A family-based method for leveraging random genetic variation to identify variance-controlling loci

Dalton Conley [1]

Rebecca Johnson [1]

Ben Domingue [2]

Christopher Dawes [3]

Jason Boardman [4]

Mark Siegal [5]

*corresponding author
Email: dconley@princeton.edu

[1] Department of Sociology, Princeton University
[2] Graduate School of Education, Stanford University
[3] Wilff Family Department of Politics, New York University
[4] Institute for Behavioral Sciences, University of Colorado, Boulder
[5] Center for Genomics and Systems Biology, New York University

Abstract

The propensity of a trait to vary within a population may have evolutionary, ecological or clinical significance. In the present study we deploy sibling models to offer a novel and unbiased—if underpowered—way to ascertain loci associated with the extent to which phenotypes vary (variance-controlling quantitative trait loci, or vQTL). Previous methods for vQTL mapping either exclude genetically related individuals or treat genetic relatedness among individuals as a complicating factor addressed by adjusting estimates for non-independence in phenotypes. The present method uses genetic relatedness as a tool to obtain unbiased estimates of variance effects rather than as a nuisance. The family-based approach, which utilizes random variation between siblings in minor allele counts at a locus, also allows controls for parental genotype, mean effects, and non-linear (dominance) effects that may spuriously appear to generate variation. Deploying this strict approach on data from the Framingham Heart Study we found several alleles that appeared to alter within-family variation in height as well as pathways that appear to be enriched. However, significant SNPs failed to replicate in an independent sample. Pathway analysis revealed one gene set, encoding members of several signaling pathways related to gap junction function, which appears significantly enriched for associations with within-family height variation in both datasets (while not enriched in analysis of mean levels). We recommend approximating laboratory random assignment of genotype using complete family data and more careful attention to the possible conflation of mean and variance effects. Because of the lower power of family-based analyses, we further advocate the creation of consortia that unite studies with family-based designs, thereby increasing the sample sizes available for such analyses.

Introduction

The extent to which a complex trait varies in a population is a product of mutation, genetic drift and natural selection, as well as environmental variation and its interaction with genotype. Moreover, genotypes may differ in their propensities to vary. That is, particular genotypes might be more sensitive than others to changes in the environment or to the effects of new mutations. Sensitivity to the environment could come in the form of phenotypic plasticity, whereby stereotyped phenotypic changes occur under particular circumstances, or in the form of developmental instability, whereby random fluctuations in the internal or external environment lead to different phenotypic outcomes [1–3].

Sensitivity to the effects of mutations is related to cryptic genetic variation. In a population in which individuals are relatively insensitive to the effects of mutations, allelic variation may accumulate, while not presenting phenotypic effects. Replacing an allele that confers less sensitivity to mutational effects with one that confers more sensitivity (or introducing an environmental perturbation with similar effect) will cause the accumulated variation to have phenotypic consequences [4,5]. This release of cryptic genetic variation might have major implications for the adaptation of populations to environmental change, as well as for the genetics of human complex traits [6–10]. Indeed, it has been proposed that release of cryptic genetic variation might be responsible for the increased prevalence of human "diseases of modernity" such as diabetes [6].

Model organisms have been used to identify genes that control sensitivity to mutations or the environment [4,11]. Two approaches have been taken. One approach is to test for differences in phenotypic variance between wild-type and mutant strains. For example, the molecular chaperone Hsp90, when impaired, reveals cryptic genetic variation in the fly *Drosophila*

3

melanogaster, the flowering plant *Arabidopsis thaliana*, the fish *Danio rerio* and the budding yeast *Saccharomyces cerevisiae* [12–15]. Other genes appear to affect the expression of cryptic genetic variation as well [16,17].

Evidence on whether Hsp90 controls environmental sensitivity is mixed [17–19]. However, screens of the *S. cerevisiae* genome identified hundreds of genes that, when mutated, cause increased variation in cell morphology among isogenic cells raised in the same environment. That is, these mutations increased sensitivity to fluctuations in the internal or external environment [20,21]. A test of one of these genes, which showed a major increase in environmental sensitivity upon deletion, revealed a high degree of epistasis with new mutations but no net increase in mutational sensitivity upon deletion [22]. A similar result was found for impairment of Hsp90 [23]. In general, the relationship between suppression of the effects of environmental variation and suppression of the effects of mutational variation is unclear [4,11].

The second approach to identifying variance-controlling genes is to use linkage or association analysis to map natural genetic variants that confer differential sensitivity. Such searches for variance-controlling quantitative trait loci (vQTL) have been conducted to identify determinants of microenvironmental sensitivity (i.e., inherent stochasticity or sensitivity to fluctuations within a nominally constant environment) and/or phenotypic plasticity (variation across controlled environments) of various phenotypes, including morphological and life-history traits as well as expression levels of individual genes, in a range of organisms including humans [24–42]. Natural genetic variation affecting sensitivity to other segregating alleles has been studied as well, and indeed there are natural variants of the Hsp90 gene that appear to reveal cryptic variation [43–45].

4

The first study claiming to map a locus that controls variance of a human trait examined body mass index (BMI) using data from 38 cohorts that participate in the GIANT consortium for genome-wide analysis (GWA) of single-nucleotide polymorphisms (SNPs) that affect human height [46] (although a prior study had examined variance in order to prioritize the search for gene-by-environment and gene-by-gene interaction effects [47]). Yang et al. [46] computed a Z-score (inverse-normal transformation) for BMI for each of 133,154 individuals then performed GWA on the squared Z-score. This squared Z-score captures the magnitude of each individual's deviation from the mean phenotype and therefore is meant to act as an individual-based measure of variance. In this discovery sample they found SNPs in the *FTO* gene and the *RCOR1* gene that appear to control variation in human BMI. In a replication sample of 36,327 individuals from 13 cohorts, one SNP in the *FTO* gene was confirmed [46].

Later research pointed to several challenges inherent in using association analyses to identify vQTL.The first issue is mean-variance confounding: a given locus can have: 1) effects on mean levels of a trait (mean effects), 2) effects on variance in a trait (variance effects), or 3) effects on both mean levels of a trait *and* variance in a trait. For association analyses that aim to isolate variance effects, it is critical to check that any detected effects on "variance" are not merely statistical artifacts of mean effects. In particular, this conflation of variance and mean effects was a concern in the analysis of Yang et al. [46] because normalized variation scores will tend to be higher for populations with higher mean levels and because *FTO* is one of the most well-established genes that affects the mean level of BMI in human populations [48–56]. Research examining MMP3 protein levels in cerebrospinal fluid found similar overlap, where SNPs in linkage disequilibrium with a locus well-established in predicting mean levels of the

5

trait (rs679620 of the *MMP3* gene) were associated with both higher mean levels and higher variance in the trait [57].

Yang et al. [46] addressed the effect of mean BMI by showing that there is no global correlation between mean effects of SNPs and their variance effects. However, it must be considered that a lack of correlation between mean and variation effects across the genome could be caused merely by the fact that the vast majority of SNPs have negligible or nonexistent effects on both mean and variation. For this reason, in the present study, we test only the top-scoring SNPs. We show that a significant confounding between mean and variance effects does in fact exist.

In addition to mean-variance confounding, methods to investigate vQTL must address several other issues, some of which are shared with the estimation of mean effects and others of which, such as mean-variance confounding, are unique to or particularly acute in the case of variance effects [39]. Methods for vQTL analysis beyond the squared Z-score method can be grouped into: 1) non-parametric methods that test whether the three genotypes at a biallelic locus (minor-allele homozygote, heterozygote, and major-allele homozygote) have equal or unequal variance, and 2) parametric methods that relax the assumption in GWA linear regressions that the residual error is identically distributed across all genotypes.

Non-parametric methods include Levene's test, which uses a test statistic derived from the squared difference between an individual's level of a trait and the genotype-level mean or median, the latter of which is used to make the method more robust to a non-normally distributed trait [47,58] and the Fligner-Killeen (FK) test, which is similar to Levene's test, in that it uses the absolute difference between an individual's level of a trait and the genotype-level median, but then computes the test statistic based on ranks of these differences. The FK test can be used

either as a standalone test for variance effects [24], or as the test statistic for the scale component of the Lepage or other joint scale-location test [59]. More recently proposed non-parametric tests consider not only the variance of the trait distribution but other features as well, such as skew [60,61].

The main drawback of non-parametric tests for detecting vQTL's is that the tests, which compute differences between discrete genotype groups, cannot directly control for important covariates, such as age, sex and population structure. Some adopt a two-stage regression procedure for including covariate controls: first the trait is regressed on covariates, then the variance test of interest is performed on the residualized dependent variable [61,62]. However, two-stage procedures have been shown in simulations to reduce power and induce bias [57]. The problem of how to include covariates in a way that does not induce bias is acute because of the importance of two types of controls when investigating variance-controlling loci: control for population stratification and controls for nonrandom association between genotype and environment. The former control is needed in all GWA analyses to separate the effect of any particular locus from the effects of all other loci shared by virtue of common ancestry. The latter control might be particularly relevant to vQTL analyses because mean effects of genotypes might impact the environment that is experienced, which might in turn impact variance [63].

Parametric approaches that use generalized linear models to jointly estimate the mean and variance of a trait address this problem [57,64,65]; these models can include controls for population stratification as well as controls for *observed* covariates that influence genetic distribution into variance-affecting environments. The double generalized linear model (DGLM) approach begins with the typical linear model for estimating mean effects where the residual variance is the same across genotypes, then the model is relaxed to allow residual variance to

differ by genotype and to incorporate non-genetic covariates that might contribute to residual variance; fitting the two models alternates until convergence [64]. DGLM thus allows joint estimation of mean effects and variance effects, attempting to address mean-variance confounding, and permits controls for population stratification directly in the model. The main drawback is that DGLM is sensitive to violations of the assumption that the trait value is normally distributed, and simulations [57] show that attempts to address this violation by transforming the dependent variable (e.g., by Box-Cox transformation [64]) lead to the high type I error rates that others note are of particular concern in vQTL analyses due to right-skewness in trait variance [66]. This drawback might be alleviated by a parametric bootstrap-based likelihood ratio test for choosing among models with and without mean and variance effects [57]. Bayesian approaches, which allow simultaneous estimation of mean and variance effects as well as inclusion of covariates, also show promise [67].

The covariate of population stratification is typically controlled in these methods by inclusion of principal components of the sample population's genotypes among the vectors of covariates predicting the mean and/or residual variance of a trait. This control is especially important for traits, such as BMI, that are expected to show considerable environment-dependence. Especially when pooling such traits across cohorts, there is a risk that systematic differences in environment correlate regionally with systematic differences in genetic variation. That is, it is plausible, due to population stratification, that any genetic signal is merely acting as a proxy for culture and environment—a potential confounder that has been well-illustrated by the "chopsticks gene" example [68].

The other critical covariate that might bias vQTL estimates is genotype-environment correlation (rGE). Genotype-environment correlations may be caused by niche construction,

whereby individual organisms shape the environment (in a genotype-dependent way) [69–73]—dynamics we might expect to occur for phenotypes that have significant behavioral and environmental etiologies, such as BMI. As a result, genotypes may be associated with variance in BMI and other traits not through direct genetic effects but through interaction with alternative environments associated with variance such as more versus less sedentary lifestyles. An analogous situation that illustrates this potentially confounding genotype-by-environment interaction effect is that of caffeine consumption. A variant in a gene that encodes a caffeine-metabolizing enzyme can lead to greater variation with no effect on the mean through a mechanism of niche construction—i.e. individuals with the minor allele avoid coffee altogether, or if they are unable to do so, they end up drinking more than those with the major allele, thus leading to greater variance thanks to the coffee "environment" [63].

Regression approaches can control for *observed* covariates that are correlated with genotype and influence construction of variance-affecting environments. However, these methods cannot control for *unobserved differences* that produce these correlations. The present paper uses a family-based model to control for these unobserved differences. Two existing family-based models *allow* for investigations of variance-based loci in samples among related as opposed to unrelated individuals, but do not *leverage* the family-based structure of the data to control for unobserved confounders that vary between families [65,74]. First is a family-based version of the likelihood ratio test that adds a random effect meant to capture familial correlation in a trait. Although the family-level random effect helps control for unobserved variation *between* families that may influence variance in a trait through pathways other than genotype, the model relies on the strong assumption of independence between these unobserved features of family and the observed covariates. As a result, the model does not control for unobserved

characteristics that vary between families and that may be correlated with construction of variance-affecting environments. Simulation studies of random effects models show that substantial bias might therefore be introduced into the estimated effects if there are likely non-zero correlations between unobserved confounders, observed genotypes, and the trait of interest (BMI in the author's case) [75]. Similarly, a model that estimates the DGLM in a sample containing monozygotic and dizygotic twins, which has the advantage of isolating non-genetic from genetic sources of variance, does not control for unobserved features that vary between families and that affect construction of variance-affecting environments [74].

We offer an alternative methodology—comparisons of variation within sibling sets while controlling for parental genotype—that does not assume independence between observed covariates and unobserved between-family differences and that, as a result, better approximates random assignment of genotype in a laboratory. Utilizing a regression-based framework, the approach retains the advantages of the DGLM and Bayesian regression approaches: the ability to include covariates and control for mean effects when estimating variance effects through simultaneous estimation of parameters capturing both. The model uses sibling pairs as the unit of analysis and regresses the standard deviation of the sibling pair's trait on the pair's count of minor alleles with sibling pair-level controls that include controls for the mean level of the trait in the sibling pair, parental genotype, pair sex (MM or FM or FF), mean pair age, and the within-pair age difference (for the full model specification, as well as alternative specifications tested, see Methods). The parental genotype control means that the quantity of interest—the additive effect of each minor allele on variance in the trait—is identified solely from random between-sibship variation in minor allele counts. Likewise, the control for the mean level of the trait in a sibling pair aims to control for mean-variance conflation.

The proposed methodology, although restricted in applicability to datasets that enjoy a family-based design with quartets (both parents and at least two offspring), sacrifices statistical power and efficiency for improved consistency of parameters. This is a particularly important trade-off to make when we wish to rule out gene-environment correlations across populations as well as population stratification as alternative explanations to variance-locus associations. As noted above, this is especially critical for a phenotype such as BMI and a locus such as *FTO* given that environment and behavior (such as sedentariness) alter the *FTO*-BMI relationship and may vary significantly across cohorts/societies [76]. In cases where the goal is not to study control of variance per se, but instead is to probe the existence of gene-environment or gene-gene interactions in a way that avoids the high-dimensional parameter space problems of traditional approaches, the use of vQTL approaches with higher statistical power might be warranted.

In addition to power, another important feature of an approach to detect vQTL is the flexibility to capture non-linear effects of alleles, which the DGLM, the parametric bootstrap-based likelihood ratio test, and Bayesian regressions allow for by allowing genotypes to be specified using three indicator variables to capture non-linearities. The present family-based approach is potentially susceptible to confounding of variance effects by non-linear effects of alleles, because the association mapping is done on the sibship unit so the genotype is represented as the total number of major or minor alleles of each sib pair. Therefore, if dominance were at play with respect to mean levels, then this might generate spurious effects on variance [77]. That is, if among heterozygotes there was no effect of an allele on mean levels but among homozygotes there was, then this itself would generate apparent, but spurious, effects on variance even when controlling for linear mean effects. However, by comparing subgroups among sibship pairs that have two minor alleles (out of four possible in total), we are able to

check for this possibility. Specifically, by comparing those 2-minor allele pairs where both

siblings are heterozygotes (i.e. each individual has one minor allele) with those where both are

homozygotes (where one sibling has zero minor alleles and the other has two), we can rule out

this possible statistical artifact of non-linear effects on mean levels. Although the possibility of

non-linearities that do not reflect true variance effects can never be totally eliminated [77], this

approach guards against a primary form of non-linearity—dominance.

Below we report GWA results for variation effects on two phenotypes: height and BMI.

When we control for parental genotype, we are unable to replicate any genome-wide statistically

suggestive hits in the Framingham Heart Study (FHS) data, our discovery sample, in our

replication sample, the Minnesota Twin Family Study (MTFS). We also test whether our

potential variance-related alleles are merely reflecting dominance effects among heterozygotes;

we find no evidence for this. Finally, we perform gene-based and pathway enrichment analysis.

We do find one pathway, related to gap junction function, that is significantly enriched in both

our discovery and replication samples for associations with variance in height. We then discuss

the implications of our findings for prior and future research.

Results

When we use data from quartets in the FHS and control for mean sibling-pair height, sex,

mean pair age, within-pair age difference, and parental genotype in genome-wide regressions on

the standard deviation of the sibling pair height we find four SNPs that are genome-wide

suggestively significant ($p < 10^{-5}$) and meet other Hardy-Weinberg equilibrium (HWE) and minor-

allele frequency (MAF) controls (see Methods): rs2804263 (MAF 30.8%); rs2073302 (MAF

39.1 %); rs8126205 (MAF 37.1 %) and rs4834078 (MAF 24.0 %) (Fig 1A; for detailed

association plots and regional linkage maps of genome-wide suggestive hits, see S1 Fig and S2 Fig, respectively). For BMI we find no loci that meet even genome-wise suggestive significance (Fig 1B). Our method, as expected, controls well for population structure: QQ plots for the height and BMI p-values do not show the telltale "early liftoff" typical of failure to control this confounder (S3 Fig).

Fig 1: Manhattan plot of sibling variation in height and BMI among FHS 3rd generation sibling pairs. A) Results for the pairwise sibling standard deviation in height regressed against the sibling-pair average height with controls for sex of sibship, mean age of siblings, age difference of siblings and parental genotype. B) Same as in (A) except for BMI instead of height.

However, the absence of genome-wide significant SNPs suggests that the method's statistical power is low. A power analysis for the sample size utilized in the discovery dataset indeed supports this suggestion. For example, a single SNP would need to explain over 4.8% of the variation in the trait for the study to achieve 80% statistical power at the given sample size (S4 Fig). An effect size of this magnitude is not expected for human complex traits (e.g., the largest effect of a single SNP for mean height explains approximately 1% of the variation). Power is very low near $R^2 = 0.01$ (S4 Fig).

Apparent effects on variance can be generated if means and variances are correlated or if effects of alleles at a locus are not additive (i.e., there is dominance). Confounding of mean and variance effects is a serious issue with the present data. Although there is no overall correlation between effects on mean and on variance (as measured by squared Z-score and as previously reported [46]), the absence of the correlation appears to be caused by the fact that the vast majority of SNPs affect neither mean nor variance and therefore noise swamps out any signal; restricting the analysis to the SNPs that are the top hits reveals a very strong correlation between mean and variance effects for BMI (S5 Fig). Moreover, the top SNP for variance in BMI found

13

by Yang et al. [46] shows clear association with mean BMI in the same data set from the GIANT consortium (S6 Fig), whereas the four suggestively significant SNPs in our analysis do not show an association with mean height in the GIANT consortium data (S7 Fig). Further highlighting the importance of our regression-based control for sibling-pair mean are the observations that: 1) sibling-pair standard deviation shows a strong positive correlation with sibling-pair mean, and 2) this correlation is not eliminated by using the sibling-pair coefficient of variation rather than the standard deviation (S8 Fig).

Our approach does not control for non-additivity of allele effects at a locus, as it assumes a linear model. However, it does allow a test of whether an effect on variation net of mean was actually an artifact of non-linear effects on average rather than an actual variance effect. If the true relationship between phenotype and MAF were non-linear (i.e. revealed dominance effects) our initial findings could be entirely driven by divergence among those sibling pairs with two minor alleles. For example, if an individual with two minor alleles were significantly taller than an individual with either one or zero minor alleles (recessive effect) then when we collapsed the sibling pairs with two minor alleles, we could generate artifactual variation effects because among those sibships with two minor alleles, some would be distributed 0-2 (and thus one sibling would be taller than the other) while other sibships would be 1-1 (and thus would be the same height). Put together, it would appear that two minor alleles increased the variation net of mean effects. And if strong enough, such a misspecified effect could exert enough leverage to make a linear effect on variation appear across all allele numbers (zero to four for the sibship). These concerns appear not to apply to our analysis: there are minor differences in the variance effects among two-minor allele sibships by zygosity, but three out of the four are not statistically significant and the other is only weakly significant ($p= 0.048$) (Fig 2).

Fig 2: Test for spurious association with variance due to non-linear effects on mean levels. Sibling-pair mean height (inches) among pairs with two minor alleles is shown separately for homozygotes (one sibling with zero minor alleles and the other sibling with two) and heterozygotes (each sibling has one minor allele), for each genome-wide suggestively significant SNP.

For replication analysis, we used respondents from the MTFS. (For list of proxy MTFS SNPs with information on MAF and linkage with FHS SNPs see S1 Table.) These families included phenotypic and genotypic information on pairs of twins as well as their parents, allowing us to replicate the sibling-based analysis with parental genetic controls so as to mimic random assignment of alleles. The MTFS has both dizygotic (DZ) and monozygotic (MZ) twins. Because MZ sibships do not vary in terms of cryptic genetic variation and may experience much more similar environments to each other than do genetically distinct siblings, we also repeated our replication analysis only with DZ twin sets but found that exclusion of MZ twins did not affect results. Another concern is that twins (even DZ twins) may experience more similar environments than singleton siblings; thus, our replication analysis may suffer from attenuation bias to the extent that the cause of variation is environmental and not cryptic genetic differences (which should, by contrast, be equivalent for singleton full siblings and DZ twins). Among the SNPs that were genome-wide suggestively significant for the height analysis, only two of the four had viable proxy SNPs in the MTFS dataset after quality control: rs2804263 and rs4834078 both had proxies whereas rs2073302 and rs8126205 did not (S1 Table). When we ran the analysis for the proxy SNPs in the MTFS dataset, none achieved statistical significance.

In an attempt to gain power (at the expense of identifiability), we also performed two analyses that pool SNPs: gene-based and pathway-based pooling (see Methods). VEGAS for the gene analysis yielded no genes that were significantly enriched compared to null (in either

15

dataset). However, *i*-GSEA4GWAS did yield some pathways that appeared to be significantly enriched (Table 1). One of these pathways, associated with within-sibship variance in height in the FHS data, replicated in the MTFS data: HSA04540 Gap Junction (p=0.002 for each data set)(S9 Fig). HSA04540 includes members of several signaling pathways, including growth factors and their receptors, although any connection between these factors and organismal growth, as manifested in ultimate height, remains to be determined. Importantly, this pathway was not significant in the GWA for mean levels effects in either dataset.

Table 1: Enriched canonical pathways for height and BMI sibling-pair standard deviations in FHS

Pathway/Gene set name	P-value	FDR	Sign. Genes/ Slctd. Genes/ All Genes
HEIGHT			
GLYCEROLIPID METABOLISM	< 0.001	0.006	20/35/45
HSA00561 GLYCEROLIPID METABOLISM	< 0.001	0.009	27/49/58
VEGFPATHWAY	< 0.001	0.015666666	10/20/28
HSA03030 DNA POLYMERASE	0.001	0.0276	9/20/25
BILE ACID BIOSYNTHESIS	0.003	0.03175	11/21/27
HSA04540 GAP JUNCTION *	**0.002**	**0.041666668**	**42/78/98**
HSA05217 BASAL CELL CARCINOMA	0.004	0.106	25/46/56
HSA02010 ABC TRANSPORTERS GENERAL	0.006	0.123875	20/40/44
HSA00564 GLYCEROPHOSPHOLIPID METABOLISM	0.005	0.24111111	20/49/68
BMI			
HSA00591 LINOLEIC ACID METABOLISM	< 0.001	0.006	11/25/31

HSA00020 CITRATE CYCLE	< 0.001	0.1145	12/22/29
HSA04520 ADHERENS JUNCTION	0.005	0.2392	37/67/75

*Replicates in MTFS data

Discussion

Our analysis extends earlier work that aimed to map variance-controlling loci in humans [46]. Although the prior work enjoyed greater statistical power, it also had more potential for bias—due both to environmental confounding and to conflation of mean and variance effects. Indeed, Yang et al. [46] identified a locus regulating BMI variability that is also strongly associated with mean levels and for which a gene-by-environment interaction effect on mean has been shown. In the present study we were able to perform within-family analysis on two samples of white Americans, completely free of population stratification, largely devoid of rGE confounding, and with controls for mean level effects as well as checks for non-linear (i.e. dominance) effects on mean levels. Like the latest methods to map variance-controlling loci in controlled crosses [64], our approach therefore avoids common confounds. At the same time it overcomes problems specifically associated with human traits, including the construction of variance-affecting environments, that existing regression-based methods for detecting vQTLs fail to address because they allow controls solely for observed confounders (e.g., [57,64,65,78]).

Though underpowered our results strongly support the benefits of approximating a randomized genetic experiment by analyzing within-family variation while controlling for parental genotype. Such an analysis addresses the possibility that it is merely cross-family environments interacting with a mean effect and/or population structure that produce apparent association with variability. Meanwhile, parameterizing the estimand as spread (SD) net of sibship mean levels provides a robust, flexible way to conceive of variation—that is, rather than parameterizing the relationship between mean and variance a priori by using the coefficient of variation or some similar summary statistic. The trade-off inherent to our approach is that environmental and phenotypic variation within sibships may be attenuated, reducing statistical

leverage; the extent of such a dynamic is wholly dependent on phenotype, of course. Although the within and between family components of the variation in the phenotype can be measured to determine whether or not the phenotype is suitable for such an approach, the extent of variation within and between sibships in the unmeasured environmental factors that matter is, of course, unknown.

In light of this discussion, we think that there is benefit to combining prior, pedigree-based approaches with newer GWAS methods to better estimate variance effects (as well as levels effects). Thus, we recommend that consortia of cohorts with genome-wide data on full pedigrees (quartets at a minimum) be formed to advance GWA to a more solid foundation of inference that approximates the unbiased estimates of lab-based genetic manipulations by taking advantage of random differences in sibling genotypes.

Materials and Methods

Data

Data for discovery analysis come from the Framingham Heart Study (FHS), second (parental) and third (sibling) generation respondents. (This dataset is publicly available through dbGaP http://www.ncbi.nlm.nih.gov/gap. QC code can be obtained from the FHS investigators [79].) The FHS is, in fact, one of the cohorts included in the GIANT meta-analysis performed by Yang et al. [46]. Height and weight were taken from clinical measurements and then BMI was calculated as (weight in kilograms)^2 / height in meters. Genotypes were assayed using the Affymetrix GeneChip Human Mapping 500K Array and the 50K Human Gene Focused Panel. Genotypes were determined using the BRLMM algorithm. Our analysis began with the original 500,568 SNPs, and resulted in 260,469 SNPs available for analysis after cleaning (e.g., HWE

19

screens and a MAF cut-off of 0.05). The screens were conducted using all available individuals with genetic data, not only those that were included in this analysis. Genome-region association plots were produced using SNAP [80], except for those of published GIANT consortium data, which were produced using LocusZoom [81]. Regional linkage maps were produced using SNAP [80] and data from the 1,000 Genomes CEU Panel [82], which also provided the reference MAFs for S1 Table.

Among third-generation respondents, the numbers in our sample by sibship size are presented in Table 2. The 200 families with only one sibling in the data drop from the sibling analysis. Those with more than two contribute multiple pairs to the data; however, our final analysis selects only one pair per second-generation family as the more complicated error structure with multiple pairs leads to early takeoff on QQ plots.

_____

Table 2: Distribution of third generation siblings included in data by sibship size:

# 3rd-G Sibs in Family:	1	2	3	4	5	6	7	8	9
N (pairs):	--	292	483	504	250	150	189	28	36
Families:	200	292	161	84	25	10	9	1	1

Total actual N of sibling pairs after random selection of one per family: 583

_____

The Minnesota Twin Family Study (MTFS) replication data were genotyped on the Illumina 660W Quad array [84]. (DNA samples are stored at the National Institutes of Health's National Genetics Repository at Rutgers University in New Jersey. This dataset will be publicly available through dbGaP http://www.ncbi.nlm.nih.gov/gap. QC code can be obtained from the MTFS investigators.) Detailed information on genotypes [84] and phenotypes can be found

20

elsewhere [85]. Quality control procedures were applied separately to each individual cohort. Individuals with a call rate <0.95 (N=22), estimated inbreeding coefficient > 0.15 (N=2), or showing evidence of non-European descent from multidimensional scaling (N=298, mainly individuals with Mexican ancestry) were removed. Individuals were considered outlying from European descent if one or more of the first four eigenvectors were more than three standard deviations removed from the mean. SNPs with MAF < 0.01, call rate < 0.95 or HWE-test p-value <0.001 were removed. For our analysis, we included both the MZ and DZ twin pairs because restricting to DZ twin pairs that more closely approximate the sibships in the FHS discovery sample does not change the substantive findings. Because the discovery sample and the replication sample were genotyped on different arrays, we deployed SNAP to find corresponding SNPs [80]. The resulting sample size for our analysis was 1,048 pairs.

In addition to the GWA analysis in the discovery and replication sample, we performed two sets of analyses that pooled SNPs: gene-based pooling using VEGAS [86] and pathway-based pooling using *i*-GSEA4GWAS [87]. For the discovery sample, gene-based and pathway-based analyses were performed using 260,434 variants input; 239,526 variants used; 14,783 genes mapped; 221 gene sets selected. For the replication sample, gene-based and pathway-based analyses were performed using 522,726 variants input; 487,692 variants used; 16,840 genes mapped; 259 gene sets selected.

Statistical Analysis

All analysis was performed using R . The power analysis depicted in S4 Fig. to estimate the R^2 of the putative effect at varying power was performed using the *pwr* package in R [88] with n = observations in discovery sample and p = 10^{-5} (heteroscedasticity-robust standard errors should not substantially affect power under the present sample size [89]).

For the main analysis, sibling-pair standard deviations (SD) were fit by linear regression to the following model, where the key regressor is the number of minor alleles for the pair of siblings at a given locus. Because this number is for two individuals, the range is 0 to 4:

$$SD_i = \alpha \ + \beta \text{MAF count}_{ik} + \gamma Z_i + \ \epsilon_{ik}$$

In the model, i indexes a sibling pair, k indexes a SNP, MAF count is the total number of minor alleles in a sibling pair and Z_i is a vector of sibling pair-level controls that includes controls for the mean level of the trait in the sibling pair, parental genotype (coded as the parent count of minor alleles from 0 to 4), pair sex (MM or FM or FF), mean pair age, and the within-pair age difference. Qualitative results do not change if we instead specify the mother's and father's genotypes separately. Results also do not change when we normalize for minor allele frequency (MAF) using the procedure in [83], where i indexes participants, k indexes a snp, p_k refers to the MAF, and count refers to the count of minor alleles for sibling pair i at snp k:

$$\frac{count_{ik} - p_k}{\sqrt{2p_k(1 - p_k)}}$$

Huber-White standard errors robust to clustering on pedigree ID (to account for correlated errors among cousins: sibling pairs that share the same grandparents but not the same parents) were calculated for the FHS analysis in the following way, where i = sibling pair 1, j = sibling pair 2, n = total sibling pairs, g = pedigree grouping, and k = SNP.

For the simple case, where $\sigma^2 = \ \sigma_i^2 \ \forall i$ and $SST_x^2 = \ \sum_{i=1}^{n}(x_i - \bar{x})^2 \ :$

$$var(\hat{\beta}_k) = \frac{\sum_{i=1}^{n}(x_i - \bar{x})^2 \sigma^2}{SST_x^2}$$

To account for errors that may be heteroskedastic and correlated within a shared pedigree, we adjust the variance to be robust to cases where $\sigma^2 \neq \sigma_i^2$ and where $E[u_{ig} \ u_{jg'}] \neq 0$ for g = g' (two sibling pairs share same grandparent/pedigree ID). For the case of pedigree ID's with two

22

or more sibling pairs, this becomes the following variance robust to heteroskedastic and correlated errors, with an indicator function for when the sibling pairs belong to the same pedigree:

$$var(\hat{\beta}_k) = \frac{\sum_{j=1}^{n} \quad \sum_{i=1}^{n}(x_i - \bar{x})^2 (x_j - \bar{x})^2 \hat{u}_{ig} \hat{u}_{jg'} \mathbf{1}[g = g']}{SST_x^2}$$

For the case of pedigree ID's with one sibling pair only, the above equation reduces to the following variance robust to heteroskedastic errors [90]:

$$var(\hat{\beta}_k) = \frac{\sum_{i=1}^{n}(x_i - \bar{x})^2 \hat{u}_{ig}^2}{SST_x^2}$$

Acknowledgements

For helpful feedback on this work, we would like to thank members of the Siegal Lab at the NYU Center for Genomics and Systems Biology as well as Justin Blau, Richard Bonneau, and Michael Purugganan of the NYU Department of Biology and Tomas Kirchoff of NYU Medical School. The authors are also grateful to participants in the University of Colorado's 2012 workshop on Integrating Genetics and the Social Sciences for helpful comments.

References

1. Levy S, Siegal M. The robustness continuum. Adv Exp Med Biol. 2012;751: 431–452.

2. Dongen S V. Fluctuating asymmetry and developmental instability in evolutionary biology: past, present and future. J Evol Biol. 2006;19: 1727–43. doi:10.1111/j.1420-9101.2006.01175.x

3. Geiler-Samerotte KA, Bauer CR, Li S, Ziv N, Gresham D, Siegal ML. The details in the distributions: why and how to study phenotypic variability. Curr Opin Biotechnol. 2013;24: 752–759. doi:10.1016/j.copbio.2013.03.010

4. Masel J, Siegal ML. Robustness: mechanisms and consequences. Trends Genet TIG. 2009;25: 395–403. doi:10.1016/j.tig.2009.07.005

5. Gibson G, Dworkin I. Uncovering cryptic genetic variation. Nat Rev Genet. 2004;5: 681–90. doi:10.1038/nrg1426

6. Gibson G. Decanalization and the origin of complex disease. Nat Rev Genet. 2009;10: 131–140.

7. Wagner A. Robustness and evolvability: a paradox resolved. Proc Biol Sci. 2008;275: 91–100. doi:10.1098/rspb.2007.1137

8. Gibson G, Reed LK. Cryptic genetic variation. Curr Biol CB. 2008;18: R989-90. doi:10.1016/j.cub.2008.08.011

9. Berger D, Bauerfeind SS, Blanckenhorn WU, Schäfer MA. High temperatures reveal cryptic genetic variation in a polymorphic female sperm storage organ. Evol Int J Org Evol. 2011;65: 2830–42. doi:10.1111/j.1558-5646.2011.01392.x

10. Paaby AB, Rockman MV. Cryptic genetic variation: evolution's hidden substrate. Nat Rev Genet. 2014;15: 247–258. doi:10.1038/nrg3688

11. Siegal ML, Leu J-Y. On the Nature and Evolutionary Impact of Phenotypic Robustness Mechanisms. Annu Rev Ecol Evol Syst. 2014;45: 496–517. doi:10.1146/annurev-ecolsys-120213-091705

12. Queitsch C, Sangster T, Lindquist S. Hsp90 as a capacitor of phenotypic variation. Nature. 2002;417: 618–624.

13. Rutherford S, Lindquist S. Hsp90 as a capacitor for morphological evolution. Nature. 1998;396: 336–342.

14. Jarosz DF, Lindquist S. Hsp90 and environmental stress transform the adaptive value of natural genetic variation. Science. 2010;330: 1820–4. doi:10.1126/science.1195487

15. Yeyati P, Bancewicz R, Maule J, Heyningen V van. Hsp90 selectively modulates phenotype in vertebrate development. PLoS Genet. 2007;3: e43.

16. Takahashi KH. Multiple capacitors for natural genetic variation in Drosophila melanogaster. Mol Ecol. 2013;22: 1356–65. doi:10.1111/mec.12091

17. Siegal ML. Crouching variation revealed. Mol Ecol. 2013;22: 1187–9. doi:10.1111/mec.12195

18. Sangster TA, Salathia N, Undurraga S, Milo R, Schellenberg K, Lindquist S, et al. HSP90 affects the expression of genetic variation and developmental stability in quantitative traits. Proc Natl Acad Sci U S A. 2008;105: 2963–8. doi:10.1073/pnas.0712200105

19. Hsieh Y-Y, Hung P-H, Leu J-Y. Hsp90 regulates nongenetic variation in response to environmental stress. Mol Cell. 2013;50: 82–92. doi:10.1016/j.molcel.2013.01.026

20. Levy S, Siegal M. Network hubs buffer environmental variation in Saccharomyces cerevisiae. PLoS Biol. 2008;6: e264.

21. Bauer CR, Li S, Siegal ML. Essential gene disruptions reveal complex relationships between phenotypic robustness, pleiotropy, and fitness. Mol Syst Biol. 2015;11: 773.

22. Richardson JB, Uppendahl LD, Traficante MK, Levy SF, Siegal ML. Histone variant HTZ1 shows extensive epistasis with, but does not increase robustness to, new mutations. PLoS Genet. 2013;9: e1003733. doi:10.1371/journal.pgen.1003733

23. Geiler-Samerotte KA, Zhu YO, Goulet BE, Hall DW, Siegal ML. Selection Transforms the Landscape of Genetic Variation Interacting with Hsp90. PLoS Biol. 2016;14: e2000465. doi:10.1371/journal.pbio.2000465

24. Fraser H, Schadt E. The quantitative genetics of phenotypic robustness. PLoS One. 2010;5: e8635.

25. Stinchcombe JR, Caicedo AL, Hopkins R, Mays C, Boyd EW, Purugganan MD, et al. Vernalization sensitivity in Arabidopsis thaliana (Brassicaceae): the effects of latitude and FLC variation. Am J Bot. 2005;92: 1701–7. doi:10.3732/ajb.92.10.1701

26. Hall MC, Dworkin I, Ungerer MC, Purugganan M. Genetics of microenvironmental canalization in Arabidopsis thaliana. Proc Natl Acad Sci U S A. 2007;104: 13717–22. doi:10.1073/pnas.0701936104

27. Ansel J, Bottin H, Rodriguez-Beltran C, Damon C, Nagarajan M, Fehrmann S, et al. Cell-to-cell stochastic variation in gene expression is a complex genetic trait. PLoS Genet. 2008;4: e1000049. doi:10.1371/journal.pgen.1000049

28. Takahashi KH, Okada Y, Teramura K, Tsujino M. Deficiency mapping of the genomic regions associated with effects on developmental stability in Drosophila melanogaster. Evol Int J Org Evol. 2011;65: 3565–77. doi:10.1111/j.1558-5646.2011.01400.x

29. Takahashi KH, Okada Y, Teramura K. Genome-wide deficiency mapping of the regions responsible for temporal canalization of the developmental processes of Drosophila melanogaster. J Hered. 2011;102: 448–57. doi:10.1093/jhered/esr026

30. Takahashi KH, Okada Y, Teramura K. Deficiency screening for genomic regions with effects on environmental sensitivity of the sensory bristles of Drosophila melanogaster. Evol Int J Org Evol. 2012;66: 2878–90. doi:10.1111/j.1558-5646.2012.01636.x

31. Jimenez-Gomez JM, Corwin JA, Joseph B, Maloof JN, Kliebenstein DJ. Genomic analysis of QTLs and genes altering natural variation in stochastic noise. PLoS Genet. 2011;7: e1002295. doi:10.1371/journal.pgen.1002295

32. Shen X, Pettersson M, Rönnegård L, Carlborg Ö. Inheritance beyond plain heritability: variance-controlling genes in Arabidopsis thaliana. PLoS Genet. 2012;8: e1002839. doi:10.1371/journal.pgen.1002839

33. Janhunen M, Kause A, Vehviläinen H, Järvisalo O. Genetics of microenvironmental sensitivity of body weight in rainbow trout (Oncorhynchus mykiss) selected for improved growth. PloS One. 2012;7: e38766. doi:10.1371/journal.pone.0038766

34. Fehrmann S, Bottin-Duplus H, Leonidou A, Mollereau E, Barthelaix A, Wei W, et al. Natural sequence variants of yeast environmental sensors confer cell-to-cell expression variability. Mol Syst Biol. 2013;9: 695. doi:10.1038/msb.2013.53

35. Yvert G, Ohnuki S, Nogami S, Imanaga Y, Fehrmann S, Schacherer J, et al. Single-cell phenomics reveals intra-species variation of phenotypic noise in yeast. BMC Syst Biol. 2013;7: 54. doi:10.1186/1752-0509-7-54

36. Tonsor SJ, Elnaccash TW, Scheiner SM. Developmental instability is genetically correlated with phenotypic plasticity, constraining heritability, and fitness. Evol Int J Org Evol. 2013;67: 2923–35. doi:10.1111/evo.12175

37. Perry GML, Nehrke KW, Bushinsky DA, Reid R, Lewandowski KL, Hueber P, et al. Sex modifies genetic effects on residual variance in urinary calcium excretion in rat (Rattus norvegicus). Genetics. 2012;191: 1003–13. doi:10.1534/genetics.112.138909

38. Hulse AM, Cai JJ. Genetic variants contribute to gene expression variability in humans. Genetics. 2013;193: 95–108. doi:10.1534/genetics.112.146779

39. Rönnegård L, Valdar W. Recent developments in statistical methods for detecting genetic loci affecting phenotypic variability. BMC Genet. 2012;13: 63. doi:10.1186/1471-2156-13-63

40. Fraser HB, Schadt EE. The quantitative genetics of phenotypic robustness. PloS One. 2010;5: e8635. doi:10.1371/journal.pone.0008635

41. Fehrmann S, Bottin-Duplus H, Leonidou A, Mollereau E, Barthelaix A, Wei W, et al. Natural sequence variants of yeast environmental sensors confer cell-to-cell expression variability. Mol Syst Biol. 2013;9: 695. doi:10.1038/msb.2013.53

42. Ayroles JF, Buchanan SM, O'Leary C, Skutt-Kakaria K, Grenier JK, Clark AG, et al. Behavioral idiosyncrasy reveals genetic control of phenotypic variability. Proc Natl Acad Sci. 2015;112: 6706–6711. doi:10.1073/pnas.1503830112

43. Sgrò CM, Wegener B, Hoffmann AA. A naturally occurring variant of Hsp90 that is associated with decanalization. Proc Biol Sci. 2010;277: 2049–57. doi:10.1098/rspb.2010.0008

44. Siegal ML, Masel J. Hsp90 depletion goes wild. BMC Biol. 2012;10: 14. doi:10.1186/1741-7007-10-14

45. Chen B, Wagner A. Hsp90 is important for fecundity, longevity, and buffering of cryptic deleterious variation in wild fly populations. BMC Evol Biol. 2012;12: 25. doi:10.1186/1471-2148-12-25

46. Yang J, Loos R, Powell J, Medland SE, Speliotes EK, Chasman DI, et al. FTO genotype is associated with phenotypic variability of body mass index. Nature. 2012;490: 267–272.

47. Paré G, Cook NR, Ridker PM, Chasman DI. On the use of variance per genotype as a tool to identify quantitative trait interaction effects: a report from the Women's Genome Health Study. PLoS Genet. 2010;6: e1000981. doi:10.1371/journal.pgen.1000981

48. Scuteri A, Sanna S, Chen W, Uda M, Albai G, Strait J, et al. Genome-wide association scan shows genetic variants in the FTO gene are associated with obesity-related traits. PLoS Genet. 2007;3: e115.

49. Andreasen C, Stender-Petersen KL, Mogensen MS, Torekov SS, Wegner L, Andersen G, et al. Low physical activity accentuates the effect of the FTO rs9939609 polymorphism on body fat accumulation. Diabetes. 2008;57: 95–101.

50. Tönjes A, Zeggini E, Kovacs P, Böttcher Y, Schleinitz D, Dietrich K, et al. Association of FTO variants with BMI and fat mass in the self-contained population of Sorbs in Germany. Eur J Hum Genet. 2009;18: 104–110.

51. Haworth C, Carnell S, Meaburn E, Davis OS, Plomin R, Wardle J. Increasing heritability of BMI and stronger associations with the FTO gene over childhood. Obesity. 2008;16: 2663–2668.

52. Cha S, Choi S, Kim K, Park B, Kim J, Shin HD. Replication of genetic effects of FTO polymorphisms on BMI in a Korean population. Obesity. 2008;16: 2187–2189.

53. Hunt S, Stone S, Xin Y, Scherer C, Magness CL, Iadonato SP, et al. Association of the FTO gene with BMI. Obesity. 2008;16: 902–904.

54. Dina C, Meyre D, Gallina S, Durand E, Körner A, Jacobson P, et al. Variation in FTO contributes to childhood obesity and severe adult obesity. Nat Genet. 2007;36: 724–726.

55. Chang Y, Liu P, Lee W, Chang T, Jiang YD, Li HY, et al. Common variation in the fat mass and obesity-associated (FTO) gene confers risk of obesity and modulates BMI in the Chinese population. Diabetes. 2008;57: 2245–2252.

56. Frayling T, Timpson N, Weedon M, Zeggini E, Freathy RM, Lindgren CM, et al. A common variant in the FTO gene is associated with body mass index and predisposes to childhood and adult obesity. Science. 2007;316: 889–894.

57. Cao Y, Wei P, Bailey M, Kauwe JSK, Maxwell TJ, for the Alzheimer's Disease Neuroimaging Initiative. A Versatile Omnibus Test for Detecting Mean and Variance Heterogeneity. Genet Epidemiol. 2014;38: 51–59. doi:10.1002/gepi.21778

58. Struchalin MV, Dehghan A, Witteman JC, van Duijn C, Aulchenko YS. Variance heterogeneity analysis for detection of potentially interacting genetic loci: method and its limitations. BMC Genet. 2010;11: 92. doi:10.1186/1471-2156-11-92

59. Soave D, Corvol H, Panjwani N, Gong J, Li W, Boëlle P-Y, et al. A Joint Location-Scale Test Improves Power to Detect Associated SNPs, Gene Sets, and Pathways. Am J Hum Genet. 2015;97: 125–138. doi:10.1016/j.ajhg.2015.05.015

60. Aschard H, Zaitlen N, Tamimi RM, Lindström S, Kraft P. A non-parametric test to detect quantitative trait loci where the phenotypic distribution differs by genotypes. Genet Epidemiol. 2013;37: 323–333. doi:10.1002/gepi.21716

61. Hong C, Ning Y, Wei P, Cao Y, Chen Y. A semiparametric model for vQTL mapping. Biometrics. 2017;73: 571–581. doi:10.1111/biom.12612

62. Wei W-H, Bowes J, Plant D, Viatte S, Yarwood A, Massey J, et al. Major histocompatibility complex harbors widespread genotypic variability of non-additive risk of rheumatoid arthritis including epistasis. Sci Rep. 2016;6. doi:10.1038/srep25014

63. Caplin A, Cesarini D, Johannesson M, Thom K. Genes, Addiction, and Economics. Conf Biol Basis Prefer Behav Univ Chic. 2012;

64. Rönnegård L, Valdar W. Detecting major genetic loci controlling phenotypic variability in experimental crosses. Genetics. 2011;188: 435–447.

65. Cao Y, Maxwell TJ, Wei P. A Family-Based Joint Test for Mean and Variance Heterogeneity for Quantitative Traits. Ann Hum Genet. 2015;79: 46–56. doi:10.1111/ahg.12089

66. Shen X, Carlborg Ö. Beware of risk for increased false positive rates in genome-wide association studies for phenotypic variability. Front Genet. 2013;4. doi:10.3389/fgene.2013.00093

67. Hill WG, Mulder HA. Genetic analysis of environmental variation. Genet Res. 2010;92: 381–395. doi:10.1017/S0016672310000546

68. Hamer DH. Beware the chopsticks gene. Mol Psychiatry. 2000;5: 11–13.

69. Odling-Smee F, Laland K, Feldman M. Niche construction: the neglected process in evolution. Princeton, NJ: Princeton University Press; 2003.

70. Day R, Laland K, Odling-Smee F. Rethinking adaptation: the niche-construction perspective. Perspect Biol Med. 2003;46: 80–95.

71. Laland K, Brown G. Niche construction, human behavior, and the adaptive-lag hypothesis. Evol Anthropol Issues News Rev. 2006;15: 95–104.

72. Laland K, Odling-Smee FJ, Feldman MW. Evolutionary consequences of niche construction and their implications for ecology. Proc Natl Acad Sci. 1999;96: 10242–10247.

73. Odling-Smee F, Laland K, Feldman M. Niche construction. Am Nat. 1996;147: 641–648.

74. Wang G, Yang E, Brinkmeyer-Langford CL, Cai JJ. Additive, Epistatic, and Environmental Effects Through the Lens of Expression Variability QTL in a Twin Cohort. Genetics. 2014;196: 413–425. doi:10.1534/genetics.113.157503

75. Dieleman JL, Templin T. Random-Effects, Fixed-Effects and the within-between Specification for Clustered Data in Observational Health Studies: A Simulation Study. PLOS ONE. 2014;9: e110257. doi:10.1371/journal.pone.0110257

76. Ahmad S, Rukh G, Varga T V., Ali A, Kurbasic A, Shungin D, et al. Gene× physical activity interactions in obesity: combined analysis of 111,421 individuals of European ancestry. PLoS Genet. 2013;9: e1003607.

77. Sun X, Elston R, Morris N, Zhu X. What Is the Significance of Difference in Phenotypic Variability across SNP Genotypes? Am J Hum Genet. 2013;93: 390–397.

78. Dumitrascu B, Darnell G, Ayroles J, Engelhardt BE. A Bayesian test to identify variance effects. ArXiv151201616 Q-Bio Stat. 2015; Available: http://arxiv.org/abs/1512.01616

79. dbGaP | phs000007.v29.p10 | Framingham Cohort [Internet]. [cited 7 Aug 2017]. Available: https://www.ncbi.nlm.nih.gov/projects/gap/cgi-bin/study.cgi?study_id=phs000007.v29.p10

80. Johnson A, Handsaker R, Pulit S, Nizzari MM, O'Donnell CJ, de Bakker PI. SNAP: a web-based tool for identification and annotation of proxy SNPs using HapMap. Bioinformatics. 2008;24: 2938–2939.

81. Pruim R, Welch R, Sanna S, Teslovich TM, Chines PS, Gliedt TP, et al. LocusZoom: regional visualization of genome-wide association scan results. Bioinformatics. 2010;26: 2336–2337.

82. A map of human genome variation from population scale sequencing. Nature. 2010;467: 1061–1073. doi:10.1038/nature09534

83. Yang J, Lee SH, Goddard ME, Visscher PM. GCTA: A Tool for Genome-wide Complex Trait Analysis. Am J Hum Genet. 2011;88: 76–82. doi:10.1016/j.ajhg.2010.11.011

84. Miller MB, Basu S, Cunningham J, Eskin E, Malone SM, Oetting WS, et al. The Minnesota Center for Twin and Family Research Genome-Wide Association Study. Twin Res Hum Genet Off J Int Soc Twin Stud. 2012;15: 767–774. doi:10.1017/thg.2012.62

85. Iacono WG, McGue M. Minnesota Twin Family Study. Twin Res Off J Int Soc Twin Stud. Australian Academic Press; 2002;5: 482–7. doi:10.1375/136905202320906327

86. Liu J, Mcrae A, Nyholt D, Medland SE, Wray NR, M. BK, et al. A versatile gene-based test for genome-wide association studies. Am J Hum Genet. 2010;37: 139–145.

87. Zhang K, Cui S, Chang S, Zhang L, Wang J. i-GSEA4GWAS: a web server for identification of pathways/gene sets associated with traits by applying an improved gene set enrichment analysis to genome-wide association study. Nucleic Acids Res. 2010;38: W90–W95.

88. Champely S. pwr: Basic functions for power analysis. R package version 1.1. 1. R Found. 2009;

89. Imbens GW, Kolesár M. Robust Standard Errors in Small Samples: Some Practical Advice. Rev Econ Stat. 2015;98: 701–712. doi:10.1162/REST_a_00552

90. Cameron AC, Miller DL. A Practitioner's Guide to Cluster-Robust Inference. J Hum Resour. 2015;50: 317–372. doi:10.3368/jhr.50.2.317

Supporting Information

S1 Fig. Regional association plots for QC-filtered SNPs that were genome-wide suggestively significant (p<10^{-5}) for association with sibling-pair standard deviation in height. Plots, produced by SNAP [80], show genotyped SNPs used in analysis of sibling variation in height. A–D) Plots for the SNPs rs2804263, rs2073302, rs8126205, and rs4834078, respectively.

S2 Fig. Regional linkage maps for FHS genome-wide suggestive SNPs for sibling-pair standard deviation in height from 1,000 Genomes, CEU Panel. Maps produced by SNAP [80]. A–D) Plots for the SNPs rs2804263, rs2073302, rs8126205, and rs4834078, respectively.

S3 Fig. **QQ plots associated with Manhattan plots in Fig. 1**. A) Observed versus expected p-value distributions for analysis of sibling-pair standard deviation in height for FHS generation-three respondents with controls for parental genotype, mean height of sibling pair, sex, and sex difference. B) Same as in (A) except for BMI instead of height. Shaded gray regions depict 95% confidence intervals.

S4 Fig. Power to detect an effect size of R^2 for the sample size used in the discovery analysis at genome-wide suggestive significance level (p<10^{-5}).

S5 Fig. Correlation between SNP mean effects and SNP association with squared Z-scores. SNPs are normalized for minor allele frequency (W) (See Data). A) For each SNP, association between the SNP and squared Z-scores for BMI is plotted against the SNP's effect on mean BMI (correlation approximately zero). B) Same as in (A) except only the top 100 SNPs (based on mean effects on BMI) are shown (correlation 0.87).
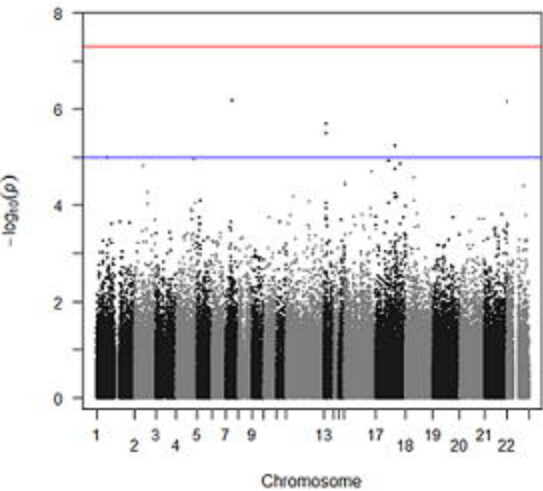
S6 Fig. Regional Association Plot of rs7202116, top hit for variance in BMI found by Yang et al. (2012), on mean level of BMI from GIANT consortium data. Figure produced using LocusZoom [81].

S7 Fig. Regional Association Plot of genome-wide suggestively significant (p<10^{-5}) hits from Fig 1 on mean height from GIANT consortium data. (A–D) Plots for the SNPs rs2804263, rs2073302, rs8126205, and rs4834078, respectively, show no markers in the respective regions that approach even genome-wide suggestive significance (p<10^{-5}). Figures produced using LocusZoom [81].

S8 Fig. Relationship between sibling-pair mean BMI and sibling-pair standard deviation (SD) or coefficient of variation (CV). A) Sibling-pair SD versus mean (correlation = 0.43). B) Sibling-pair CV versus mean (correlation = 0.25).

S9 Fig. Manhattan plots for enriched pathway HSA04540 Gap Junction for height variability. A) FHS discovery sample; B) MTFS replication sample.