**
- 919 young males, 480k SNPs
- Greek schizophrenia study (Stefanis et al., 2004)

Nikos Stefanis, Alex Hatzimanolis, Nikolaos Smyrnis,
Dimitrios Avramopoulos, University of Athens

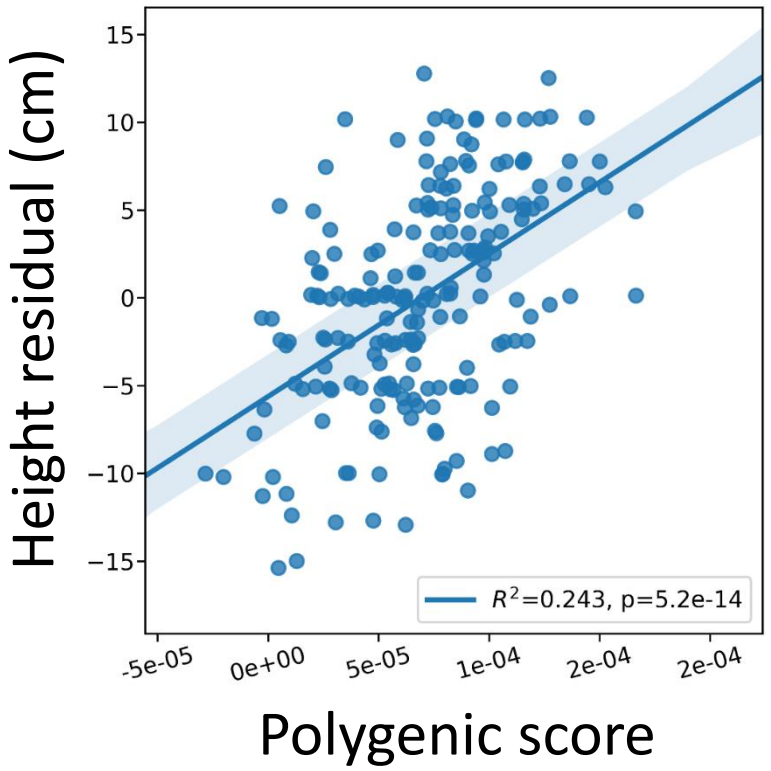
Polygenic scores

- **Height**
- Yengo et al., 2018
- 700k individuals

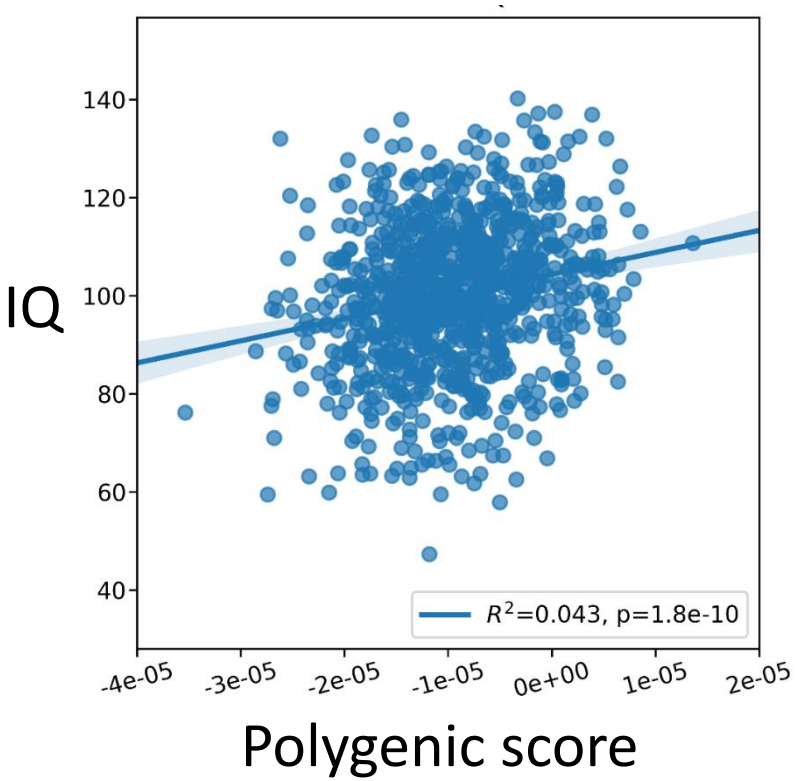
- **IQ**
- Savage et al., 2018
- 270k individuals

- Cross-validation for tuning parameters

Height, $r_{ps}^2 \approx 24\%$

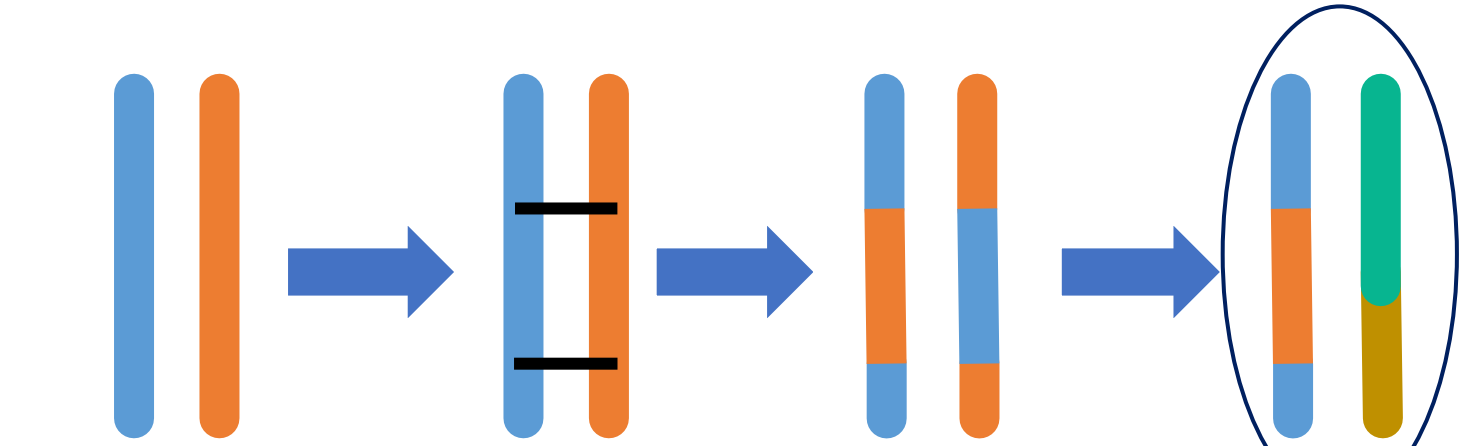


IQ, $r_{ps}^2 \approx 4.3\%$



Simulating embryos

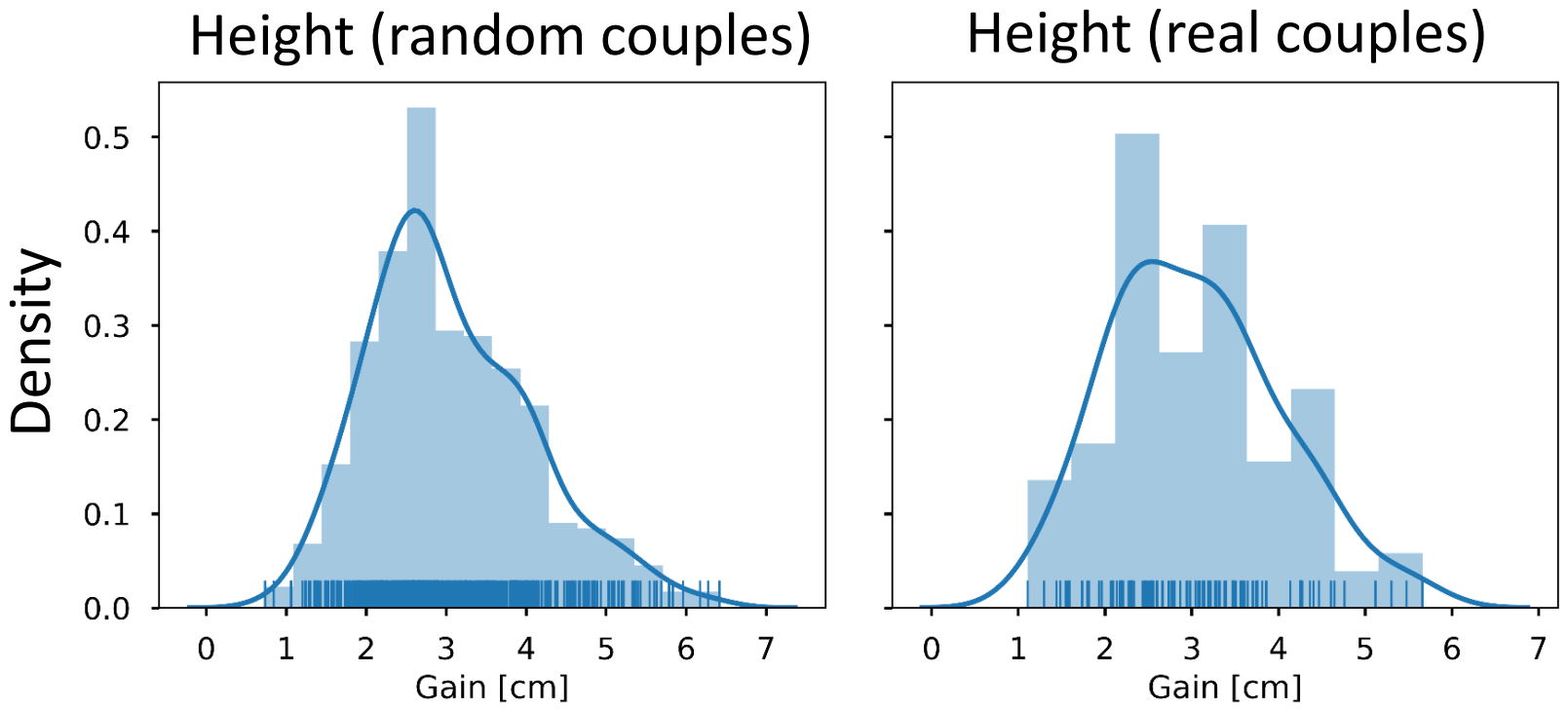
- To simulate a gamete:
  - Poisson number of crossovers, random placement
  - Independent random segregation
- Gametes paired to form diploids
- Polygenic scores computed and phenotypes predicted



Parental chrs

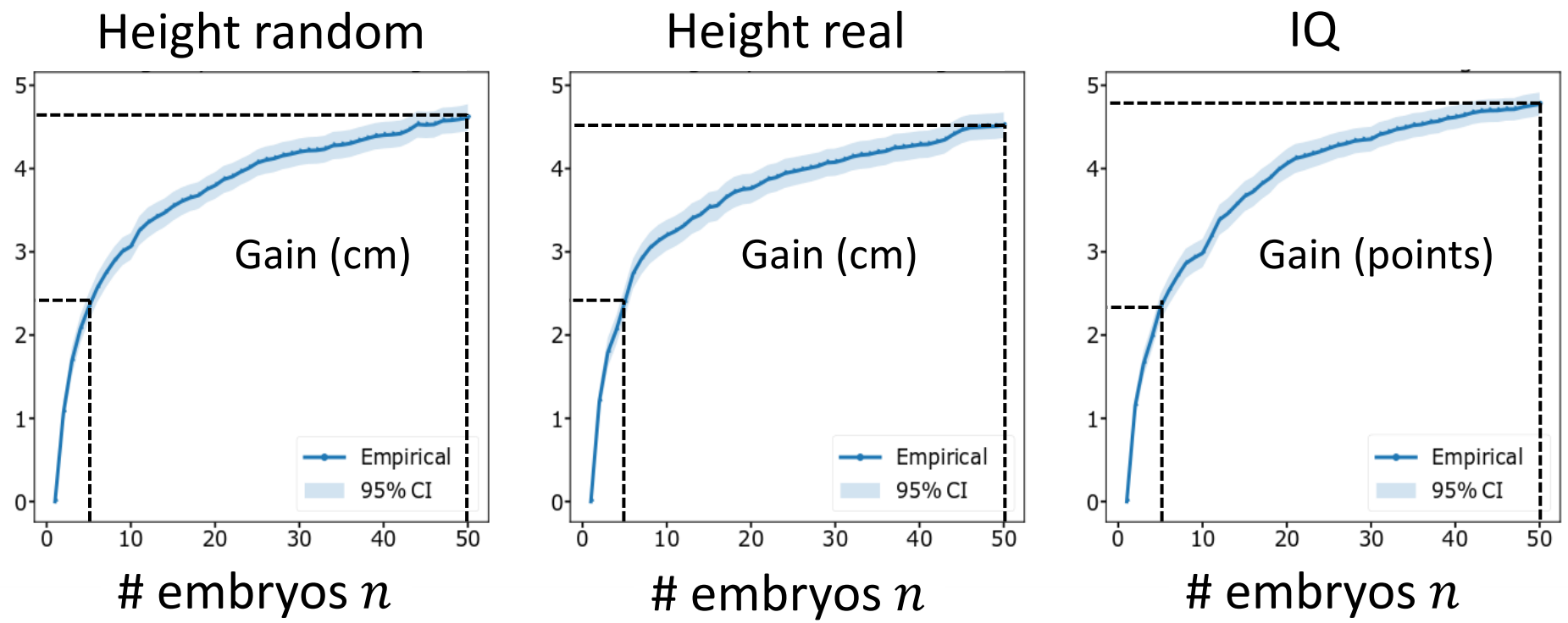
Experiments ($n = 10$ embryos)

- Gain in height: **2-4 cm**
- Real/random families behave similarly
- Gain in IQ: **2-4 points**



More/less embryos?

- $n = 50$: gain is ≈4.5 cm/IQ points
- $n = 5$: gain is ≈2.5 cm/IQ points



Questions

- Are these results expected?
- What happens for even more embryos?
- What happens if we have better predictors?

- We need a quantitative genetic model

A model for the polygenic scores of sibs

- $PS_1, PS_2, \ldots, PS_n \sim MVN(\mathbf{0}_n, \mathbf{\Sigma})$
- $\mathbf{\Sigma} = \sigma_z^2 r_{ps}^2 \begin{bmatrix} 1 & \cdots & 1/2 \\ \vdots & \ddots & \vdots \\ 1/2 & \cdots & 1 \end{bmatrix}$

$\boldsymbol{\sigma_z^2}$: variance of trait

$\boldsymbol{r_{ps}^2}$: proportion of variance explained by PS

For siblings: $\mathrm{Cov}(PS_i, PS_j) = \frac{1}{2}\mathrm{Var}(PS_i)$

The gain under the model

- The **gain**: (PS of best embryo) minus (average PS for the family)
- $G = \max(PS_1, PS_2, \dots, PS_n) - \frac{1}{n}(PS_1 + PS_2 + \cdots + PS_n)$
- $E(G) = E(\max(PS_1, PS_2, \dots, PS_n))$

Decomposing the polygenic scores

- $\boldsymbol{PS} \sim MVN(\mathbf{0}_n, \boldsymbol{\Sigma}) = \boldsymbol{Y} + \boldsymbol{Z}$
- $\boldsymbol{\Sigma} = \sigma_z^2 r_{ps}^2 \begin{bmatrix} 1 & \cdots & 1/2 \\ \vdots & \ddots & \vdots \\ 1/2 & \cdots & 1 \end{bmatrix} = \frac{1}{2}\sigma_z^2 r_{ps}^2 \begin{bmatrix} 1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & 1 \end{bmatrix} + \frac{1}{2}\sigma_z^2 r_{ps}^2 \begin{bmatrix} 1 & \cdots & 1 \\ \vdots & \ddots & \vdots \\ 1 & \cdots & 1 \end{bmatrix}$

$\boldsymbol{Y}$ — Independent normals

$\boldsymbol{Z}$ — Identical normals

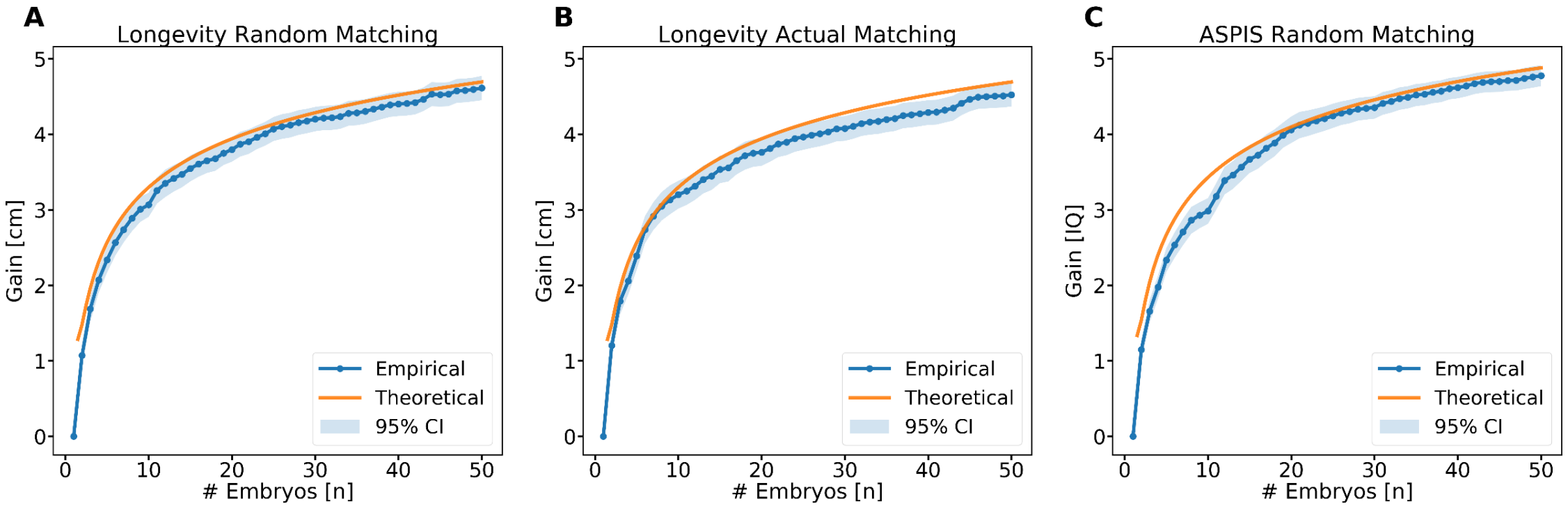
Do not affect the max

The mean gain

- $E(G) = E(\max(Y_1, Y_2, \dots, Y_n))$
- $Y_i \sim N\left(0, \frac{1}{2}\sigma_z^2 r_{ps}^2\right)$ are independent
- Using extreme value theory:
- $$E(G) = \frac{\sigma_z r_{ps}}{\sqrt{2}}\left[\Phi^{-1}\left(1-\frac{1}{n}\right) + \frac{\gamma}{n\phi\left(\Phi^{-1}\left(1-\frac{1}{n}\right)\right)}\right]$$

$\boldsymbol{\Phi^{-1}}$: inverse normal CDF
$\boldsymbol{\phi}$: normal PDF
$\boldsymbol{\gamma}$: Euler-Mascheroni constant

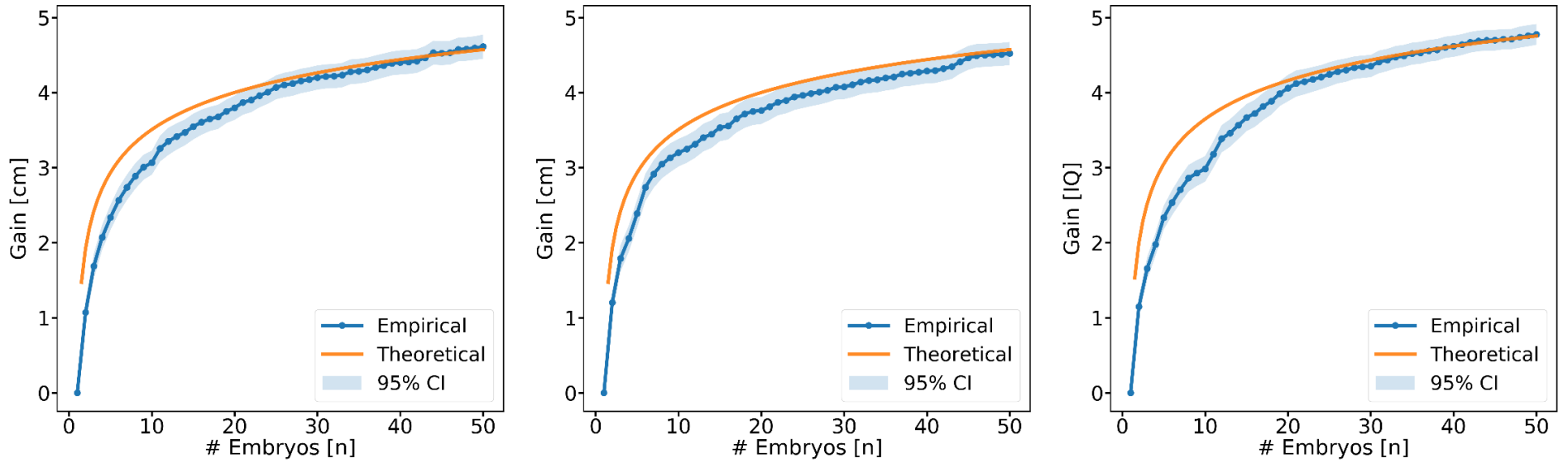
Confirming the theory



A useful approximation

σ_z: SD of the trait

$$E(Gain) \approx \underbrace{\sigma_z r_{ps}}_{\text{Accuracy of the score}} \cdot \underbrace{O\left(\sqrt{\log n}\right)}_{\text{Contribution of the number of embryos}}$$



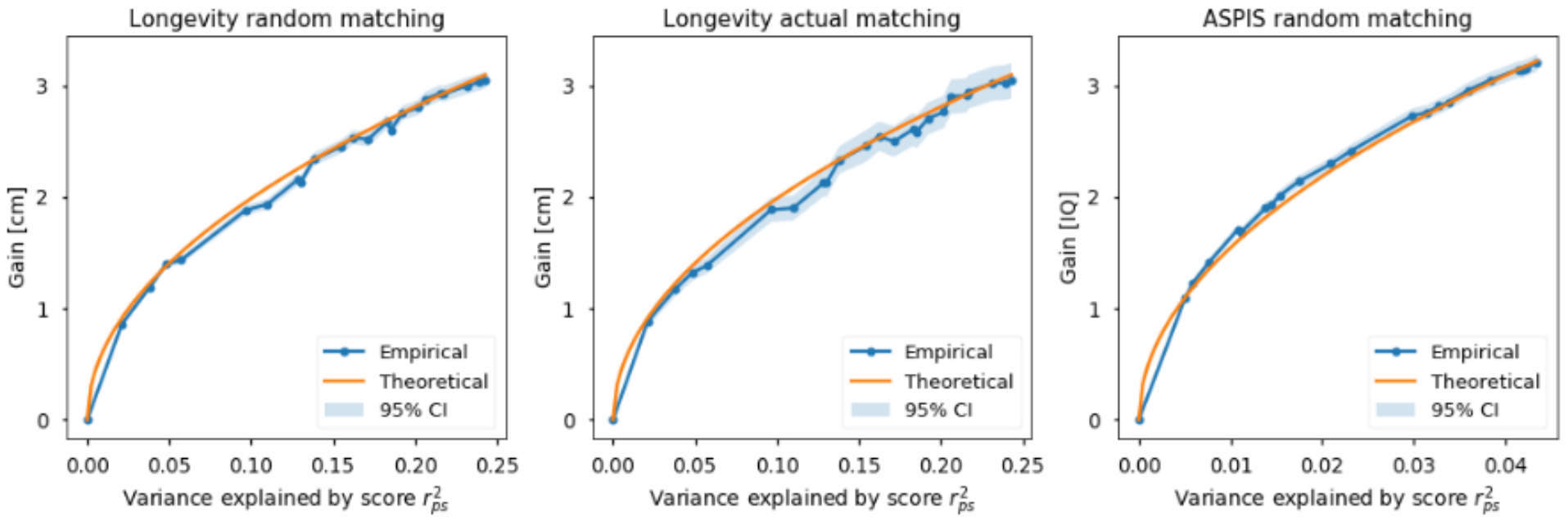
The effect of the predictor

- $E(G) \propto r_{ps} = \sqrt{r_{ps}^2}$

Can estimate the gain for future predictors!

Using growing subsets of chrs

$n = 10$ embryos

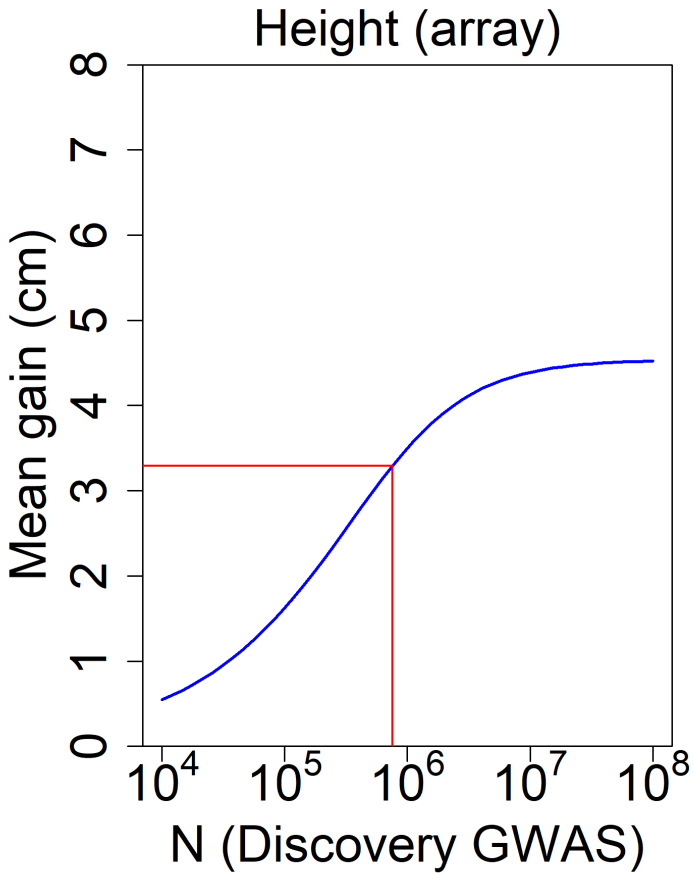


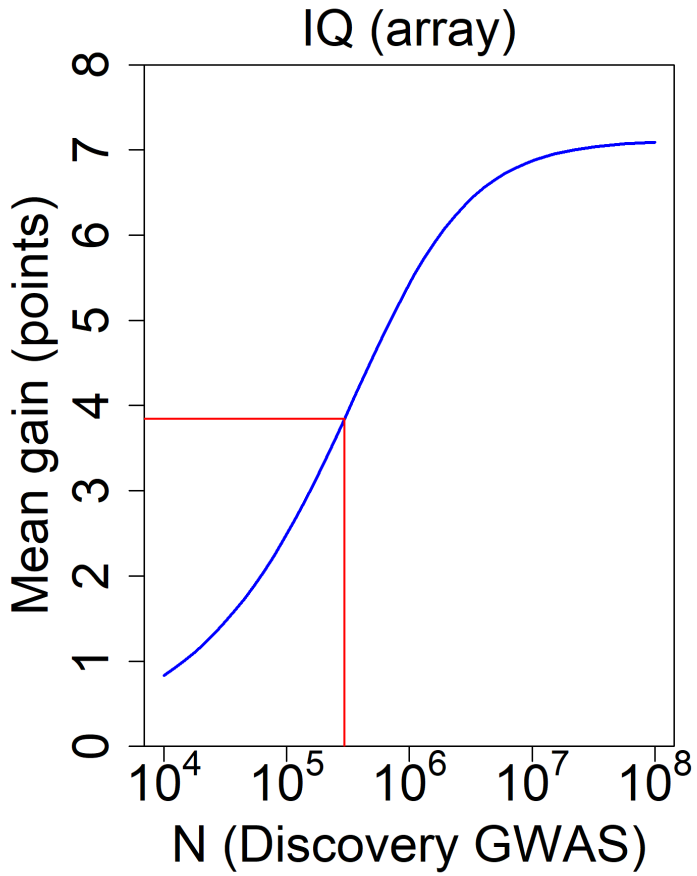
Larger GWASs?

- $r_{ps}^2 = \frac{h_{snp}^2}{1+\frac{M}{Nh_{snp}^2}}$

- M: Effective #markers
- N: GWAS sample size
- h_{snp}^2: chip-based heritability

- Wray et al., 2019
- Pasaniuc and Price, 2017





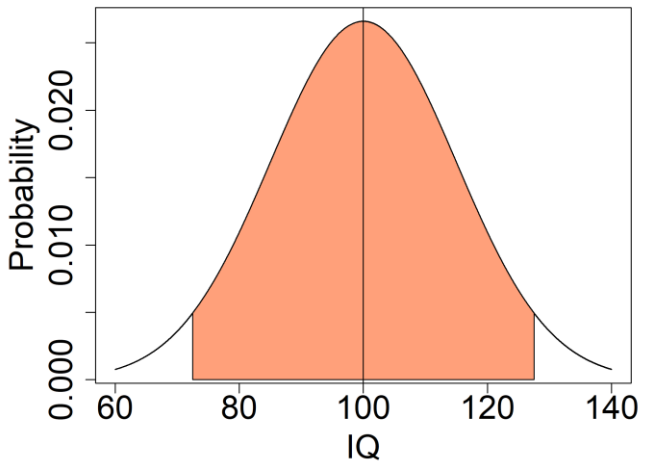
10 embryos

The actual gain is uncertain

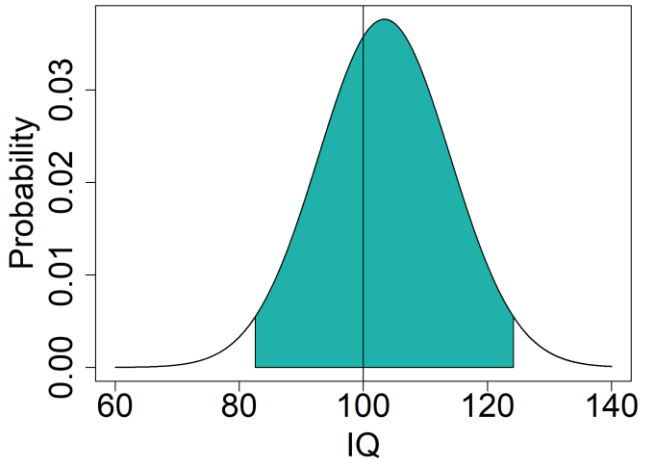
1. Embryos are random
2. Non-score genetic factors
3. Non-genetic factors

$n = 10, h^2 = 0.5, r_{ps}^2 = 4.3\%$

No selection
[72,128]

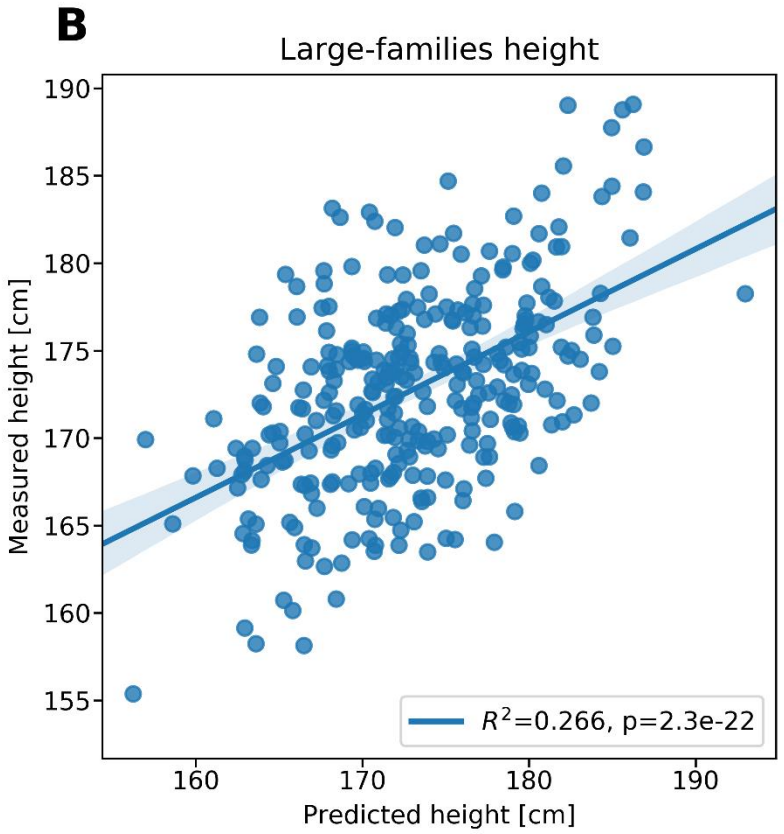


Selection
[83,124]



But what about an actual experiment?...

- Ethical issues + too long
- But we can consider large nuclear families
- 28 Ultra-orthodox Jewish families (Israel/US)
- 9.6 **adult** children per family (range: 3-20)
- Genotype + height available!
- Danny Zeevi, Leonid Kruglyak, UCLA

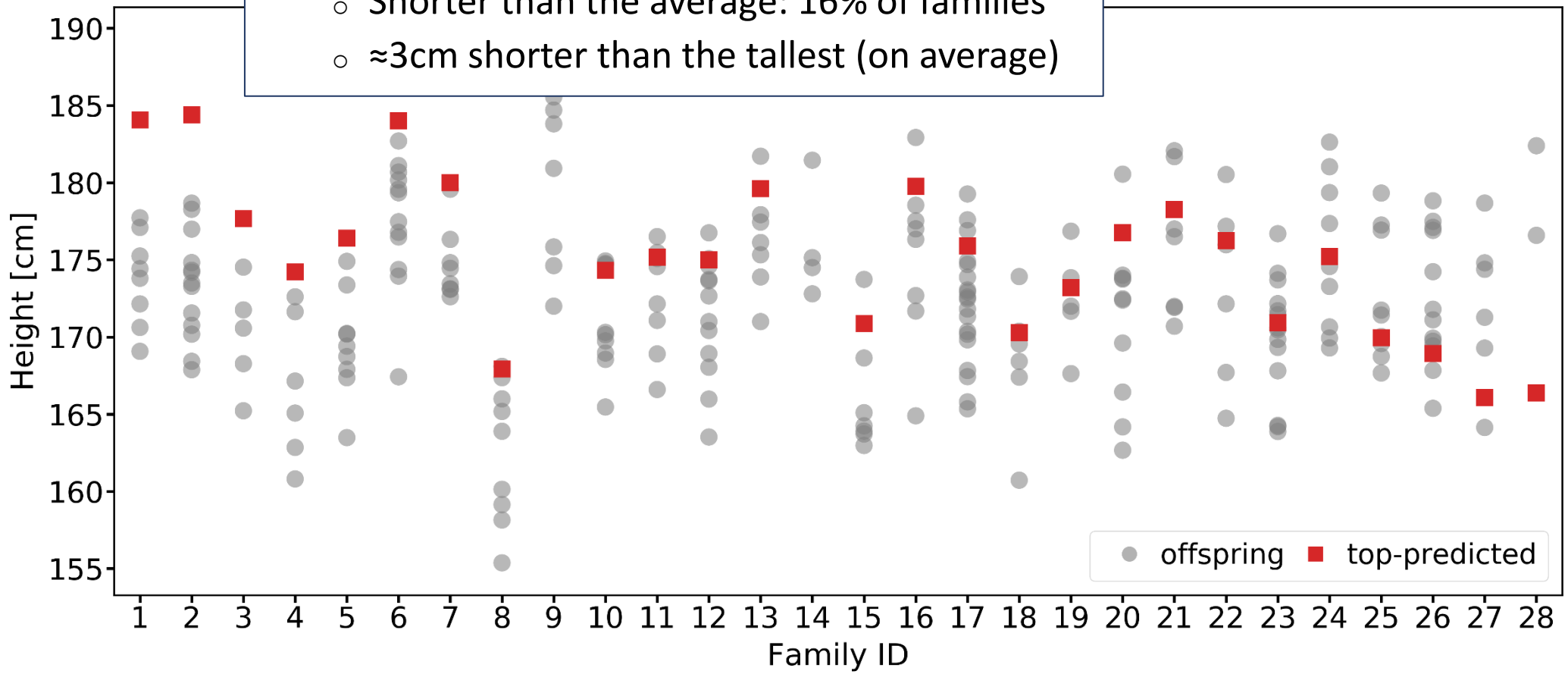


An “experiment”

Top-scoring embryo ($n = 7$) is:

- Tallest: 32% of families
- Shorter than the average: 16% of families
- ≈3cm shorter than the tallest (on average)

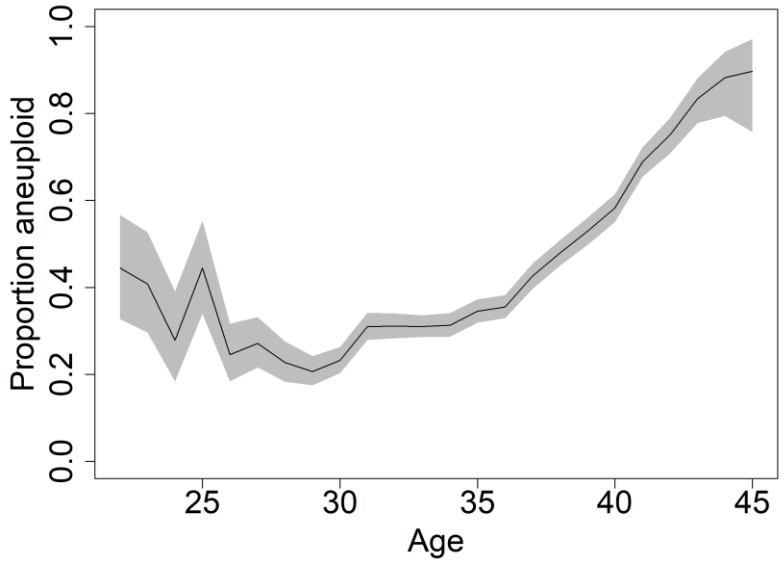
Danny Zeevi, Leonid Kruglyak, UCLA



Many practical limitations

- Prediction accuracy lower within families and across populations, assortative mating
- Advanced maternal age: less viable embryos
- Selection for multiple traits
  - Gain decreases by $\sqrt{T}$-fold
- Risk of unknown health issues

From Franasiak et al., 2014



Summary

- Current gain: ≈2-3 cm/IQ points
- Improved predictors will increase the gain substantially ($\propto r_{ps}$, but only up to the heritability)
- More embryos will not ($\propto \sqrt{\log n}$)
- Actual gain uncertain and practically limited

Theory

Cell

Screening Human Embryos for Polygenic Traits Has Limited Utility

Ehud Karavani,[1,18] Or Zuk,[2,18] Danny Zeevi,[3] Nir Barzilai,[4,5] Nikos C. Stefanis,[6,7,8] Alex Hatzimanolis,[6,8] Nikolaos Smyrnis,[6,7] Dimitrios Avramopoulos,[9,10] Leonid Kruglyak,[3,11,12] Gil Atzmon,[4,5,13] Max Lam,[14,15,16] Todd Lencz,[14,15,17,*] and Shai Carmi[1,19,*]

Acknowledgements:
Ehud Karavani, HUJI
Or Zuk, HUJI
Danny Zeevi, UCLA
Todd Lencz, Northwell

Abisch-Frenkel Foundation