Augustine Kong

Heritability Estimates and Genetic Nurture

University of Oxford

In the 2008 meeting, we asked: Is heritability being overestimated?

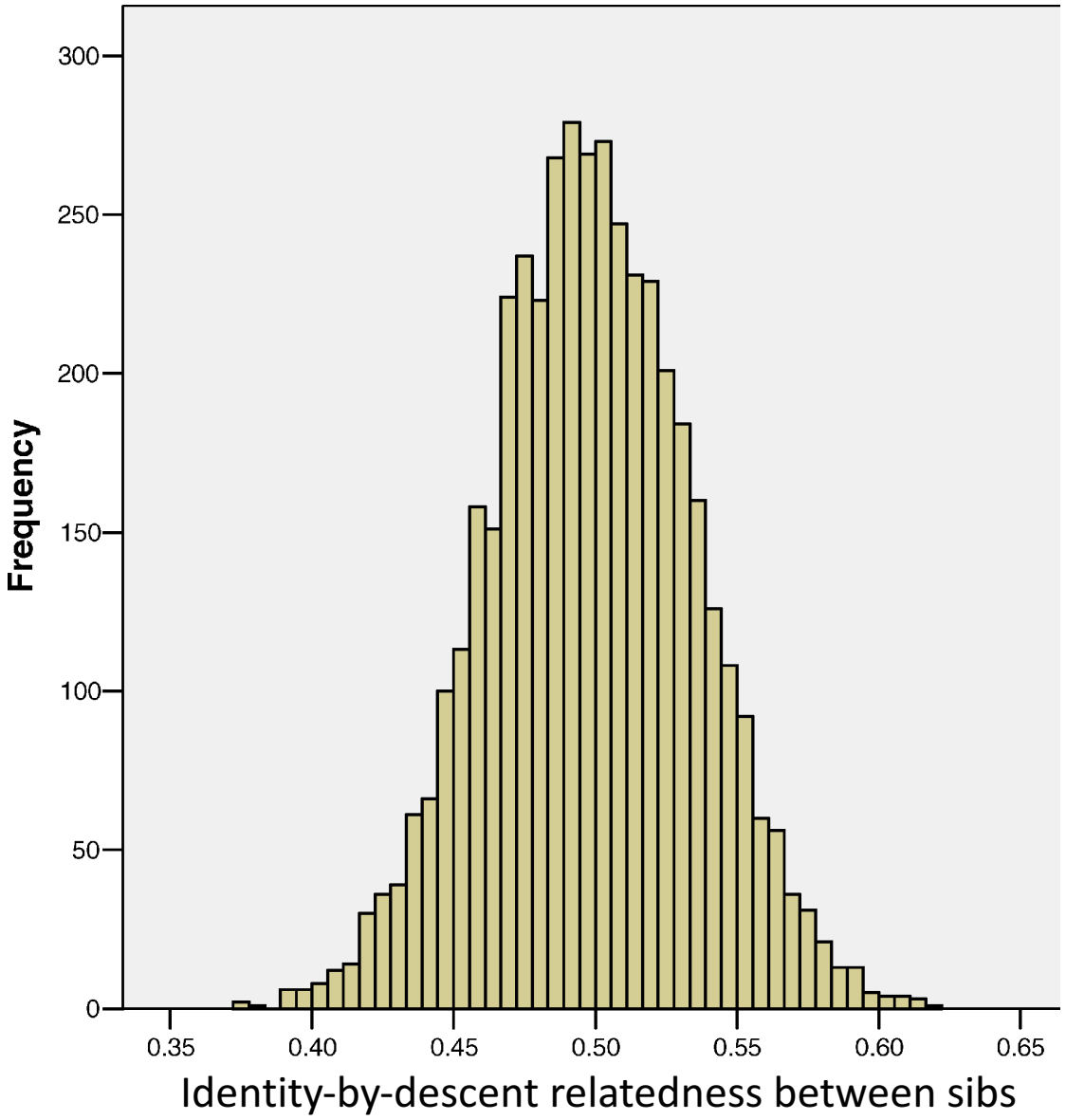
Sib-Regression for height heritability

Study	# sib-pairs	estimate	95% CI
Visscher, …, Martin (2006) PLoS Genetics*	3,375	0.80	0.46-0.85
Visscher, …, Martin (2007) AJHG	11,214	0.86	0.49-0.95
Hemani, …, Visscher (2013) AJHG	20,240	0.69	0.42-0.96
Young, …, Stefansson, Kong (2017) bioRxiv**	64,874	0.68	0.49-0.87

*result cited in Manolio, ..., Visscher (2009) Finding the missing heritability of complex diseases. Nature.

**Icelandic data from deCODE genetics.

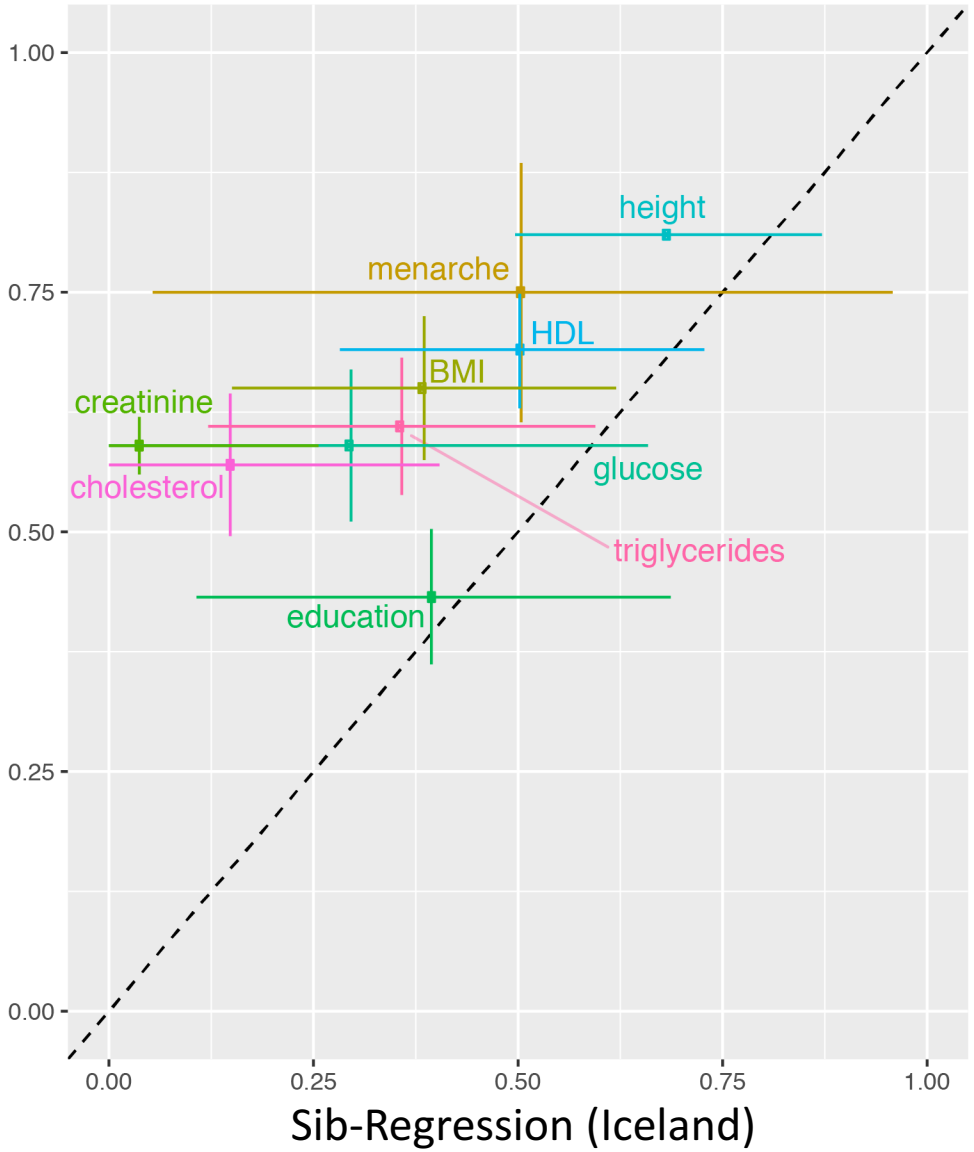
Sib-Regression



Taken from Visscher, …, Martin (2006) Assumption-free estimation of heritability from genome-wide identity-by-descent sharing between full siblings. PLoS Genetic.

Sib-Regression for traits in Iceland

Scandinavian Twin Studies

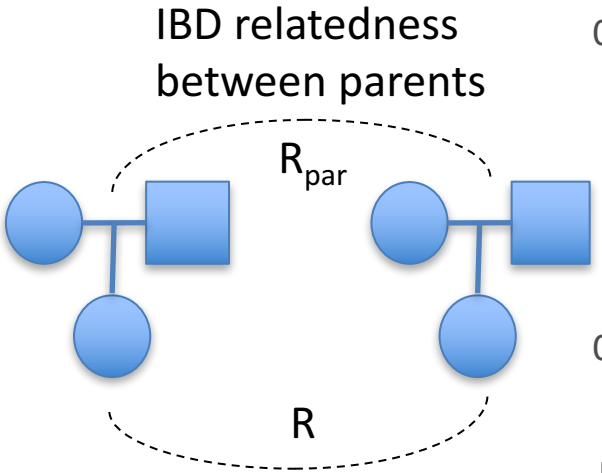


Results suggest that twin estimates could overestimate heritability in the general population (at least in Iceland)

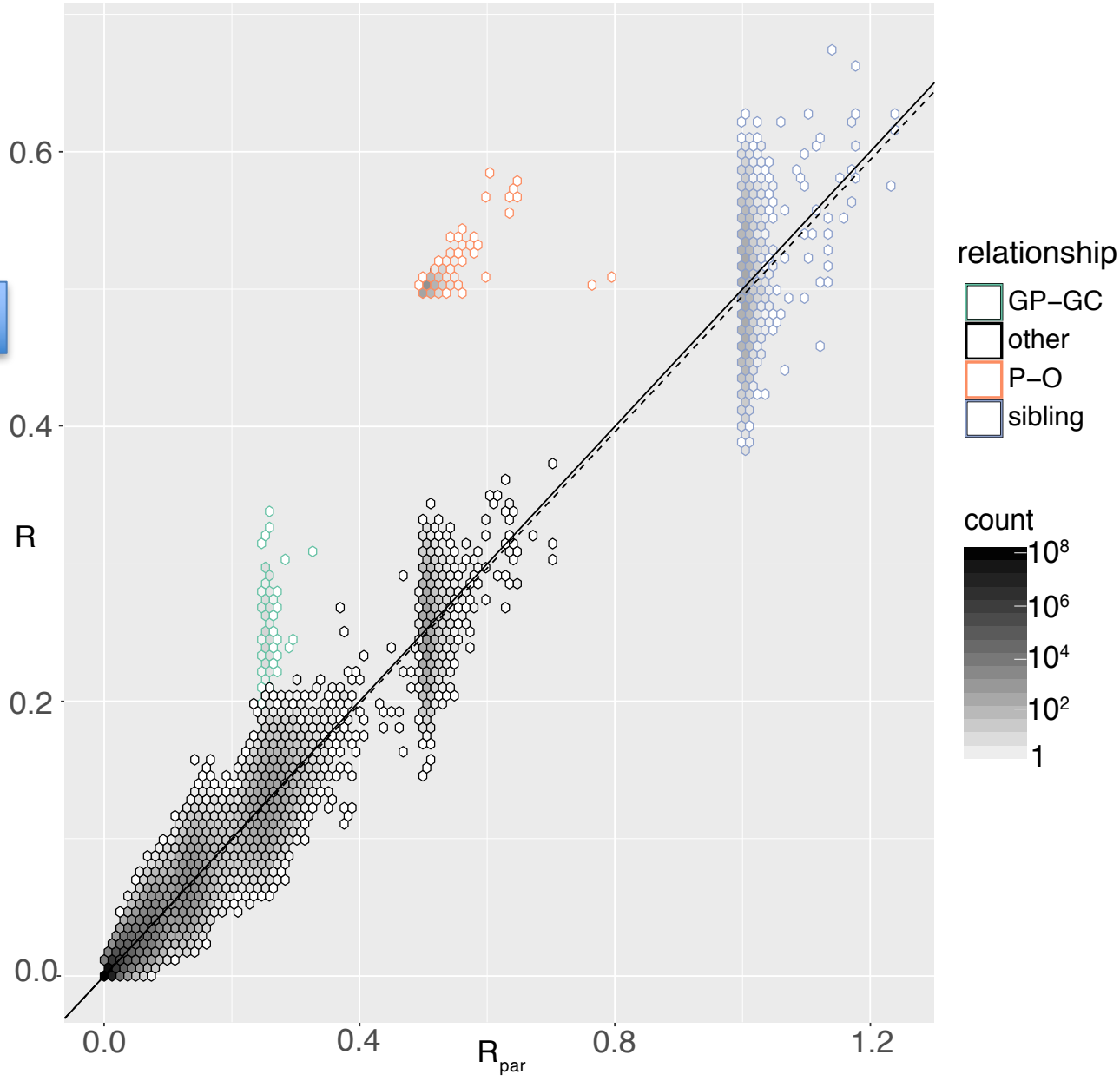
Mean difference Twin-Sib: 26%.

Twin estimate higher ($p<0.05$) for BMI, Total Chol., Triglycerides, and Creatinine.

Relatedness Disequilibrium Regression (RDR)



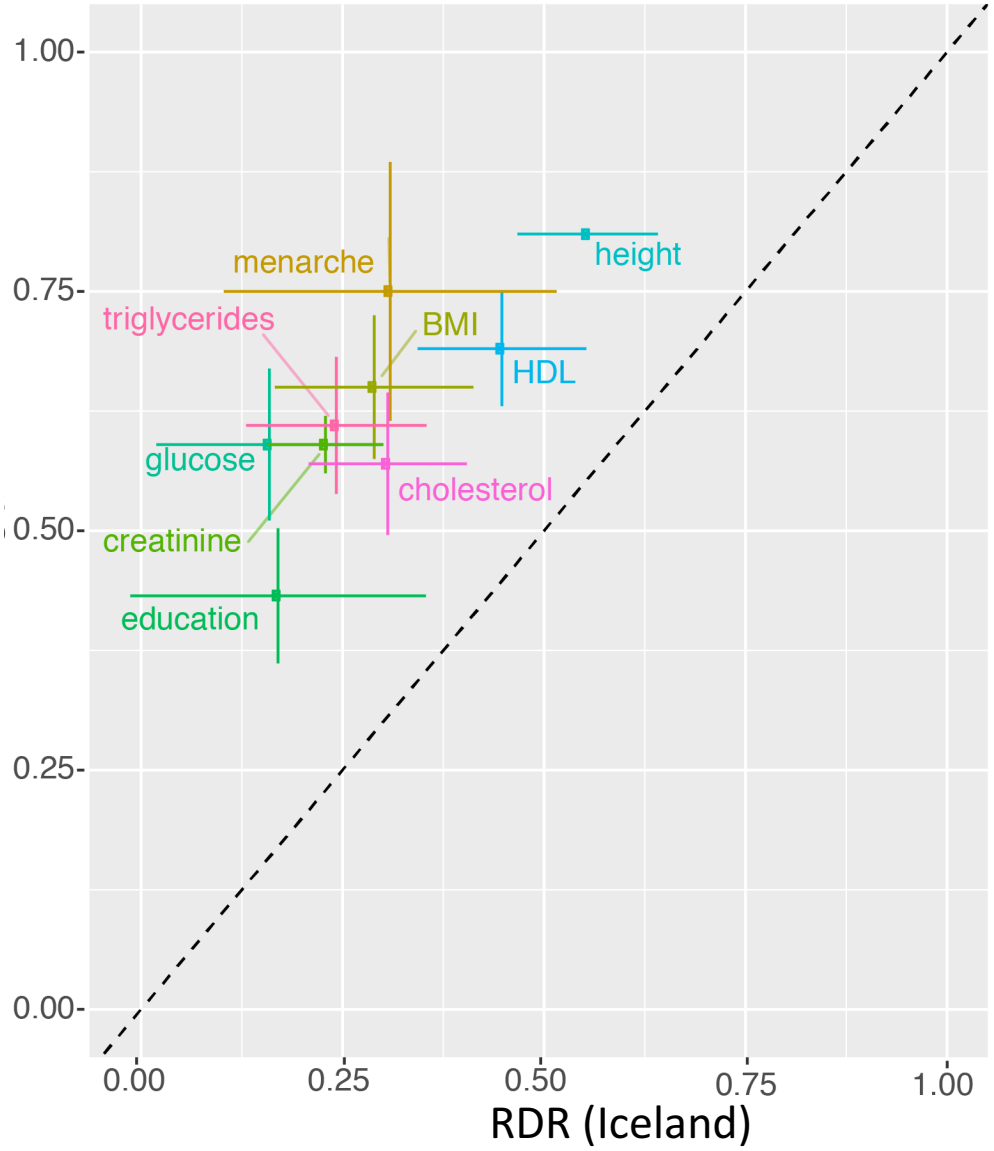
For two probands, RDR uses the IBD relatedness between the two pairs of parents as baseline/control for the IBD relatedness between the probands



Young, .., Stefansson, Kong (2017) Estimating heritability without environmental bias. bioRxiv

RDR for traits in Iceland

Scandinavian Twin Studies



If these RDR estimates are to be believed, this is evidence that the Scandinavian Twin estimates tend to be too high when applied to the general population of Iceland.

Mean difference Twin-RDR: 33%.
Twin estimate higher ($p<0.05$) for all traits.

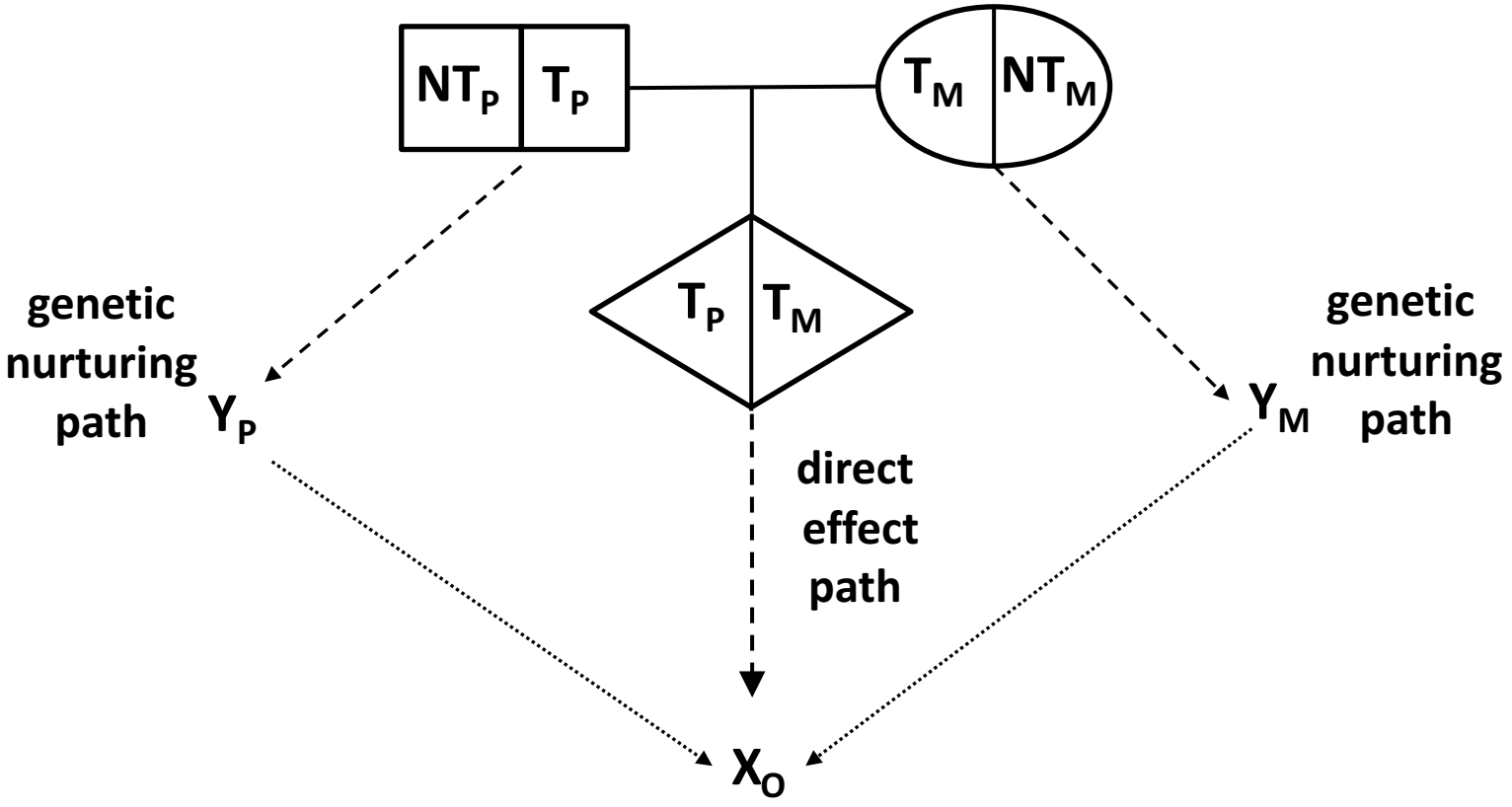
RDR and missing polygenic heritability

Trait	h^2_{RDR}	h^2_{poly}	$h^2_{\mathrm{RDR}} - h^2_{\mathrm{poly}}$	$1 - h^2_{\mathrm{poly}}/h^2_{\mathrm{RDR}}$
height	55% (SE 4%)	16.2%	39.2%	71%
BMI	29% (SE 6%)	9.7%	19.2%	66%
educational attainment (years)	17% (SE 9%)	2.5%	14.5%	85%

- h^2_{RDR} is the RDR heritability estimate from Icelandic data
- h^2_{poly} is the heritability explained by the direct effect of a polygenic score in Iceland.

Kong, ……, Stefansson. The Nature of Nurture (Jan, 2018) *Science*

Genetic Nurture --- A form of Indirect Genetic Effect



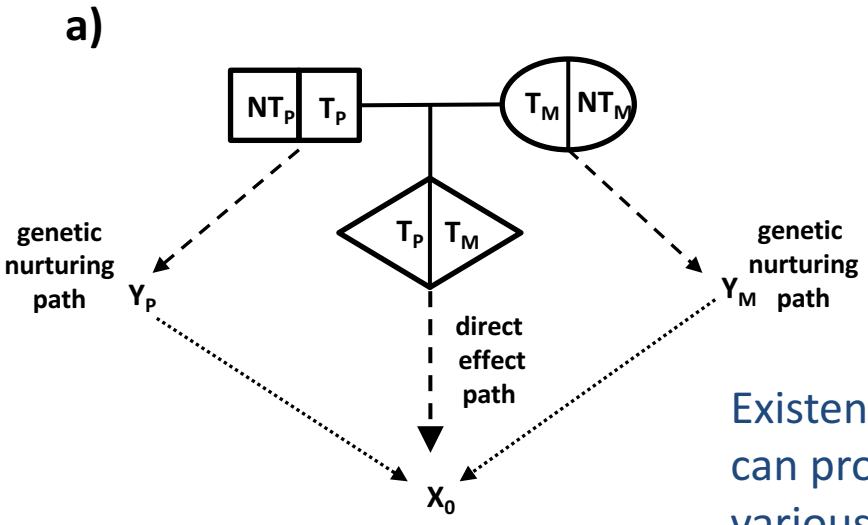
Non-transmitted (NT) alleles only have nurturing effects and transmitted (T) alleles have both direct and nurturing effects. Thus, basic GWAS effect estimates would tend to be overestimates of the direct effects when there is genetic nurturing.

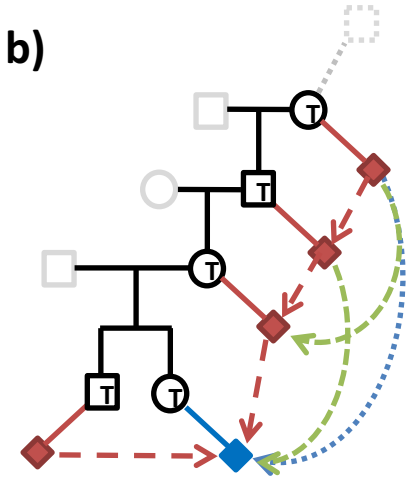
Decomposition of the observed effect of the EA polygenic score into direct, genetic nurturing, and confounding effects (Table from Kong, …., Stefansson 2018, *Science*)

				Transmitted			Nontransmitted						
				T (T = T_P + T_M)			NT (NT = NT_P + NT_M)						
Trait	N	N_{NTP}	N_{NTM}	$\widehat{\theta}_T$	*P*	R^2 (%)	$\widehat{\theta}_{NT}$	*P*	R^2_δ (%)	$\widehat{\delta}$ / $\widehat{\theta}_T$	$\widehat{\phi}_\delta$ / $\widehat{\theta}_T$	$\widehat{\eta}$ / $\widehat{\theta}_T$	$\widehat{\phi}_\eta$ / $\widehat{\theta}_T$
EA	21637	13948	19012	0.223	1.6×10^{-174}	4.98	0.067	1.6×10^{-14}	2.45	0.701	0.046	0.224	0.029
AGFC	54372	35294	47052	0.108	9.7×10^{-110}	1.17	0.039	2.9×10^{-13}	0.48	0.640	0.052	0.264	0.043
HDL	46872	30855	40788	0.065	9.0×10^{-29}	0.42	0.027	6.0×10^{-6}	0.14	0.586	0.046	0.319	0.050
BMI	39078	26433	34533	-0.060	1.0×10^{-22}	0.36	-0.017	0.0077	0.19	0.718	0.055	0.197	0.030
FG	34767	22959	30222	-0.051	7.6×10^{-18}	0.26	-0.018	0.0059	0.11	0.655	0.052	0.252	0.040
HT	39270	26563	34703	0.052	6.6×10^{-14}	0.28	0.030	1.5×10^{-5}	0.05	0.422	0.031	0.476	0.071
CPD	18887	12371	16589	-0.055	1.4×10^{-12}	0.31	-0.030	5.3×10^{-4}	0.06	0.461	0.035	0.439	0.066
HLTH	62328	41996	54546	0.082	2.7×10^{-60}	0.67	0.033	8.9×10^{-11}	0.23	0.592	0.051	0.305	0.052

The notion that for Educational Attainment (EA) the variance explained by the polygenic score is magnified by a substantial genetic nurturing effect is further supported by GWAS studies of EA (*Science* 2013, *Nature* 2016, ….) which noticed that within-family variance (from sib-pairs) explained by the polygenic score is substantially smaller than the usual variance explained.

There is clearly also a genetic nurturing effect for height (HT), although its magnitude is much smaller. Indeed, variants that are genome-wide significant in a HT GWAS are probably mostly 'height variants'. However, the genetic nurturing effects on HT of the EA-related variants would be captured by a polygenic score constructed from a HT GWAS -- and by GREML heritability estimates. These effects are small individually, but could accumulate to a meaningful amount. This would also have an effect on the apparent effect size distribution.





Genetic nurture
is not only manifested
through parents.

Existence of genetic nurture can profoundly affect how various heritability estimates should be the interpreted, e.g. GREML estimates would unavoidably also capture the genetic nurturing effects.

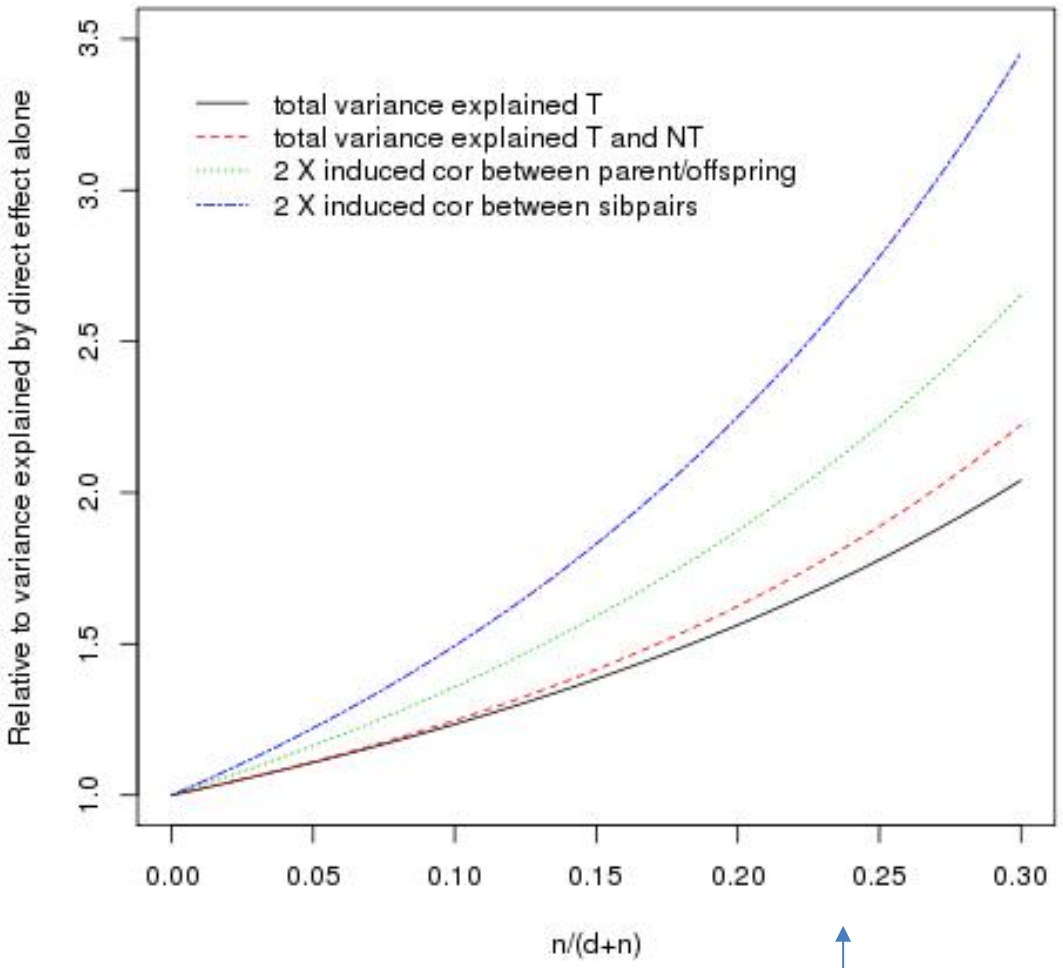
Twin estimates and Sib-regression are not affected by parental/ancestral genetic nurture, but can be biased due to genetic nurture from siblings.

Genetic nurture of EA from siblings: $P = 0.015$

Summary

- Heritability estimates based on twins, for whatever reason, appear to be too high for the general population. Sib-regression has its appeal, but requires very large sample size. The RDR method might work well for probands with parents also genotyped. RDR and sib-regression can complement each other.
- Genetic nurture can lead to positive bias of both effect estimates and heritability estimates from GREML. If 'explained heritability' only count GW significant markers and GREML heritability estimates are used, this could inflate missing heritability for many health related traits.
- The genetic components of EA and BMI are estimated to have a correlation of -0.13 (Bulik-Sullivan et al *NG* 2015). A part of that could be shared genetic nurturing components.

Magnification of the direct effect through nurturing



Indirect genetic effects from relatives bias 'GREML' estimates

Effect	Contribution to GREML	Contribution to RDR
Direct (δ)	δ^2	δ^2
Genetic nurturing (η)	η^2	0
Cov between direct & nurturing	$2\delta\eta$	0
Total	$(\delta+\eta)^2$	δ^2

Trait	$h^2_{\mathrm{RELT-SNP}}/h^2_{\mathrm{RDR-SNP}}$	Average η/δ
Educational attainment	1.69 (>1 with p=0.027)	0.30
Age at first child (women)	1.72 (>1 with $p=7.6\times10^{-3}$)	0.31

RELT-SNP, a method closely related to GREML, was applied to 'unrelated' pairs to estimate the heritability explained by SNPs in Iceland.