

The Population Genetic Signature of Polygenic Local Adaptation

Jeremy J. Berg^{1,2,3} and Graham Coop^{1,2,3}

¹ Graduate Group in Population Biology, University of California, Davis.

² Center for Population Biology, University of California, Davis.

³ Department of Evolution and Ecology, University of California, Davis

To whom correspondence should be addressed: jjberg@ucdavis.edu, gmcoop@ucdavis.edu

Abstract

Adaptation in response to selection on polygenic phenotypes occurs via subtle allele frequencies shifts at many loci. Current population genomic techniques are not well posed to identify such signals. In the past decade, detailed knowledge about the specific loci underlying polygenic traits has begun to emerge from genome-wide association studies (GWAS). Here we combine this knowledge from GWAS with robust population genetic modeling to identify traits that have undergone local adaptation. Using GWAS data, we estimate the mean additive genetic value for a give phenotype across many populations as simple weighted sums of allele frequencies. We model the expected differentiation of GWAS loci among populations under neutrality to develop simple tests of selection across an arbitrary number of populations with arbitrary population structure. To find support for the role of specific environmental variables in local adaptation we test for correlations with the estimated genetic values. We also develop a general test of local adaptation to identify overdispersion of the estimated genetic values among populations. This test is a natural generalization of QST /FST comparisons based on GWAS predictions. Finally we lay out a framework to identify the individual populations or groups of populations that contribute to the signal of overdispersion. These tests have considerably greater power than their single locus equivalents due to the fact that they look for positive covariance between like effect alleles. We apply our tests to the human genome diversity panel dataset using GWAS data for six different traits. This analysis uncovers a number of putative signals of local adaptation, and we discuss the biological interpretation and caveats of these results.

1 Introduction

From the earliest days, the study of the genetic basis of evolutionary change has generated a great deal of controversy over the nature of the variation that contributes to evolution and adaptation (PEARSON, 1897; BATESON, 1902; WELDON, 1902, 1903; PEARSON, 1903; DARBISHIRE, 1905). The earliest debates pitted Mendelians against biometricians, and often centered on the question of whether evolution proceeds by continuous or discontinuous changes (see PROVIN, 2001, for a full account). The resolution of this conflict (YULE, 1902; WEINBERG, 1909; FISHER, 1918) was central to the development of early population and quantitative genetics, with both fields seeded in large part by FISHER's (1918) demonstration that the inheritance and evolution of quantitative characters could be explained by small contributions from many independent Mendelian loci.

While still theoretically aligned (TURELLI and BARTON, 1990), these two fields have often been divergent in empirical practice. Evolutionary quantitative geneticists have historically focused either on mapping the genetic basis of relatively simple traits (SLATE, 2004), or in the absence of any such knowledge, on understanding the evolutionary dynamics of phenotypes in response to selection over relatively short time-scales (KINGSOLVER *et al.*, 2001). Population geneticists, on the other hand, have focused on understanding the subtle signals left in genetic data by selection over longer time scales (HUDSON *et al.*, 1987; McDONALD and KREITMAN, 1991; BEGUN and AQUADRO, 1992), usually at the expense of a clear relationship between these patterns of genetic diversity and evolution at the phenotypic level. Given this still imperfect understanding, it should come as no surprise that much debate remains

over the nature of the genetic variation that matters for evolution and adaptation (e.g. ORR and COYNE, 1992; ROCKMAN, 2012).

Enabled by technological advancements allowing for the generation of vast quantities of DNA polymorphism or sequence data, population genomic techniques have revolutionized our ability to identify recent individual selective events. These approaches are typically based on identifying unusually large allele frequency changes either by observed differentiation among populations; by identifying the diversity reducing signatures of selective sweeps (MAYNARD SMITH and HAIGH, 1974; KAPLAN *et al.*, 1989; SABETI *et al.*, 2002, 2007; BERSAGLIERI *et al.*, 2004; NIELSEN *et al.*, 2005; WANG *et al.*, 2006; VOIGHT *et al.*, 2006; SABETI *et al.*, 2006; NIELSEN *et al.*, 2007; FUJIMOTO *et al.*, 2008; PICKRELL *et al.*, 2009; AKEY, 2009; CHEN *et al.*, 2010) or by identifying large allele frequency changes along geographic or environmental gradients (COOP *et al.*, 2010; HANCOCK *et al.*, 2010, 2011). While the interpretation of these tests remains difficult (e.g. COOP *et al.*, 2009; HERNANDEZ *et al.*, 2011), they provide the first insights into the mechanistic and ecological basis of adaptive evolutionary change at a genome-wide scale.

Such approaches are nonetheless limited because they rely on identifying individual loci that look unusual, and thus are only capable of identifying selection on traits where an individual locus has a large effect on fitness. When selection acts weakly on a phenotype that is underwritten by a large number of loci, the response at any given locus is expected to be modest (PRITCHARD *et al.*, 2010), and the signal instead manifests as a coordinated shift in allele frequency across many loci, with the phenotype increasing (decreasing) alleles all on average shifting in the same direction (LATTA, 1998, 2003; LE CORRE and KREMER, 2003, 2012; KREMER and LE CORRE, 2011). Because this signal is so weak at the level of the individual locus, it is impossible to identify against the genome-wide background without a very specific annotation of which sites are the target of selection on a given trait.

The advent of well-powered genome wide association studies with large sample sizes (RISCH and MERIKANGAS, 1996) has allowed for just this sort of annotation, enabling the mapping of many small effect alleles associated with phenotypic variation down to the scale of linkage disequilibrium in the population. The development and application of these methods in human populations has identified thousands of loci associated with a wide array of traits, largely confirming the polygenic view of phenotypic variation (VISSCHER *et al.*, 2012).

Although the field of human medical genetics has been the largest and most rapid to adopt such approaches, evolutionary geneticists studying non-human model organisms have also carried out GWAS for a wide array of fitness-associated traits, and the development of further resources is ongoing (ATWELL *et al.*, 2010; FOURNIER-LEVEL *et al.*, 2011; MACKAY *et al.*, 2012). In human populations, the cumulative contribution of these loci to the additive variance so far only explain a fraction of the narrow sense heritability for a given trait (usually less than 15%), a phenomenon known as the missing heritability problem (MANOLIO *et al.*, 2009; BLOOM *et al.*, 2013). Nonetheless, these GWAS hits represent a rich source of information about the loci underlying phenotypic variation.

Many researchers have begun to test whether the loci uncovered by these studies tend to be enriched for signals of selection, in the hopes of learning more about how adaptation has shaped phenotypic diversity and disease risk (e.g. MYLES *et al.*, 2008; CASTO and FELDMAN, 2011; JOSTINS *et al.*, 2012). The tests applied are generally still predicated on the idea of identifying individual loci that look unusual, such that a positive signal of selection is only observed if some subset of the GWAS loci have experienced strong enough selection to make them individually distinguishable from the genomic background. As noted above, it is unlikely that such a signature will exist if adaptation is truly polygenic, and thus many selective events will not be identified by this approach.

Here we develop and implement a general method based on simple quantitative and population genetic principals, using allele frequency data at GWAS loci to test for a signal of selection on the phenotypes they underwrite while accounting for the hierarchical structure among populations induced by shared history and genetic drift. Our work is most closely related to the recent work of TURCHIN *et al.* (2012); ? and CORONA *et al.* (2013), who look for co-ordinated shifts in allele frequencies of GWAS alleles for particular traits. Our approach constitutes an improvement over the methods implemented in these studies as it provides a high powered and theoretically grounded approach to investigate selection in an arbitrary number of populations with an arbitrary relatedness structure.

Using the set of GWAS effect size estimates and genome wide allele frequency data, we estimate the mean genetic value (FISHER, 1930; FALCONER, 1960) for the trait of interest in a diverse array of human populations. These genetic values may in some cases be poor predictors of the actual phenotypes for reasons we address below and in the Discussion. As such, we place little weight on their prediction

ability, and focus instead on population genetic modeling of their joint distribution across populations, which provides a strong and robust way of indirectly investigating the effects of selection on the genetic basis of complex phenotypes.

We develop a framework to describe how genetic values covary across populations based on a flexible model of genetic drift and population history. Using this null model, we implement simple test statistics based on transformations of the genetic values that remove this covariance among populations. We judge the significance of the departure from neutrality by comparing to a null distribution of test statistics constructed from well matched sets of control SNPs. Specifically, we test for local adaptation by asking whether the transformed genetic values show excessive correlations with environmental or geographic variables. We also develop and implement a less powerful but more general test, which asks whether the genetic values are over-dispersed among population compared to our null model of drift. We show that this overdispersion test, which is closely related to Q_{ST} (the phenotypic analog of F_{ST} ; PROUT and BARKER, 1993; SPITZE, 1993), as well as LEWONTIN and KRAKAUER’s (1973) test and subsequent extensions (BONHOMME *et al.*, 2010; GÜNTHER and COOP, 2012), gains considerable power to detect selection over single locus tests by looking for unexpected covariance among loci in the deviation they take from neutral expectations. Lastly, we develop an extension of our model that allows us to identify individual populations or groups of populations whose genetic values deviate from their neutral expectations given the values observed for related populations, and thus have likely been impacted by selection.

2 Results

Estimating Genetic Values with GWAS Data

Consider a trait of interest where L loci (SNPs) have been identified from a genome-wide association study. We arbitrarily label the alleles A_1 and A_2 at each locus. These loci have additive effect size estimates $\alpha_1, \dots, \alpha_L$, where α_ℓ is the average effect on an individual’s phenotype of replacing an A_2 allele with an A_1 allele at locus ℓ . We have allele frequency data for M populations at our L SNPs, and denote by $p_{m\ell}$ the observed sample frequency of allele A_1 at the ℓ^{th} locus in the m^{th} population. From these, we estimate the mean genetic value in the m^{th} population as

$$Z_m = 2 \sum_{\ell=1}^L \alpha_\ell p_{m\ell} \quad (1)$$

and we take \vec{Z} to be the vector containing the mean genetic values for all M populations.

The relationship between our estimated genetic values and true phenotypic values is imperfect, and thus it is worth taking a moment to cover these differences:

1. We assume that the L loci act strictly in an additive manner, i.e. no dominance or epistasis. This is likely to be a reasonable assumption for small effect loci, as typically found by GWASs (HILL *et al.*, 2008; CROW, 2010). However, if this assumption is violated, differences in allele frequencies across populations will result in differences in the additive effect size, and may lead to inconsistencies between estimated genetic values and observed phenotypes.
2. For the majority of GWAS associations we do not know the causal variant(s) at a locus, but rather a SNP that is in linkage disequilibrium (LD) with the causal variant(s). As the structure of LD may vary between populations, our estimates will be better in some populations than in others.
3. We assume that the effect sizes are perfectly estimated, when they may in fact be biased upward or downward due to the winner’s curse (ZOLLNER and PRITCHARD, 2007) or incomplete LD (SPENCER *et al.*, 2011) respectively.
4. The genetic values strictly apply only to the environment of the population in which the association study was done. They will only be representative of the phenotypes of the populations in their “natural” environment to the extent that environmental differences between populations are unimportant.

5. We do not have all of the causal variants in the population that the GWAS was performed in, and we may be missing other loci that contribute to variation in other populations but not the one in which the GWAS was done.

These limitations may result in inconsistencies between our estimated genetic values and the true genetic values, as well as the actual observed phenotypic values. However, because the methods we develop below rely only on the rejection of a neutral null model, none of them can lead to us infer a false positive signal of selection if the mapped loci evolve neutrally. They do, however, result in a loss of statistical power, and significantly complicate the biological interpretation of a positive signal. We note that one can obtain a false positive, however, due to unaccounted for population structure in the original GWAS ascertainment panel. We return to both of these issues in greater depth in the Empirical Applications and Discussion sections.

A Model of Genetic Value Drift

We are chiefly interested in developing a framework for testing the hypothesis that the joint distribution of \vec{Z} is driven by neutral processes alone, with rejection of this hypothesis implying the action of selection. We first describe a general model for the expected joint distribution of estimated genetic values (\vec{Z}) across populations under neutrality, accounting for genetic drift and shared population history.

A simple approximation to a model of genetic drift is that the current frequency of an allele in a population is normally distributed around some ancestral frequency (ϵ). Under a Wright-Fisher model of genetic drift, the variance of this distribution is approximately $f\epsilon(1 - \epsilon)$, where f is a property of the population shared by all loci, reflecting the compounded effect of many generations of binomially sampling (see NICHOLSON *et al.*, 2002, for more justification). Note also that for small values, f is approximately equal to the inbreeding coefficient of the present day population relative to the defined ancestral population, and thus has an interpretation as the correlation between two randomly chosen alleles relative to the ancestral population (NICHOLSON *et al.*, 2002).

We can expand this framework to describe the distribution of allele frequencies in an arbitrary number of populations for an arbitrary demographic history by assuming that the vector of allele frequencies in M populations follows a multivariate normal distribution

$$\vec{p} \sim MVN\left(\epsilon\vec{1}, \epsilon(1 - \epsilon)\mathbf{F}\right). \quad (2)$$

where \mathbf{F} is an M by M positive definite matrix describing the correlation structure of allele frequencies across populations relative to the mean/ancestral frequency. Note again that for small values it is also approximately the matrix of inbreeding coefficients (on the diagonal) and kinship coefficients (on the off-diagonals) describing the relatedness among populations (WEIR and HILL, 2002; BONHOMME *et al.*, 2010). This flexible model was introduced, to our knowledge, by CAVALLI-SFORZA *et al.* (1964) (see FELSENSTEIN, 1982, for a review), and has subsequently been used as a computationally tractable model for population history inference (e.g. NICHOLSON *et al.*, 2002; ?), and as a null model for signals of selection (e.g. COOP *et al.*, 2010; BONHOMME *et al.*, 2010; GÜNTHER and COOP, 2012; FARIELLO *et al.*, 2013). So long as the multivariate normal assumption of drift holds reasonably well, this framework can summarize arbitrary population histories, including tree-like structures with substantial gene flow between populations (?), or even those which lack any coherent tree-like component (e.g. isolation by distance models: GUILLOT, 2012; BRADBURY *et al.*, 2013).

Recall that our estimated genetic values (\vec{Z}) are merely a sum of sample allele frequencies weighted by effect size. If the underlying allele frequencies are well described by the multivariate normal model described above, then the distribution of \vec{Z} is a weighted sum of multivariate normal distributions, such that this distribution is itself multivariate normal:

$$\vec{Z} \sim MVN\left(\mu\vec{1}, V_A\mathbf{F}\right) \quad (3)$$

where $\mu = \frac{1}{L} \sum_{\ell=1}^L 2\alpha_{\ell}\epsilon_{\ell}$ and $V_A = 4 \sum_{\ell=1}^L \alpha_{\ell}^2\epsilon_{\ell}(1 - \epsilon_{\ell})$ are respectively the expected phenotypic value and additive genetic variance of the ancestral (global) population. The covariance matrix describing the distribution of \vec{Z} therefore differs from that describing the distribution of frequencies at individual loci only by a scaling factor that can be interpreted as the contribution of the associated loci to the additive

genetic variance present in a hypothetical population with allele frequencies equal to those of the entire metapopulation.

The assumption that the drift of allele frequencies around their shared mean is normally distributed (2) may be problematic if there is substantial drift. However, even if that is the case, the estimated genetic values may still be assumed to follow a multivariate normal distribution by appealing to the central limit theorem, as each estimated genetic value is a sum over many loci. We show in the Results that this assumption often holds in practice.

Note that in light of the caveats discussed above, the elements of \vec{Z} are likely to be imperfect predictions both of the true genetic values, and the observed phenotypic values. However these caveats do not invalidate our null model, as it describes only the expected allele frequency change under neutrality at the associated loci, which should be independent from their identification in a GWAS (provided that ascertainment schemes are not too complicated) and from the effect sizes estimated. As such, if the GWAS loci truly are evolving neutrally, then their evolution should be well described by our model regardless of which of the assumptions outlined above are violated.

Fitting the Model and Standardizing the Estimated Genetic Values

As described above, we obtain the vector \vec{Z} by summing allele frequencies across loci while weighting by effect size. We do not get to observe the ancestral genetic value of the sample (μ), so we assume that this is simply equal to the mean genetic value across populations ($\mu = \frac{1}{M} \sum_m Z_m$). This assumption costs us a degree of freedom, and so we must work with a vector \vec{Z}' , which is the vector of estimated genetic values for the first $M - 1$ populations, centered at the mean of the M (see Methods for details). Note that this procedure will be the norm for the rest of this paper, and thus we will always work with vectors of length $M - 1$ that are obtained by subtracting the mean of the M vector and dropping the last component.

To estimate the null covariance structure of the $M - 1$ populations we sample a large number K random unlinked SNPs. In our procedure, the K SNPs are sampled so as to match certain properties of the L GWAS SNPs (the specific matching procedure is described in more depth below and in the Methods section). Setting ϵ_k to be the mean sample allele frequency across populations at the k^{th} SNP, we standardize the sample allele frequency in population m as $(p_{mk} - \epsilon_k) / (\epsilon_k (1 - \epsilon_k))$. We then calculate the sample covariance matrix (\mathbf{F}) of these standardized frequencies, accounting for the $M - 1$ rank of the matrix (see Methods). We estimate the scaling factor of this matrix \mathbf{F} as

$$V_A = 4 \sum_{\ell=1}^L \alpha_{\ell}^2 \epsilon_{\ell} (1 - \epsilon_{\ell}). \quad (4)$$

We now have an estimated genetic value for each population, and a simple null model describing their expected covariance due to shared population history. Under this multivariate normal framework, we can transform the vector of mean centered genetic values (\vec{Z}') so as to remove this covariance. First, we note that the Cholesky decomposition of the \mathbf{F} matrix is

$$\mathbf{F} = \mathbf{C}\mathbf{C}^T \quad (5)$$

where \mathbf{C} is a lower triangular matrix, and \mathbf{C}^T is its transpose. Informally, this can be thought of as taking the square root of \mathbf{F} , and so \mathbf{C} can be thought of as analogous to the standard deviation matrix.

Using this matrix \mathbf{C} we can transform our estimated genetic values as:

$$\vec{X} = \frac{1}{\sqrt{V_A}} \mathbf{C}^{-1} \vec{Z}'. \quad (6)$$

If $\vec{Z}' \sim \text{MVN}(\vec{0}, V_A \mathbf{F})$ then $\vec{X} \sim \text{MVN}(0, \mathbb{I})$, where \mathbb{I} is the identity matrix. Therefore, under the assumptions of our model, these standardized genetic values should be independent and identically distributed $\sim N(0, 1)$ random variates.

It is worth spending a moment to consider what this transformation has done to the allele frequencies at the loci underlying the estimated genetic values. As our original genetic values are written as a

weighted sum of allele frequencies, our transformed genetic values can be written as a weighted sum of transformed allele frequencies (which have passed through the same transform). We can write

$$\vec{X} = \frac{1}{\sqrt{V_A}} \mathbf{C}^{-1} \vec{Z}' = \frac{2}{\sqrt{V_A}} \sum_{\ell} \alpha_{\ell} \mathbf{C}^{-1} \vec{p}_{\ell} \quad (7)$$

and so we can simply define the vector of transformed allele frequencies at locus ℓ to be

$$\vec{p}'_{\ell} = \mathbf{C}^{-1} \vec{p}_{\ell}. \quad (8)$$

This set of transformed frequencies exist within a set of transformed populations, which by definition have zero covariance with one another under the null, and are related by a star-like population tree with branches of equal length.

As such, we can proceed with simple, straightforward and familiar statistical approaches to test for the impact of spatially varying selection on the estimated genetic values. Below we describe three simple methods for identifying the signature of polygenic adaptation, which arise naturally from this observation.

Environmental correlations

Assume we have a vector \vec{Y} , containing measurements of a specific environmental variable of interest in each of the M populations. We mean-center this vector and put it through a transform identical to that which we applied to the estimated genetic values in (7). This gives us a vector \vec{Y}' , which is in the same frame of reference as the transformed genetic values.

There are many possible models to describe the relationship between a trait of interest and a particular environmental variable that may act as a selective agent. We first consider a simple linear model, where we model the distribution of transformed genetic values (\vec{X}) as a linear effect of the transformed environmental variables (\vec{Y}')

$$\vec{X} \sim \beta \vec{Y}' + \vec{e} \quad (9)$$

where \vec{e} under our null is a set of normal, independent and identically distributed random variates (i.e. residuals), and we can estimate β by finding the value that minimizes the variance of \vec{e} (i.e. ordinary least squares regression). We can also calculate the associated squared Pearson correlation coefficient (r^2) as a measure of the fraction of variance explained by our variable of choice, as well as the non-parametric Spearman's rank correlation $\rho(\vec{X}, \vec{Y}')$, which is robust to outliers that can mislead the linear model. In the Methods we show that the linear environmental model applied to our transformed genetic values has a natural interpretation in terms of the underlying individual loci. Therefore, exploring the environmental correlations of estimated genetic values nicely summarizes information in a sensible way at the underlying loci identified by the GWAS.

In order to assess the significance of these measures, we implement an empirical null hypothesis testing framework, using β , r^2 , and ρ as test statistics. We sample many sets of L SNPs randomly from the genome, again applying a matching procedure discussed below and in the Methods. With each set of L SNPs we construct a vector \vec{Z}_{null} , which represents a single draw from the genome-wide null distribution for a trait with the given ascertainment profile. We then perform an identical set of transformations and analyses on each \vec{Z}_{null} , thus obtaining an empirical genome-wide null distribution for all test statistics.

Excess Variance Test

As an alternative to testing the hypothesis of an effect by a specific environmental variable, one might simply test whether the estimated genetic values exhibit more variance among populations than expected due to drift. Here we develop a simple test of this hypothesis.

As \vec{X} is composed of $M - 1$ independent, identically distributed standard normal random variables, a natural choice of test statistic is given by

$$Q_X = \vec{X}^T \vec{X} = \frac{\vec{Z}'^T \mathbf{F}^{-1} \vec{Z}'}{V_A}. \quad (10)$$

This Q_X statistic represents a standardized measure of the among population variance in estimated genetic values that is not explained by drift and shared history. It is also worth noting that by comparing the rightmost form in (10) to the multivariate normal likelihood function, we find that Q_X is proportional to the negative log likelihood of the estimated genetic values under the model, and is thus the natural measurement of the model's ability to explain their distribution. Values of this statistic that are in the upper tail correspond to an excess of variance among populations (i.e. local adaptation), while values in the lower tail correspond to a paucity of variance, and thus potentially to stabilizing selection. Multivariate normal theory predicts that this statistic should follow a χ^2 distribution with $M - 1$ degrees of freedom under the null hypothesis. Nonetheless, we use a similar approach to that described for the linear model, generating the empirical null distribution by resampling SNPs genome-wide. As discussed below, we find that in practice the empirical null distribution tends to be very closely matched by the theoretically predicted χ^2_{M-1} distribution.

The Relationship of Q_X to Previous Tests Our Q_X statistic is closely related to Q_{ST} , the phenotypic analog of F_{ST} , which measures the fraction of the genetic variance that is among populations relative to the total genetic variance (WRIGHT, 1951; SPITZE, 1993; PROUT and BARKER, 1993). Q_{ST} is typically estimated in traditional local adaptation studies via careful measurement of phenotypes from related individuals in multiple populations in a common garden setting. If the loci underlying the trait act in a purely additive manner and are experiencing only neutral genetic drift, then $\mathbb{E}[Q_{ST}] = \mathbb{E}[F_{ST}]$ (LANDE, 1992; WHITLOCK, 1999).

If both quantities are well estimated, and we also assume that there is no hierarchical structure among the populations, then $\frac{(M-1)Q_{ST}}{F_{ST}}$ is known to have a χ^2_{M-1} distribution under a wide range of models (ROGERS and HARPENDING, 1983; WHITLOCK, 2008; WHITLOCK and GUILLAUME, 2009). Thus the statistic $\frac{(M-1)Q_{ST}}{F_{ST}}$ is a natural phenotypic extension of the LEWONTIN and KRAKAUER (1973) F_{ST} based-test (LK test).

To see the close correspondence between Q_X and Q_{ST} , consider the case of a starlike population tree with branches of equal length (i.e. $f_{mm} = F_{ST}$ and $f_{m \neq n} = 0$). Under this demographic model, we have

$$\begin{aligned} Q_X &= \frac{(\bar{Z} - \mu)^T \mathbf{F}^{-1} (\bar{Z} - \mu)}{V_A} = \frac{(Z_1 - \mu)^2}{V_A F_{ST}} + \dots + \frac{(Z_{M-1} - \mu)^2}{V_A F_{ST}} \\ &= \frac{(M-1) \text{Var}(\bar{Z})}{V_A F_{ST}} \\ &= \frac{(M-1) \hat{Q}_{ST}}{F_{ST}} \end{aligned} \quad (11)$$

where \hat{Q}_{ST} is an estimated value for Q_{ST} obtained from our estimated genetic values. This relationship between Q_X and Q_{ST} breaks down when some pairs of populations do not have zero covariance in allele frequencies under the null, in which case the χ^2 distribution of the LK test also breaks down (ROBERTSON, 1975; NEI and MARUYAMA, 1975). BONHOMME *et al.* (2010) recently proposed an extension to the LK test that accounts for a population tree, thereby recovering the χ^2 distribution (see also GÜNTHER and COOP, 2012, who relax the tree-like assumption), and our Q_X statistic is the natural phenotypic extension of this enhanced statistic.

Given that our estimated genetic values are simple linear sums of allele frequencies, it is natural to ask how Q_X can be written in terms of these frequencies. Again, restricting ourselves to the case where \mathbf{F} is diagonal, (i.e. $f_{mm} = F_{ST}$ and $f_{m \neq n} = 0$), we can express Q_X as

$$Q_X = \frac{4}{V_A F_{ST}} \sum_{m=1}^{M-1} \sum_{\ell, \ell'} \alpha_\ell \alpha_{\ell'} (p_{m\ell} - \epsilon_\ell)(p_{m\ell'} - \epsilon_{\ell'}), \quad (12)$$

which can be rewritten as

$$Q_X = \frac{M-1}{F_{ST}} \left(\frac{\sum_\ell \alpha_\ell^2 \text{Var}(\vec{p}_\ell)}{\sum_\ell \alpha_\ell^2 \epsilon_\ell (1 - \epsilon_\ell)} + \frac{\sum_{\ell \neq \ell'} \alpha_\ell \alpha_{\ell'} \text{Cov}(\vec{p}_\ell, \vec{p}_{\ell'})}{\sum_\ell \alpha_\ell^2 \epsilon_\ell (1 - \epsilon_\ell)} \right). \quad (13)$$

The numerator of the first term inside the parentheses is the weighted sum of the variance among populations over all GWAS loci, scaled by the contribution of those loci to the additive genetic variance in the total population. As such this first term is similar to F_{ST} calculated for our GWAS loci, but instead of just averaging the among population and total variances equally across loci in the numerator and denominator, these quantities are weighted by the squared effect size at each locus. This weighting nicely captures the relative importance of different loci to the trait of interest.

The second term in (13) is less familiar; the numerator is the weighted sum of the covariance of allele frequencies between all pairs of GWAS loci, and the denominator is again the contribution of those loci to the additive genetic variance in the total population. This term is thus a measure of the correlation across loci in the deviation they take from the ancestral value, or the across population component of linkage disequilibrium (for a more in depth discussion of this relationship in the context of Q_{ST} , see LATTA, 1998, 2003).

Under the neutral null hypothesis, the expectation of this second term is equal to zero, as drifting loci are equally likely to covary in either direction. With differential selection among populations, however, we expect loci underlying a trait not only to vary more than we would expect under a neutral model, but also to covary in a consistent way, e.g. we expect the alleles with a positive effect on the trait to increase or decrease in concert with one another. Models of local adaptation predict that it is this covariance among alleles that is primarily responsible for differentiation at the phenotypic level (LATTA, 1998, 2003; LE CORRE and KREMER, 2003, 2012; KREMER and LE CORRE, 2011), and as such, we expect this test statistic to offer considerably higher power than simply averaging F_{ST} over loci, or inspecting F_{ST} at individual GWAS loci, a point which we investigate with simulations below.

As noted above (8), when \mathbf{F} is non-diagonal, our transformed genetic values can be written as a weighted sum of transformed allele frequencies. Consequently, we can obtain a similar expression to (13) when population structure exists, but now expressed in terms of the covariance of a set of allele frequencies in transformed populations that have no covariance with each other under the null hypothesis. Specifically, when the covariance is non-diagonal we can write:

$$Q_X = \frac{M-1}{V_A} \sum_{\ell} \alpha_{\ell}^2 \text{Var}(\vec{p}_{\ell}) + \frac{M-1}{V_A} \sum_{\ell \neq \ell'} \alpha_{\ell} \alpha_{\ell'} \text{Cov}(\vec{p}_{\ell}, \vec{p}_{\ell'}). \quad (14)$$

We refer to the first term in this sum as the standardized F_{ST} -like component and the second term as the standardized LD-like component. We use this decomposition below to demonstrate for empirical examples that significant Q_X statistics are due to unexpected covariance between like effect alleles, rather than unexpectedly large variance at individual loci.

Identifying Outlier Populations

Having detected a putative signal of selection for a given trait, one may wish to identify individual regions and populations which contribute to the signal. Here we rely on our multivariate normal model of relatedness among populations, along with well understood methods for generating conditional multivariate normal distributions, in order to investigate specific hypotheses about individual populations or groups of populations.

Briefly, using standard results from multivariate normal theory, we can generate the expected joint conditional distribution of genetic values for an arbitrary set of populations given the observed genetic values in some other set of populations. These conditional distributions allow for a convenient way to ask whether the estimated genetic values observed in certain populations or groups of populations differ significantly from the values we would expect them to take given the values observed in related populations.

Specifically, we exclude a set of populations and then calculate the expected mean and variance of genetic values in these excluded populations given the remaining populations. Using this mean and variance we calculate a Z-score to describe how well fit the estimated genetic values of the excluded populations are by our model of drift, conditional on the values in the remaining populations. The details of our procedure are given in the Methods section, and we pursue implementations of this approach at two levels of population structure in our Empirical Applications.

Datasets

We conducted power simulations and an empirical application of our methods based on the Human Genome Diversity Panel (HGDP) population genomic dataset (Li *et al.*, 2008), and a number of GWAS SNP sets. To ensure that we made the fullest possible use of the information in the HGDP data, we took advantage of a genome wide allele frequency dataset of ~ 3 million SNPs imputed from the Phase II HapMap into the 52 populations of the HGDP. These SNPs were imputed as part of the HGDP phasing procedure in PICKRELL *et al.* (2009); see our Methods section for a recap of the details. We applied our method to test for signals of selection in six human GWAS datasets identifying SNPs associated with height, skin pigmentation, body mass index (BMI), type 2 diabetes (T2D), Crohn’s Disease (CD) and Ulcerative Colitis (UC).

Choosing null SNPs Various components of our procedure involve sampling random sets of SNPs from across the genome. While we control for biases in our test statistics introduced by population structure through our \mathbf{F} matrix, we are also concerned that subtle ascertainment effects of the GWAS process could lead to biased test statistics, even under neutral conditions. We control for this possibility by sampling null SNPs so as to match the joint distribution of certain properties of the ascertained GWAS SNPs. Specifically, we chose our random SNPs to match the GWAS SNPs in each study in terms of their minor allele frequency (MAF) in the ascertainment population and the imputation status of the allele in our population genomic dataset. In addition, we were concerned that GWAS SNPs might be preferentially found close to genes and in low recombination regions, the latter due to better tagging, and as such may be subject to a high rate of drift due to background selection, leading to higher levels of differentiation at these sites (CHARLESWORTH *et al.*, 1997). Therefore, in addition to MAF and imputation status, we also matched our random SNPs to an estimate of the background selection environment experienced by the GWAS SNPs, as measured by B value (McVICKER *et al.*, 2009), which is a function of both the density of functional sites and recombination rate calibrated to match the reduction in genetic diversity due to background selection. We detail the specifics of the binning scheme for matching the discretized distributions of GWAS and random SNPs in the Methods.

Power Simulations

To assess the power of our methods in comparison to other possible approaches, we conducted a series of power simulations. As statistical power is necessarily a function of the data, we sample 1000 null sets of SNPs for each phenotype according to the same matching procedure described above and in the Methods. We then gradually increase the effect of selection on these underlying loci, and measure the corresponding increase in power.

For each set of matched SNPs, we compared two of our statistics, (r^2 and Q_X) against their naive counterparts, which are not adjusted for population structure (nr^2 and Q_{ST}). For Q_X and Q_{ST} , we count a given simulation as having detected the signal of selection if it lies in the upper 5% tail of the null distribution, whereas for the environmental correlation statistics (r^2 and nr^2) we use a two-tailed 5% test. We also compared our tests to two single locus enrichment tests, the first where we counted the number of loci with individual SNP environmental correlations in the 5% tail using the single locus r^2 (slr^2) and in the second we counted the number of loci in the upper 5% tail of F_{ST} (slF_{ST}). We then considered these enrichment tests to have detected the signal of selection if the number of individual loci in the 5% tail was itself in the 5% tail using a binomial test. We do not include our alternative linear model statistics β and ρ in these plots for the sake of figure legibility, but they generally had very similar power to that of r^2 . While slightly more powerful versions of the single locus tests that account for sampling noise are available (e.g. WEIR and HILL, 2002; COOP *et al.*, 2010), note that our tests could be extended similarly as well, so the comparison is fair.

We base our power simulations on empirical data altered to have an increasing effect of selection along a latitudinal gradient. In order to mimic the effect of selection, we generate a new set of allele frequencies ($p_{s,m\ell}$) by taking the original frequency ($p_{m\ell}$) and adding a small shift according to

$$p_{s,m\ell} = p_{m\ell} + p_{m\ell} (1 - p_{m\ell}) \alpha_{\ell} \delta Y_m \quad (15)$$

where α_{ℓ} is the effect size assigned to locus ℓ , and Y_m is the mean centered absolute latitude of the population. We used the 1000 simulations at $\delta = 0$ to form null distribution for each of our test

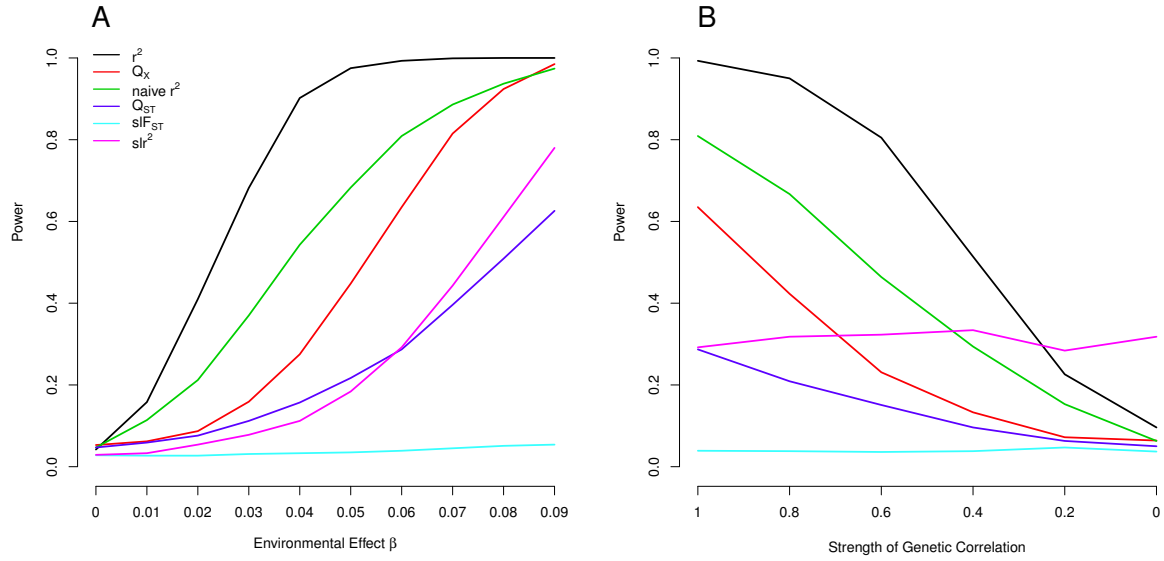


Figure 1: Power of our statistics as compared to alternative approaches (A) Across a range of effect sizes (β) of latitude, and (B) when we hold β constant at 0.06 and decrease ϕ , the genetic correlation between the trait of interest on the selected trait.

statistics, and from this established the 5% significance level. We then increment δ and give the power of each statistic as the fraction of simulations whose test statistic falls beyond this cutoff. While this approach to simulating selection is obviously naive to the way selection actually operates, it captures most of the important effects on the loci underlying a given trait. Namely, loci will have greater shifts if they experience rather extreme environments, have large effects on the phenotype, or are at intermediate frequencies. Because we add these shifts to allele frequencies sampled from real, putatively neutral loci, the effect of drift on their joint distribution is already present, and thus does not need to be simulated. Figure 1 shows simulations matched to the set of p_ℓ and α_ℓ of the Type 2 Diabetes dataset (see Figures S1-S5 for all other traits).

Our population structure adjusted statistics clearly outperform tests that do not account for structure, as well as single locus outlier based tests. Notably, tests based on an excess of F_{ST} outliers or average F_{ST} (not shown here) have essentially zero power for cases in which our tests have nearly 100% power. This nicely illustrates the fact that for highly polygenic traits, nearly all of the differentiation at the trait level arises as a consequence of across population linkage disequilibrium among the underlying loci, and not as a result of substantial differentiation at the loci themselves. While our environment-genetic value correlation tests outperform Q_X , this is somewhat artificial as it assumes that we know the environmental variable responsible for our allele frequency shift. In reality we may at best hope to guess a correlate of the selection pressure, so Q_X may outperform the environmental correlations.

In many cases, it may be that selection has not acted directly on the phenotype in question, but rather on one that is genetically correlated. Therefore, we are interested in the power of our tests to detect selection on traits genetically correlated with those for which we have GWAS data. The question of how to properly simulate this relationship is largely an empirical one, as the correlation structure underlying any two traits will depend on their shared genetic architecture. We therefore chose a simple and general approach to capture a flavor of this situation. We simulate the effect of selection as above (15), but using effect sizes (α'_ℓ) for an unobserved phenotype which are only partially correlated with our observed effect sizes. For simplicity we assume that α_ℓ and α'_ℓ have a bivariate normal distribution around zero with equal variance and correlation parameter ϕ , and then simulate α'_ℓ from its conditional distribution given α_ℓ (i.e. $\alpha'_\ell | \alpha_\ell \sim N(\phi\alpha_\ell, (1 - \phi^2)Var(\alpha))$). For each SNP ℓ in (15) we replaced α_ℓ by its effect α'_ℓ on the unobserved phenotype. Here ϕ can be thought of as the genetic correlation between

Table 1: The contribution of each geo-climatic variable to each of our four principal components, scaled such that the absolute value of the entries in each column sum to one (up to rounding error). We also show for each principal component the percent of the total variance across all eight variables that is explained by the PC.

Geo-Climatic Variable	SUMPC1	SUMPC2	WINPC1	WINPC2
Latitude	-0.16	-0.10	-0.17	-0.01
Longitude	0.02	0.12	0.04	0.05
Maximum Temp	0.24	-0.08	0.17	-0.03
Minimum Temp	0.24	0.07	0.17	0.08
Mean Temp	0.25	-0.03	0.17	0.03
Precipitation Rate	-0.01	0.16	0.07	0.32
Relative Humidity	-0.06	0.21	-0.06	0.34
Short Wave Radiation Flux	-0.03	-0.22	0.15	-0.13
Percent of Variance Explained	38%	35%	58%	20%

our phenotypes if this simple multivariate form held true for all of the loci contributing to our two traits.

In Figure 1 we hold the value of β constant at 0.06 and vary this genetic correlation ϕ from one down to zero. Predictably, our GWAS prediction-based statistics lose power as the two traits become more and more genetically independent, but do retain reasonable power out to quite low genetic correlations. In contrast, counting the number of SNPs that are significantly correlated with a given environmental variable remains equally powerful across all effect sizes, as these loci are still altered from neutral expectations along the environmental gradient, and the loci’s effect sizes and directionality do not enter into our calculation.

In some ways this may be considered a desirable property of the environmental outliers enrichment approach, it can still detect the effect of selection on these loci, even if the genetic effect sizes of the two traits are uncorrelated. On the other hand, this is also problematic, as it implies that without detailed knowledge of the various phenotypic associations at each locus, such tests may often be detecting selection on pleiotropically related phenotypes. Our approaches, however, are more suited to determining whether the genetic basis of a trait of interest has been affected by selection on a trait that is at least somewhat genetically correlated.

Empirical Applications

We estimated genetic values for each of six traits from the subset of GWAS SNPs that were present in the HGDP dataset, as described above. We discuss the analysis of each dataset in detail below, and address general points first. For each dataset, we constructed the covariance matrix from a sample of approximately 20,000 appropriately matched SNPs, and the null distributions of our test statistics from a sample of 10,000 sets of null genetic values, which were also constructed according to a similar matching procedure (as described in the Methods).

To test for an effect of major climate variables on the genetic basis of a given trait, we used the set of six climate variables (measured in both summer and winter seasons) considered by HANCOCK *et al.* (2008) in their analysis of adaptation to climate at the level of individual SNPs. We followed their general procedure by running principal components (PC) analysis for both seasons on a matrix containing the six climate variables, as well as latitude and longitude (following HANCOCK *et al.*’s (2008) rationale that these two geographic variables may capture certain elements of the long term climatic environment experienced by human populations). The percent of the variance explained by these PCs and their weighting (eigenvectors) of the different environmental variables are given in Table 1.

We also applied our Q_X test to identify traits whose underlying loci showed consistent patterns of unusual differentiation across populations. In Figure 2 we show for each GWAS set the observed value of Q_X and the empirical null distribution of Q_X using SNPs matched to the GWAS loci as described above. We also plot the expected null distribution of the Q_X statistic ($\sim \chi^2_{51}$). The expected null distribution closely matches the empirical distribution, suggesting that our multivariate normal framework provides a good null model for the data (although we will use the empirical null distribution to obtain measures of statistical significance).

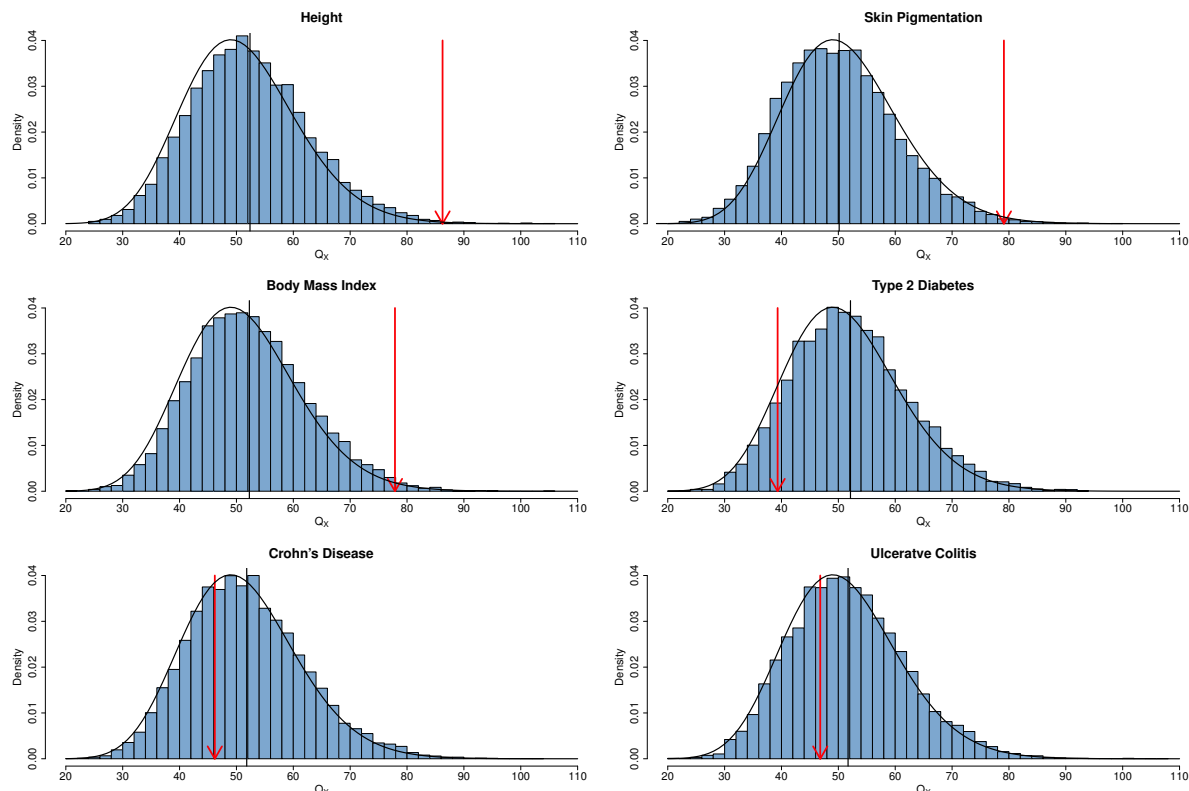


Figure 2: Histogram of the empirical null distribution of Q_X for each trait, obtained from genome-wide resampling of well matched SNPs. The mean of each distribution is marked with a vertical black bar and the observed value is marked by a red arrow. The expected χ^2_{M-1} density is shown as a black curve.

For each GWAS SNP set we can also separate our Q_X statistic into its two components (an F_{ST} -like component and an LD-like term), as described in (14). In Figure 3 we plot the null distributions of these two components for the height dataset as histograms, with the observed value marked by red arrows. As with all of the significant Q_X statistics we have examined, the signal of unusual differentiation comes largely from the standardized LD-like component (see Figures S6-S10 for all other traits at the global level). This highlights the fact that the information about departures from neutrality comes largely from a subtle signal of covariance across loci of like effect alleles.

As discussed above and in the Methods, we can use our multivariate normal framework to identify specific regions or populations that are poorly fit by the neutral model when predicted from other populations. We do this for each of our GWAS SNP sets at two different levels of population structure: seven broad geographic regions (as identified by ROSENBERG *et al.*, 2002) as well as individual populations. We first calculate the average of the estimated genetic values within one of the seven broad regions. We then calculate the expected joint conditional distribution for the populations in that region, given the estimated genetic values observed in each of the other six regions. This results in a vector of expectations for each population, and a conditional covariance matrix relating them, which we can convert into the expected conditional mean and variance of our regional average. We use this expected conditional mean and variance and the observed regional average to calculate a Z score as a measure of how differentiated that region is from the rest of the world. In Figures 5 and 6, we present the conditional expectation for each region as a light grey dashed line, and the observed mean as a solid bar, with darkness and thickness proportional to Z score.

We also calculate the conditional expectation and variance for each population in turn, given the values observed for all other populations in the dataset, allowing us to calculate population specific Z scores. Whereas our regional level Z scores are designed to detect selection events occurring over broader geographical regions (and potentially deeper time scales), our populations specific Z scores are designed to identify individual populations that are not well predicted by the model, and thus may have experienced

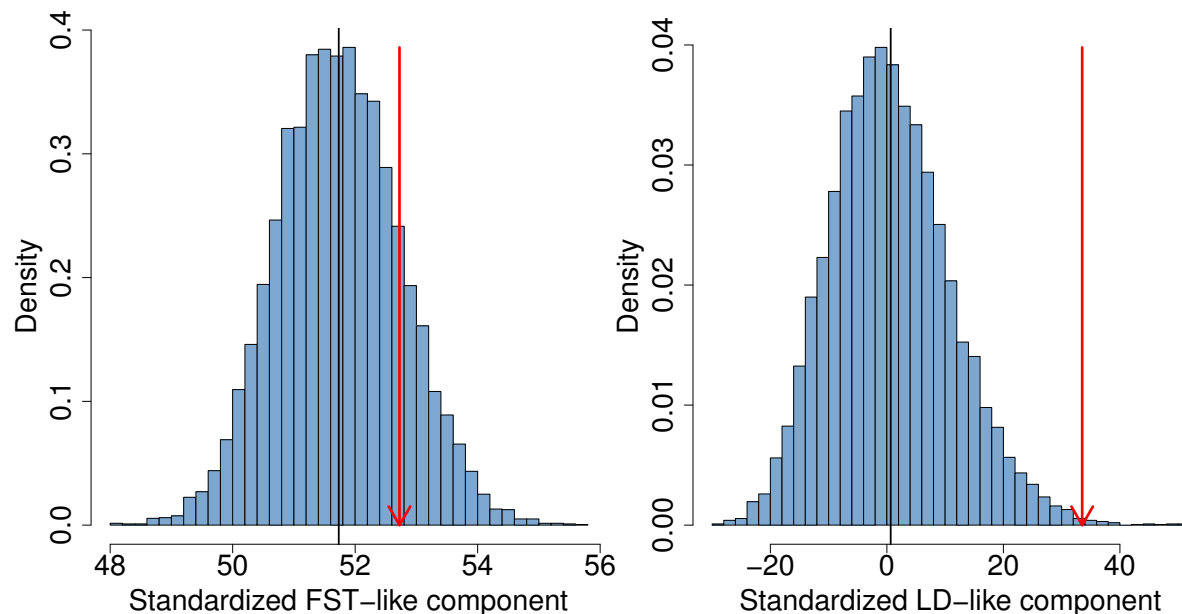


Figure 3: The two components of Q_X for the height dataset, as described by the left and right terms in (14). The null distribution of each statistic is shown as a histogram. The mean value is shown as a black bar, and the observed value as a red arrow.

more recent selection. In Figures 5 and 6, we present each population’s estimated genetic value as a colored circle, with its size proportional to its population specific Z score. Full characterizations of these results are available in Tables S3-S14.

Height We first analyzed the set of 180 height associated loci identified by LANGO ALLEN *et al.* (2010), which explain approximately 14% of the additive genetic variance for height (ZAITLEN *et al.*, 2013). This dataset is an ideal first test for our methods because it contains the largest number of associations identified for a single phenotype to date, maximizing our power gain over single locus methods (Figure S1). In addition, TURCHIN *et al.* (2012) have already identified a signal of pervasive weak selection at these same loci in European populations, and thus we should expect our methods to replicate this observation.

Of the 180 loci identified by LANGO ALLEN *et al.* (2010), 161 were present in our HGDP dataset. We used these 161 loci in conjunction with the allele frequency data from the HGDP dataset to estimate genetic values for height in the 52 HGDP populations. Although these genetic values are correlated with the observed heights of these populations, as expected they are far from perfect predictions (see Figure S11 and Table S1, which compares our estimated genetic values to observed sex-average heights for the subset of HGDP populations with a close proxy in the dataset of GUSTAFSSON and LINDENFORS, 2009). We find a signal of excessive correlation with winter PC2 (Figure 4 and Table 2), but find no effect of any other climatic variables. Our Q_X test also strongly rejects the neutral hypothesis, suggesting that our estimated genetic values are overly dispersed compared to the null model of neutral genetic drift and shared population history (Figure 2 and Table 2),

We followed up on this signal of differentiation by conducting regional level analyses, which indicate that the worldwide signal of excess variance arises from both inter- and intra-regional processes. We find that the mean genetic height in Africa is lower than expected given the values observed in the rest of the world ($Z = -2.0$, $p = 0.045$; see also Figure 5 and Table S3). We also find evidence for differential selection within Europe, as suggested by TURCHIN *et al.* (2012), and to a somewhat lesser extent in the Middle East and Central Asia (Figure 5 and Table 3).

Skin Pigmentation We next analyzed data from a recent GWAS for skin pigmentation in an African-European admixed population of Cape Verdeans (BELEZA *et al.*, 2013), which identified four loci of major effect that explain approximately 35% of the variance in skin pigmentation in that population

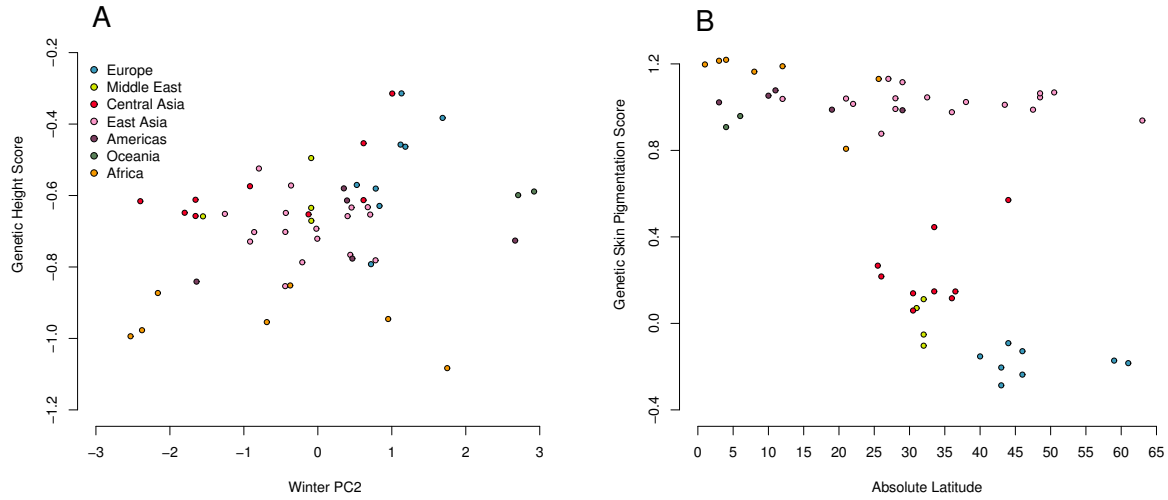


Figure 4: Estimated genetic height (A) and skin pigmentation score (B) plotted against winter PC2 and absolute latitude respectively. Both correlations are significant against the genome wide background after controlling for population structure (Table 2).

after controlling for admixture proportion. BELEZA *et al.* (2013) report effect sizes in units of modified melanin (MM) index, which is calculated as $100 \times \log(1/\% \text{melanin reflectance at } 650 \text{ nm})$, i.e. a higher MM index corresponds to darker skin, and a lower value to lighter skin.

We used these four loci to calculate a genetic skin pigmentation score in each of the HGDP populations. As expected, we identified a strong signal of excess variance among populations, as well as a strong correlation with latitude (Figure 4 and Table 2). Note, however, that this signal was driven entirely by the fact that populations of western Eurasian descent have a lower genetic skin pigmentation score than populations of African descent. Using only the markers from BELEZA *et al.* (2013), light skinned populations in East Asian and the Americas have a genetic skin pigmentation score that is almost as high as that of most African populations, an effect that is clearly visible when we plot the measured skin pigmentation and skin reflectance of HGDP populations (JABLONSKI and CHAPLIN, 2000; LAO *et al.*, 2007) against their genetic values (see Figures S12 and S13).

The correlation with latitude is thus weaker than one might expect, given the known phenotypic distribution of skin pigmentation in human populations (JABLONSKI and CHAPLIN, 2000, 2010). To illustrate this point further, we re-ran the analysis on a subsample of the HGDP consisting of populations from Europe, the Middle East, Central Asia, and Africa. In this subsample, the correlation with latitude, and signal of excess variance, was notably stronger ($r^2 = 0.2$, $p = 0.019$; $Q_X = 60.1$, $p = 8 \times 10^{-4}$).

This poor fit to observed skin pigmentation is due to the fact that we have failed to capture all

Table 2: Climate Correlations and Q_X statistics for all six phenotypes in the global analysis. We report $\text{sign}(\beta)r^2$, for the correlation statistics, such that they have an interpretation as the fraction of variance explained by the environmental variable, after removing that which is explained by the relatedness structure, with sign indicating the direction of the correlation. P-values are two-tailed for r^2 and upper tail for Q_X . Values for β and ρ are reported in Tables S15 and S16.

Phenotype	SUMPC1	SUMPC2	WINPC1	WINPC2	Latitude	Q_X
Height	-0.03 (0.21)	10^{-5} (0.99)	-0.008 (0.52)	0.086 (0.035)	0.009 (0.50)	86.9 (0.002)
Skin Pigmentation	0.061 (0.073)	0.003 (0.69)	0.048 (0.13)	-0.008 (0.51)	-0.085 (0.038)	79.1 (0.006)
Body Mass Index	-0.034 (0.19)	0.001 (0.82)	-0.022 (0.31)	0.044 (0.14)	0.031 (0.22)	67.2 (0.087)
Type 2 Diabetes	0.014 (0.40)	0.012 (0.45)	0.025 (0.27)	-0.006 (0.573)	-0.05 (0.11)	39.3 (0.902)
Crohn's Disease	0.07 (0.062)	-0.099 (0.022)	0.0001 (0.94)	-0.09 (0.039)	0.01 (0.55)	47.1 (0.68)
Ulcerative Colitis	0.03 (0.21)	-0.087 (0.034)	0.004 (0.67)	-0.049 (0.12)	0.01 (0.43)	48.58 (0.61)

of the loci that contribute to variation in skin pigmentation across the range of populations sampled, likely due to the partial convergent evolution of light skin pigmentation in Western and Eastern Eurasian populations (NORTON *et al.*, 2006). Including other loci putatively involved in skin pigmentation (MILLER *et al.*, 2007; EDWARDS *et al.*, 2010) decreases the estimated genetic pigmentation score of the other Eurasian populations (Figures S12 and S13), but we do not include these in our main analyses as they differ in ascertainment (and the role of KITLG in pigmentation variation has been contested by BELEZA *et al.*, 2013).

Within Africa, the San population has a decidedly lower genetic skin pigmentation score than any other HGDP African population. This is in accordance with the observation by JABLONSKI and CHAPLIN (2000) that the San are more lightly pigmented than other African populations represented by the HGDP and the observation that other putative light skin pigmentation alleles have higher frequency in the San than other African populations (NORTON *et al.*, 2006). A regional analysis of the six African populations alone identified a marginally significant correlation with latitude ($r^2 = 0.62$, $p = 0.0612$), and a signal of excess variance among populations ($Q_X = 16.19$, $p = 0.01$), suggesting a possible role for selection in shaping pigmentation variation within Africa.

Body Mass Index We next investigate two traits related to metabolic phenotypes (BMI and Type 2 diabetes), as there is a long history of adaptive hypotheses put forward to explain phenotypic variation among populations, with little conclusive evidence emerging thus far. We first focus on the set of 32 BMI associated SNPs identified by SPELIOTES *et al.* (2010) in their Table 1, which explain approximately 2-4% of the genetic variance for BMI. Of these 32 associated SNPs, 28 were present in the HGDP dataset, which we used to calculate a genetic BMI score for each HGDP population. We identified no significant signal of selection at the global level (Table 2).

Our regional level analysis indicated that the mean genetic BMI score is significantly lower than expected in East Asia ($Z = -2.48$, $p = 0.01$; see also Figure 5 and Table S7), while marginal Q_X statistics identify excess intraregional variation within East Asia and the Americas (Table 3).

Type 2 Diabetes We next investigated the 65 loci reported by MORRIS *et al.* (2012) as associated with T2D in their supplementary Table 2. Of these 65 SNPs, 61 were present in the HGDP dataset. We used effect sizes from stage 1 meta-analysis, and where MORRIS *et al.* (2012) reported a range of allele frequencies (due to differing sample frequencies among cohorts), we simply used the average. Where multiple SNPs were reported per locus we used the lead SNP from the combined meta-analysis. Also note that MORRIS *et al.* (2012) report effects in terms odds ratios (OR), which can be converted into additive effects by taking the logarithm (the same is true of the IBD data from JOSTINS *et al.*, 2012, analyzed below).

The distribution of genetic T2D risk scores showed no significant correlations with any of the five geographic axes we tested, and was in fact fairly underdispersed worldwide relative to the null expectation due to population structure (Table 2).

Our regional level analysis revealed that while T2D genetic risk is well explained by drift in Africa, Central and Eastern Asia, Oceania, and the Americas, European populations have far lower T2D genetic risk than expected ($Z = -2.79$, $p = 0.005$) and Middle Eastern populations a higher genetic risk than expected ($Z = 2.37$, $p = 0.018$). It's not necessarily clear, however, that these observations should be interpreted as evidence for selection either in Europe or the Middle East. While the dichotomous regional labels "Europe" vs. "Middle East" and genetic T2D risk explains the majority of the variance not accounted for by population structure ($r^2 = 0.77$, $p = 5 \times 10^{-4}$), this is essentially the same signal detected by the regional Z scores, and our Q_X statistic finds no signal of excess variance ($Q_X = 10.9$, $p = 0.48$) among the twelve HGDP populations in these two regions. Expanding to the next most closely related region, we tested for a signal of excess differentiation between Central Asia and either Europe or Middle East, but find no convincing signal in either case ($r^2 = 0.13$, $p = 0.21$; $Q_X = 12.0$, $p = 0.75$ and $r^2 = 0.15$, $p = 0.19$; $Q_X = 9.8$, $p = 0.63$ respectively).

To the extent that our results are consistent with an impact of selection on the genetic basis of T2D risk, they appear to be consistent only with a scenario in which selection has pushed the frequency of alleles that increase T2D risk up in Middle Eastern populations and down in European populations. This is an extremely subtle signal, which arises only after deep probing of the data, and as such we are skeptical that our results represent a meaningful signal of selection.

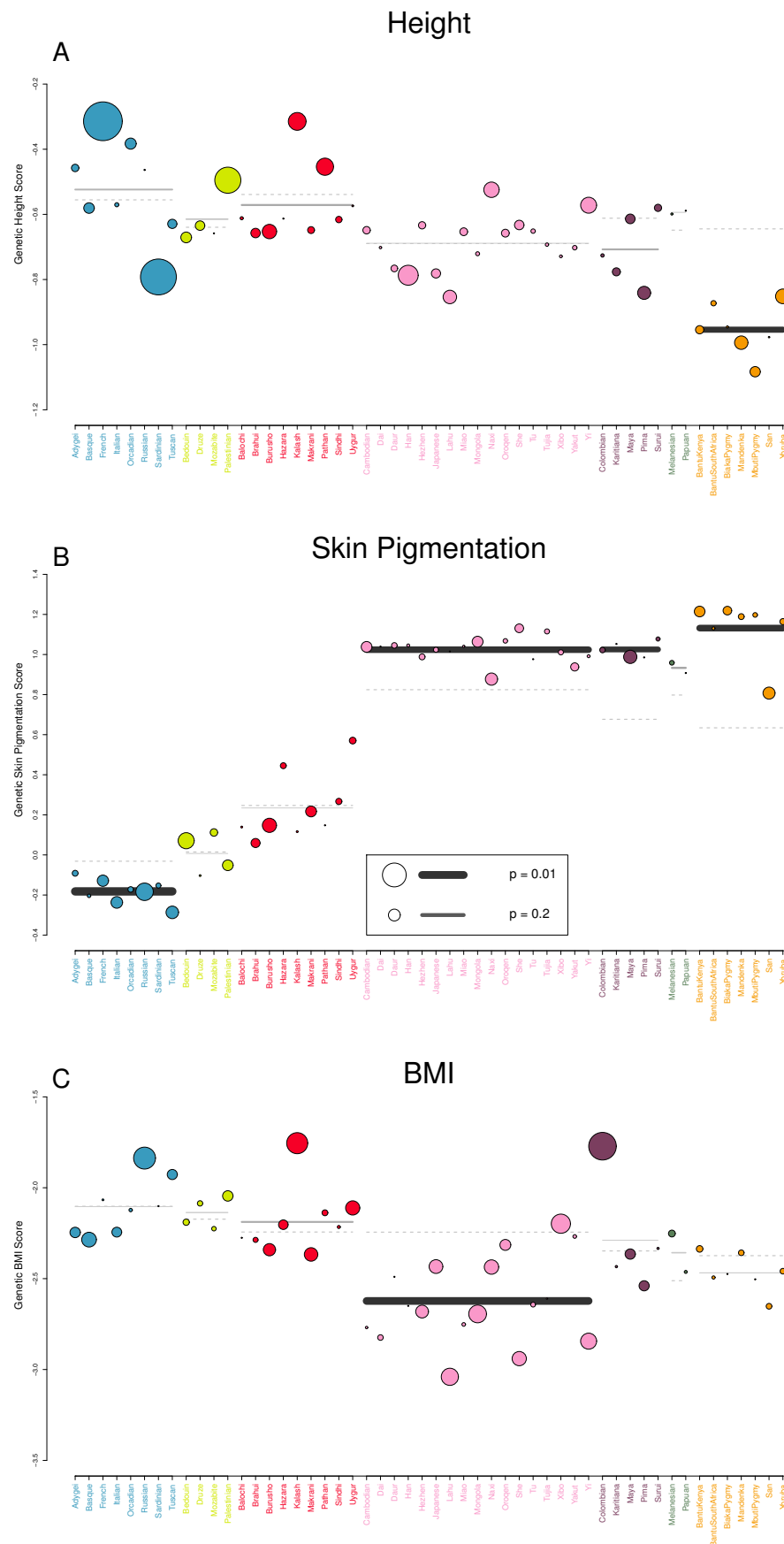
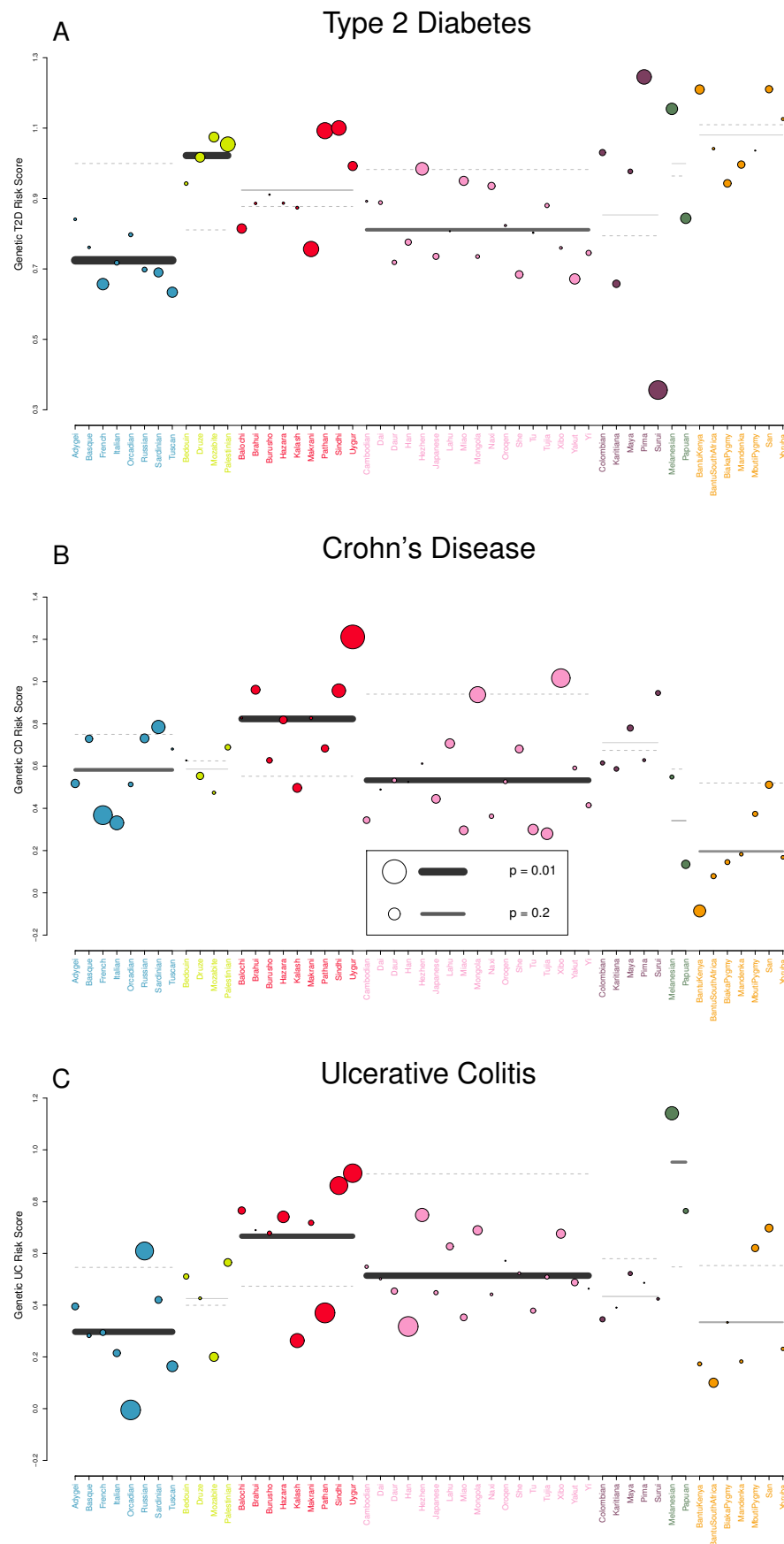


Figure 5: Visual representation of outlier analysis at the regional and individual population level for (A) Height, (B) Skin Pigmentation, and (C) Body Mass Index. For each geographic region we plot the expectation of the regional average, given the observed values in the rest of the dataset as a grey dashed line. The true regional average is plotted as a solid bar, with darkness and thickness proportional to the regional Z score. For each population we plot the observed value as a colored circle, with circle size proportional to the population specific Z score. For example, in (A), one can see that estimated genetic height is systematically lower than expected across Africa. Similarly, estimated genetic height is significantly higher (lower) in the French (Sardinian) population than expected, given the values observed for all other populations in the dataset.



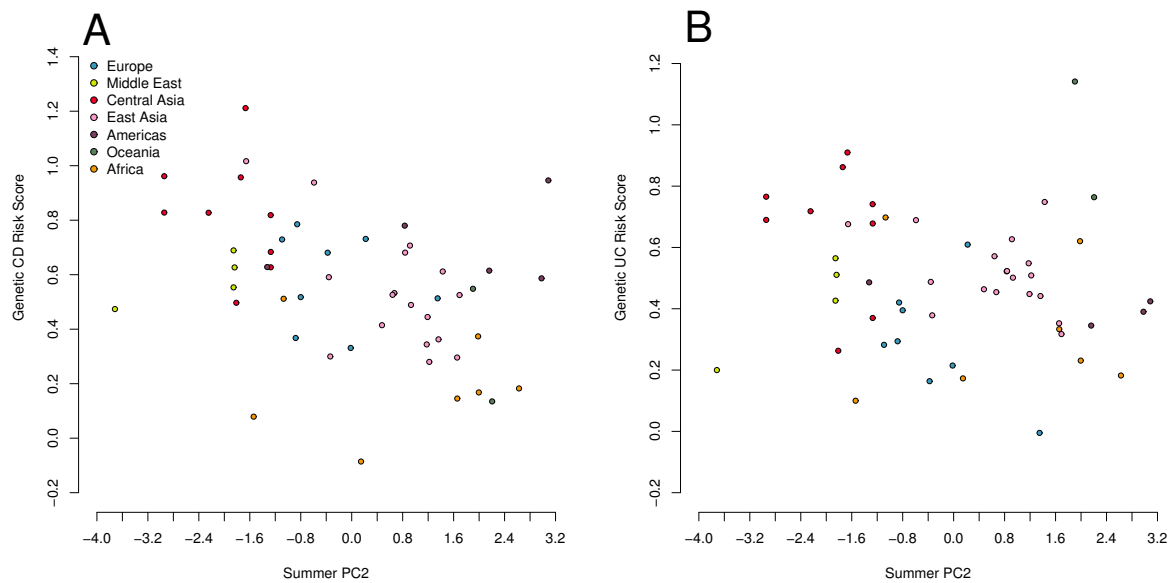


Figure 7: Estimated genetic risk score for Crohn's Disease (A) and Ulcerative Colitis (B) risk plotted against summer PC2. Both correlations are significant against the genome wide background after controlling for population structure (Table 2). Since a large proportion of SNPs underlying these traits are shared, we note that these results are not independent.

IBD Finally, we analyzed the set of associations reported by JOSTINS *et al.* (2012) for Crohn's Disease (CD) and Ulcerative Colitis (UC). Because CD and UC are closely connected phenotypes that share much of their genetic etiology, JOSTINS *et al.* (2012) used a likelihood ratio test of four different models (CD only, UC only, both CD and UC with equal effects on each, both CD and UC with independent effects) to distinguish which SNPs were associated with either or both phenotypes, and to assign effect sizes to SNPs (see their supplementary methods section 1d). We take these classifications at face value, resulting in two partially overlapping lists of 140 and 135 SNPs associated with CD and UC, of which 95 and 89 respectively were present in our HGD dataset.

We used these sets of SNPs to calculate genetic risk scores for CD and UC across the 52 HGD populations. Both CD and UC showed strong negative correlations with summer PC2 (Figure 7), while CD also showed a significant correlation with winter PC1, and a marginally significant correlation with summer PC1 (Table 2).

We did not observe any significant Q_X statistics for either trait, either at the global or the regional level, suggesting that our environmental correlation signals most likely arise from subtle differences between regions, as opposed to divergence among closely related populations. Indeed, we find moderate signals of regional level divergence in Europe (UC: $Z = -2.08, p = 0.04$), Central Asia (CD: $Z = 2.21, p = 0.03$), and East Asia (CD: $Z = -1.90, p = 0.06$ and UC: $Z = -2.12, p = 0.03$; see also Figure 6 and Tables S13 and S14).

3 Discussion

In this paper we have developed a powerful framework for identifying the influence of local adaptation on the genetic loci underlying variation in polygenic phenotypes. Below we discuss two major issues related to the application of such methods, namely the effect of the GWAS ascertainment scheme on our inference, and the interpretation of positive results.

Table 3: Q_X statistics and their empirical p-values for each of our six traits in each of the seven geographic regions delimited by ROSENBERG *et al.* (2002). The theoretical expected value of the statistic under neutrality for each region is equal to $M - 1$, where M is the number of populations in the region. We report the value of $M - 1$ next to each region for reference.

	Europe (7)	Middle East (3)	Central Asia (8)	East Asia (16)	Americas (4)	Oceania (1)	Africa (6)
Height	32.6 ($< 10^{-4}$)	7.3 (0.07)	15.5 (0.05)	18.2 (0.33)	4.2 (0.43)	0.007 (0.94)	5.4 (0.53)
Skin Pigmentation	9.7 (0.22)	9.6 (0.026)	23.4 (0.002)	13.8 (0.62)	1.3 (0.89)	0.38 (0.57)	16.2 (0.011)
Body Mass Index	9.1 (0.24)	1.6 (0.66)	9.3 (0.32)	28.4 (0.03)	13.1 (0.016)	1.2 (0.31)	1.9 (0.94)
Type 2 Diabetes	2.0 (0.96)	0.90 (0.83)	8.1 (0.43)	7.5 (0.96)	8.0 (0.13)	2.5 (0.15)	2.5 (0.88)
Crohn's Disease	6.6 (0.47)	0.87 (0.84)	7.56 (0.48)	15.5 (0.52)	1.3 (0.88)	2.5 (0.13)	2.6 (0.82)
Ulcerative Colitis	8.4 (0.30)	2.6 (0.48)	10.9 (0.21)	9.2 (0.907)	0.43 (0.986)	2.6 (0.12)	3.5 (0.77)

Ascertainment and Population Structure

Among the most significant potential pitfalls of our analysis (and the most likely cause of a false positive) is the fact that the loci used to test for the effect of selection on a given phenotype have been obtained through a GWAS ascertainment procedure, which can introduce false signals of selection if potential confounds are not properly controlled. We condition on simple features of the ascertainment process via our allele matching procedure, but deeper issues may arise from artifactual associations that result from the effects of population structure in the GWAS ascertainment panel. Given the importance of addressing this issue to the broader GWAS community, a range of well developed methods exist for doing GWAS in structured populations, and we refer the reader to the existing literature for a full discussion (FREEDMAN *et al.*, 2004; CAMPBELL *et al.*, 2005; PRICE *et al.*, 2006; KANG *et al.*, 2008; PRICE *et al.*, 2010; DIAO and CHEN, 2012; LIU *et al.*, 2013). Here, we focus on two related issues. First, the propensity of population structure in the GWAS ascertainment panel to generate false positives in our selection analysis, and second, the difficulties introduced by the sophisticated statistical approaches employed to deal with this issue when GWAS are done in strongly structured populations.

The problem of population structure arises generally when there is a correlation in the ascertainment panel between phenotype and ancestry (due to either genetic or environmental effects), such that SNPs that are ancestry informative will appear to be associated with the trait, even when no causal relationship exists (e.g. CAMPBELL *et al.*, 2005). To make matters worse, multiple false positive associations will often line up with same axis of population structure. If the populations being tested with our methods lie at least partially along the same axis of structure present in the GWAS ascertainment panel, then the ascertainment process will serve to generate the very signal of positive covariance among like effect alleles that our methods rely on to detect the signal of selection.

The primary takeaway from this observation is that the more diverse the array of individuals sampled for a given GWAS are with respect to ancestry, the greater the possibility that failing to control for population structure will generate false associations and hence false positives for our method.

What bearing do these complications have on our empirical results? The GWAS datasets we used can be divided into those conducted within populations of European descent and the skin pigmentation dataset (which used an admixed population). We will first discuss our analysis of the former.

The European GWAS loci we used were found in relatively homogeneous populations, in studies with rigorous standards for replication and control for population structure. Therefore, we are reasonably confident that these loci are true positives. Couple this with the fact that they were ascertained in (European) populations that are fairly homogenous relative to the global scale of our analyses, it is unlikely that population structure in the ascertainment panels is driving our positive signals. One might worry that we could still generate false signals by including European populations in our analysis, however many of the signals we see are driven by patterns outside of Europe (where the influence of structure within Europe should be much lessened). For height, where we do see a strong signal from within Europe, we use a set of loci that TURCHIN *et al.* (2012) have independently verified using a family based design that is immune to the effects of population structure.

We further note that for a number of our GWAS datasets, studies of non-European populations have replicated many of these loci (CHO *et al.*, 2009, 2011; VOIGHT *et al.*, 2010; KOONER *et al.*, 2011; N'DIAYE *et al.*, 2011; CARTY *et al.*, 2012; ?). More broadly, a recent study by MARIGORTA and NAVARRO (2013) found that replication rates for disease GWAS in European populations were quite high in East Asian populations, consistent with simple power considerations, an unlikely result if associations are driven

primarily by population structure. This suggests that, at least for GWAS done in relatively homogenous human populations, structure is unlikely to be a major confounding factor.

The issue of population structure may be more profound for our style of approach when GWAS are conducted using individuals from more strongly structured populations. In some cases it is desirable to conduct GWAS in such populations as locally adaptive alleles will be present at intermediate frequencies in these broader samples. A range of methods have been developed to adjust for population structure in these setting (KANG *et al.*, 2010; ZHOU and STEPHENS, 2012; LIU *et al.*, 2013). While generally effective in their goal, these methods present their own issues for our selection analysis. Consider the extreme case, such as that of ATWELL *et al.* (2010), who carried out a GWAS in *Arabidopsis thaliana* for 107 phenotypes across an array of 183 inbred lines of diverse geographical and ecological origin. ATWELL *et al.* (2010) used the genome-wide mixed model program EMMA (KANG *et al.*, 2008, 2010; ZHOU and STEPHENS, 2012) to control for the complex structure present in their ascertainment panel. This practice helps ensure that many of the identified associations are likely to be real, but also means that the loci found are likely to have unusual frequencies patterns across the species range. This follows from the fact that the loci identified as associated with the trait must stand out as being correlated with the trait in a way not predicted by the individual kinship matrix (as used by EMMA and other mixed model approaches). Our approach is predicated on the fact that we can use genome-wide patterns of kinship to adjust for population structure, but this correction is exactly the null model that loci significantly associated with phenotypes by mixed models have overcome. For this reason, both the theoretical χ^2 distribution of the Q_X statistic, as well as the empirical null distributions we construct from resampling, are likely to be inappropriate.

The Cape Verde skin pigmentation data we used could qualify as this second type of study. The Cape Verde population is an admixed population of African/European descent, and has substantial inter-individual variation in admixture proportion. Due to its admixed nature, the population segregates alleles which would not be at intermediate frequency in either parental population, and the four loci mapped in the study explain approximately 35% of the variance in skin pigmentation within the population (BELEZA *et al.*, 2013).

Despite the considerable population structure, the fact that interbreeding continues to mix genotypes in this population means that much of the LD due to the African/European population structure has been broken up (and the remaining LD is well predicted by an individual's genome-wide admixture coefficient). Population structure seems to have been well controlled for in this study, and a number of the loci have been replicated in independent admixed populations. While it is thus unlikely that the four loci we use are false associations, they could in principle suffer from the structured ascertainment issues described above, so it is unclear that the null distributions we use are strictly appropriate. That said, provided that BELEZA *et al.* (2013) have indeed appropriately controlled for population structure, under neutrality there would be no reason to expect that the correlation among the loci should be strongly positive with respect to the sign of their effect on the phenotype, and thus the pattern observed is at least consistent with a history of selection, especially in light of the multiple alternative lines of evidence for adaptation on the basis of skin pigmentation (JABLONSKI and CHAPLIN, 2000; SABETI *et al.*, 2007; LAO *et al.*, 2007; WILLIAMSON *et al.*, 2007; STURM, 2009; JABLONSKI and CHAPLIN, 2010).

Further work is needed to determine how best to modify the tests proposed herein to deal with GWAS performed in structured populations.

Complications of Interpretation

Our understanding of the genetic basis of most complex traits remains very incomplete, and as such the results of these analyses must be interpreted with care. That said, because our methods are based simply on the rejection of a neutral null model, an incomplete knowledge of the genetic basis of a given trait should only lead to a loss of statistical power, and not to a high false positive rate.

For all traits analyzed here except for skin pigmentation, the within population variance for genetic value is considerably larger than the variance between populations. This suggests that much of what we find is relatively subtle adaptation even on the level of the phenotype, and emphasizes the fact that for most genetic and phenotypic variation in humans, the majority of the variance is within populations rather than between populations (see Figures S14–S19). The influence of the environment likely plays a stronger role in the differences between populations for true phenotypes than the subtle differences we find here (as demonstrated by the rapid change in type 2 diabetes incidence with changing diet, e.g.

FRANCO *et al.*, 2013). That said, an understanding of how adaptation has shaped the genetic basis of a wide variety of phenotypes is clearly of interest, even if environmental differences dominate as the cause of present day population differences, as it informs our understanding of the biology and evolutionary history of these traits.

The larger conceptual issues relate to the interpretation of our positive findings, which we detail below. A number of these issues are inherent to the conceptual interpretation of evidence for local adaptation (KAWECKI and EBERT, 2004).

Environmental Correlations Significant correlation with environmental variables must be interpreted with caution. It is unlikely that the environmental variables included in our PC set will include the agent of selection and rather may simply be a correlate of the actual selection pressure. Further, given the large geographic and temporal scales considered here, it seems unlikely that many of our positive findings are the result of any single selection pressure, but rather may result from a combination of many different pressures, with perhaps different pressures acting in different geographic regions.

Consider also that we investigate only a limited set of phenotypes, which will each covary genetically with many other phenotypes. As such, any evidence for selection acting on the genetic basis of one of our phenotypes may represent selection on some correlated phenotype(s) not included in our study.

As an example, consider the fact that both UC and CD show a negative correlation with summer PC2, which include humidity and rainfall as primary components. However, it seems very unlikely that UC or CD have been the direct target of selection. It has been proposed that UC and CD might be understood as pleiotropic side effects of pathogen response (HANSEN *et al.*, 2010; VATN, 2012; CAGLIANI *et al.*, 2013), in which case the signal we see might correspond to local adaptation to variation in pathogen load/diversity (with humidity and/or rainfall as a possible predictor of these differences). Our results are likely to be compatible with many different hypotheses, however, and thus we consider this only one possible explanation.

Gene-by-Environment Interactions Gene by environment interactions may also complicate considerably the interpretation of a positive result. For example, we find evidence for adaptation in height, particularly in western Eurasia (consistent with the results of TURCHIN *et al.*, 2012), along with a worldwide correlation with winter PC2. However, our effect estimates are specific to a modern, northern European environment (although we note that the loci thus far identified in non European populations tend to overlap those discovered in Europe: CHO *et al.*, 2009; N'DIAYE *et al.*, 2011; CARTY *et al.*, 2012). It is possible that these alleles have experienced selection due to different effects in different environments, and so evidence for differentiation at the genetic level may not align with the observed phenotypic distribution. In an extreme case, different sets of alleles could be locally selected to maintain a constant phenotype across populations due to gene-by-environment interactions. Such a scenario could lead to a signal of local adaptation on a genetic level but no change in the phenotype across populations, a phenomenon known as countergradient variation (CONOVER and SCHULTZ, 1995).

It is also the case that GWAS loci as yet only explain a fraction of the total genetic variation for any given trait in any given environment. Because statistical power to detect the effect of a given locus is inversely related to its effect size, they are also likely those with the largest effect on the additive variance of a phenotype, and thus may also exhibit a great deal of pleiotropy. In this scenario, selection on a correlated trait could be responsible for the signals we observe, while loci not included in the GWAS set may evolve in a compensatory manner to maintain the same phenotypic optimum.

Missing variants The missing variant problem may also arise simply due to the interaction between demographic history and adaptation. A particularly dramatic example of this is offered by our analysis of skin pigmentation associated loci, where the alleles found by a GWAS in the Cape Verde population completely fail to predict the skin pigmentation of East Asians and Native Americans. This reflects the fact that the alleles responsible for light skin pigmentation in those populations are not variable in Cape Verde due to the independent and convergent adaptive evolution of light skin pigmentation in the ancestral East Asian population (NORTON *et al.*, 2006). It is unclear how prevalent convergent evolution by geographically separated populations is, but there are a number of examples (FRENCH-CONSTANT *et al.*, 2000; ARENDT and REZNICK, 2008; PEARCE *et al.*, 2009; NOVEMBRE and DI RIENZO, 2009), and theory

suggests that it might be surprisingly easy for traits with a more moderate number of loci contributing to variation (RALPH and COOP, 2010).

While such cases should not create a false signal of selection if only drift is involved, they do complicate the interpretation of positive signals. As an illustration, when we take a subset of all the Eurasian HGDP populations we see a significant correlation between genetic skin pigmentation score and longitude ($r^2 = 0.15, p = 0.015$), despite the fact that no such phenotypic correlation exists. While the wrong interpretation is easy to avoid here because we have a good understanding of the true phenotypic distribution, for the majority of GWAS studies such complications will be subtler and so care will have to be taken in their interpretation.

Loss of constraint and mutational pressure One further complication in the interpretation of our results is in how loss of constraint may play a role in driving apparent signals of local adaptation. Traits evolving under uniform stabilizing selection across all populations should be less variable than predicted by our covariance model of drift, due to negative covariances among loci, and so should be underrepresented in the extreme tails of our environmental correlation statistics and the upper tail of Q_X . As such, loss of constraint (i.e. weaker stabilizing selection in some populations than others), should not on its own create a signal of local adaptation. While the loci underpinning the phenotype can be subject to more drift in those populations, there is no systemic bias in the direction of this drift, and thus it should be well accommodated by our model. Loss of constraint, therefore, will not tend to create significant environmental correlations or systematic covariance between alleles of like effect.

An issue may arise, however, when loss of constraint is paired with biased mutational input (i.e. new mutations are more likely to push the phenotype in one direction than another; ZHANG and HILL, 2008) or if it is asymmetric (i.e. stabilizing selection is relaxed on one tail of the phenotypic distribution, but not the other). Under these two scenarios, mutations that push the phenotype in one particular direction would tend to drift up in frequency in populations with loss of constraint, creating systematic trends and positive covariance among like effect alleles at different loci, and resulting in a positive signal under our framework. While one would be mistaken to assume that the signal was necessarily that of recent positive directional selection, these scenarios do still imply that selection pressures on the genetic basis of the phenotype vary across space. Positive tests under our methods are thus fairly robust in being signals of differential selection among populations, but are themselves agnostic about the specific processes involved. Further work is needed to establish whether these scenarios can be distinguished from recent directional selection based on only allele frequencies and effect sizes, and as always, claims of recent adaptation should be supported by multiple lines of evidence beyond those provided by population genomics alone.

Limited evidence of adaptation on the genetic basis of variation in type 2 diabetes susceptibility Given the strong interest and history of investigation into whether selection has shaped phenotypic variation and disease risk, even the cases where we find limited evidence of local adaptation are potentially interesting. For example, a number of investigators have claimed that individual GWAS loci for Type 2 Diabetes show signals of selection (HELGASON *et al.*, 2007; HANCOCK *et al.*, 2008, 2010; KLIMENTIDIS *et al.*, 2010), a fact that is seen as support for the idea that genetic variation for type 2 diabetes risk has been shaped by local adaptation, potentially consistent with the thrifty genotype hypothesis (NEEL, 1962, and variants on this hypothesis). However, we find no evidence of environmental correlations and only a very subtle signal of significant differentiation for type 2 diabetes genetic risk in the Europe/Middle East comparison.

This suggests that local adaptation has not had a large role in shaping the present day world-wide distribution of type 2 diabetes susceptibility alleles (as mapped to date in Europe). One explanation of this discrepancy is that it seems biologically unrealistic that the phenotype of type 2 diabetes susceptibility would exhibit strong adaptive differentiation. Rather, local adaptation may be shaping some pleiotropically related phenotype (which shares only some of the loci involved). However, as seen in Figure 1, our methods have better power than single locus statistics, even when the genetic correlation between the GWAS phenotype and the phenotype under selection is weak. As such our results are more consistent with the idea that local adaptation has had little influence on the genetic basis of type 2 diabetes or a closely correlated phenotype, based on our current (admittedly poor) understanding of the genetic basis of this phenotype. This is not to say that some of the patterns seen at individual type 2

diabetes loci are not due to selection, but rather that it seems unlikely that they have been selected for due to their role in variation in a phenotype closely related to type 2 diabetes.

Future directions In this article we have focused on methods development and so have only just scratched the surface of the wide-range of populations and phenotypes that these types of methods could be applied to. Of particular interest is the possibility of applying these methods to GWAS performed in other species where the ecological determinants of local adaptation are better understood (ATWELL *et al.*, 2010; FOURNIER-LEVEL *et al.*, 2011).

One substantial difficulty with our approach, particularly in its application to other organisms, is that genome-wide association studies of highly polygenic phenotypes require very large sample sizes to map even a fraction of the total genetic variance. One promising way to partially sidestep this issue is by applying methods recently developed in animal breeding. In these genomic prediction/selection approaches, one does not attempt to map individual markers, but instead concentrates on predicting an individual's genetic value for a given phenotype using all markers simultaneously (MEUWISSEN *et al.*, 2001; HAYES *et al.*, 2009; MEUWISSEN *et al.*, 2013). This is accomplished by fitting simple linear models to genome-wide genotyping data, in principle allowing common SNPs to tag the majority of causal sites throughout the genome without attempting to explicitly identify them (see ZHOU *et al.*, 2013, for an overview of approaches). These methods have been applied to a range of species, including humans (YANG *et al.*, 2010; DAVIES *et al.*, 2011; YANG *et al.*, 2011; LEE *et al.*, 2012; DE LOS CAMPOS *et al.*, 2010, 2012, 2013), demonstrating that these predictions can potentially explain a relatively high fraction of the additive genetic variance within a population (and hence much of the total genetic variance). As these predictions are linear functions of genotypes, and hence allele frequencies, we might be able to predict the genetic values of sets of closely related populations for phenotypes of interest and apply very similar methods to those developed here. Such an approach may allow us to assess signals of local adaptation acting on the genetic basis of a large fraction of the total genetic variance for traits across a range of species. However, we note that if the only goal is to establish evidence for local adaptation in a given phenotype, then common garden experiments and traditional Q_{ST} -based analyses (or intermediate approaches which use genome wide data to estimate relationships among populations but still focus on phenotypic measurements as opposed to explicitly genetic estimates, e.g. OVASKAINEN *et al.*, 2011) may offer a simpler approach in species where such comparisons are possible.

As discussed in various places above, it is unlikely that all of the loci underpinning the genetic basis of a trait will have been subject to the same selection pressures, due to their differing roles in the trait and their pleiotropic effects. One potential avenue of future investigation is whether, given a large set of loci involved in a trait we can identify sets of loci in particular pathways or with a particular range of effect sizes that drive the signal of selection on the additive genetic basis of a trait.

Another promising extension of our approach is to deal explicitly with multiple correlated phenotypes. With the increasing number of GWAS efforts both empirical and methodological work are beginning to focus on understanding the shared genetic basis of various phenotypes. This raises the possibility that we may be able to disentangle the genetic basis of which phenotypes are more direct targets of selection, and which are responding to correlated selection on these direct targets (see KREMER *et al.*, 1997; BLOWS, 2007; CHENOWETH and BLOWS, 2008, for progress along these lines using Q_{ST}). Such tools may also offer a way of incorporating GxE interactions, as multiple GWAS for the same trait in different environments could be treated as correlated traits.

As association data for a greater variety of populations, species, and traits becomes available, we view the methods described out here, along with future extensions, as a productive way forward in developing a quantitative framework to explore the genetic and phenotypic basis of local adaptation.

4 Acknowledgements

We would like to thank Gideon Bradburd, Yaniv Brandvain, Luke Jostins, Chuck Langley, Joe Pickrell, Jonathan Pritchard, Peter Ralph, Jeff Ross-Ibarra, Alisa Sedghifar, Michael Turelli and Michael Whitlock for helpful discussion and/or comments on earlier versions of the manuscript.

5 Methods

Mean Centering and Covariance Matrix Estimation

Written in matrix notation, the procedure of mean centering the estimated genetic values and dropping one population from the analysis can be expressed as

$$\bar{\mathbf{Z}}' = \mathbf{T}\bar{\mathbf{Z}} \quad (16)$$

where \mathbf{T} is an $M - 1$ by M matrix with $\frac{M-1}{M}$ on the main diagonal, and $-\frac{1}{M}$ elsewhere.

In order to calculate the corresponding expected neutral covariance structure about this mean, we use the following procedure. Let \mathbf{G} be an M by K matrix, where each column is a vector of allele frequencies across the M populations at a particular SNP, randomly sampled from the genome according to the matching procedure described below. Let ϵ_k be the mean allele frequency in column k of \mathbf{G} , and let \mathbf{S} be a diagonal matrix such that $s_{kk} = \frac{1}{\epsilon_k(1-\epsilon_k)}$ and $s_{k \neq i} = 0$. With this data, we can estimate \mathbf{F} as

$$\mathbf{F} = \mathbf{T}\mathbf{G}\mathbf{S}\mathbf{G}^T\mathbf{T}^T. \quad (17)$$

This transformation performs the operation of centering the matrix at the mean value, and rooting the analysis with one population by dropping it from the covariance matrix (the same one we dropped from the vector of estimated genetic values), resulting in a covariance matrix describing the relationship of the remaining $M - 1$ populations. This procedure thus escapes the singularity introduced by centering the matrix at the observed mean of the sample.

Note that as we do not get to observe the population allele frequencies, the entries of \mathbf{G} are the sample frequencies at the randomly chosen loci, and thus the covariance matrix \mathbf{F} also includes the effect of finite sample size. Because the noise introduced by the sampling of individuals is uncorrelated across populations (in contrast to that introduced by drift and shared history), the primary effect is to inflate the diagonal entries of the matrix (see the supplementary material of ?, for discussion).

Standardized environmental variable Given a vector of environmental variable measurements for each population, we apply both the \mathbf{T} and Cholesky transformation as for the estimated genetic values

$$\mathbf{Y}' = \mathbf{C}^{-1}\mathbf{T}\mathbf{Y}. \quad (18)$$

This provides us with a set of $M - 1$ adjusted observations for the environmental variable which can be compared to the transformed genetic values for inference. This step is necessary as we have rotated the frame of reference of the estimated genetic values, and so we must do the same for the environmental variables to keep them both in a consistent reference frame.

Identifying Outliers with Conditional MVN Distributions

As described in the Results, we can use our multivariate normal model of relatedness to obtain the expected distribution of genetic values for an arbitrary set of populations, conditional on the observed values in some other arbitrary set.

We first partition our populations into two groups, those for which we want to obtain the expected distribution of genetic values (group 1), and those on which we condition in order to obtain this distribution (group 2). We then re-estimate the covariance matrix such that it is centered on the mean of group 2. This is a necessary step, as the amount of divergence between the populations in group 1 and the mean of group 2 will always be greater than the amount of divergence from the global mean, even under the neutral model, and our covariance matrix needs to reflect this fact in order to make accurate predictions. We can obtain this re-parameterized \mathbf{F} matrix as follows. If M is the total number of populations in the sample, then let q be the number of populations in group one, and let $M - q$ be the number of populations in group 2. We then define a new \mathbf{T}_R matrix such that the q columns corresponding the populations in group one have 1 on the diagonal, and 0 elsewhere, while the $M - q$ columns corresponding to group two have $\frac{M-q-1}{M-q}$ on the diagonal, and $-\frac{1}{M-q}$ elsewhere. We can then re-estimate a covariance matrix that is centered at the mean of the $M - q$ populations in group 2. Recalling our matrices \mathbf{G} and \mathbf{S} from (17), this matrix is calculated as

$$\mathbf{F}_R = \mathbf{T}_R\mathbf{G}\mathbf{S}\mathbf{G}^T\mathbf{T}_R^T \quad (19)$$

where we write \mathbf{F}_R to indicate that it is a covariance matrix that has been re-centered on the mean of group two.

Once we have calculated this re-centered covariance matrix, we can use well known results from multivariate normal theory to obtain the expected joint distribution of the genetic values for group one, conditional on the values observed in group two.

We partition our vector of genetic values and the re-centered covariance matrix such that

$$\vec{X} = \begin{bmatrix} \vec{X}_1 \\ \vec{X}_2 \end{bmatrix} \quad (20)$$

and

$$\mathbf{F}_R = \begin{bmatrix} \mathbf{F}_{11} & \mathbf{F}_{12} \\ \mathbf{F}_{21} & \mathbf{F}_{22} \end{bmatrix} \quad (21)$$

where \vec{X}_1 and \vec{X}_2 are vectors of genetic values in group 1 and 2 respectively, and \mathbf{F}_{11} , \mathbf{F}_{22} and $\mathbf{F}_{12} = \mathbf{F}_{21}^T$ are the marginal covariance matrices of populations within group 1, within group 2, and across the two groups, respectively. Letting $\mu_1 = \mu_2 = \frac{1}{M-q} \sum_{m=M-q}^M X_m$ (i.e. the sum of the elements of \vec{X}_2), we wish to obtain the distribution

$$\vec{X}_1 | \vec{X}_2, \mu_1, \mu_2 \sim MVN(\vec{\xi}, \mathbf{\Omega}), \quad (22)$$

where $\vec{\xi}$ and $\mathbf{\Omega}$ give the expected means and covariance structure of the populations in group 1, conditional on the values observed in group 2. These can be calculated as

$$\vec{\xi} = \mathbb{E}[\vec{X}_1 | \vec{X}_2, \mu_1, \mu_2] = \mu_1 \vec{1} + \mathbf{F}_{12} \mathbf{F}_{22}^{-1} (\vec{X}_2 - \mu_2 \vec{1}) \quad (23)$$

and

$$\mathbf{\Omega} = \text{Cov}[X_1 | X_2, \mu_1, \mu_2] = \mathbf{F}_{11} - \mathbf{F}_{12} \mathbf{F}_{22}^{-1} \mathbf{F}_{21}. \quad (24)$$

where the one vectors in line (23) are of length q and $M - q$ respectively.

This distribution is itself multivariate normal, and as such this framework is extremely flexible, as it allows us to obtain the expected joint distribution for arbitrary sets of populations (e.g. geographic regions or continents), or for each individual population. Further,

$$\mathbb{E} \left[\frac{1}{q} \sum_{m=1}^q X_m \mid \vec{\xi}, \mathbf{\Omega} \right] = \frac{1}{q} \sum_{m=1}^q \xi_m \quad (25)$$

and

$$\text{Var} \left[\frac{1}{q} \sum_{m=1}^q X_m \mid \vec{\xi}, \mathbf{\Omega} \right] = \frac{V_A}{q^2} \sum_{m=1}^q \sum_{n=1}^q \omega_{mn}. \quad (26)$$

where ω_{nm} denotes the elements of $\mathbf{\Omega}$. In words, the conditional expectation of the mean estimated genetic value across group 1 is equal to the mean of the conditional expectations, and its variance is equal to the mean value of the elements of the conditional covariance matrix. As such we can easily calculate a Z score (and corresponding p value) for group one as a whole as

$$Z = \frac{\frac{1}{q} \sum_{m=1}^q X_m - \frac{1}{q} \sum_{m=1}^q \xi_m}{\frac{1}{q} \sqrt{V_A \sum_{m=1}^q \sum_{n=1}^q \omega_{m,n}}}. \quad (27)$$

This Z score is a normal random variable with mean zero, variance one under the null hypothesis, and thus measures the divergence of the genetic values between the two populations relative to the null expectation under drift. Note that the observation of a significant Z score in a given population or region cannot necessarily be taken as evidence that selection has acted in that population or region, as selection in some of the populations on which we condition (especially the closely related ones) could be responsible for such a signal. As such, caution is warranted when interpreting the output of these sort of analyses, and is best done in the context of demographic history information (e.g. from TreeMix, ?, which uses the same basic model to draw inference about population history as we use to control for it and to generate conditional distributions).

The Linear Model at the Individual Locus Level

It is natural to ask about the relationship between our environmental correlation tests performed at the genetic value level, and tests that could be performed on individual SNPs and then aggregated.

As noted in (8), we can obtain for each underlying locus a set of transformed allele frequencies, which have passed through the same transformation as the estimated genetic values. We assume that each locus ℓ has a regression coefficient

$$\beta_\ell = \alpha_\ell \gamma \quad (28)$$

where γ is shared across all loci

$$p'_{m\ell} \sim \alpha_\ell \gamma Y'_m + e_{m\ell} \quad (29)$$

where the $e_{m\ell}$ are independent and identically distributed residuals. We can find the maximum likelihood estimate $\hat{\gamma}$ by treating $\alpha_\ell Y'_m$ as the linear predictor, and taking the regression of the combined vector $\vec{p'}$, across all populations and loci, on the combined vector $\vec{\alpha Y'}$. As such

$$\hat{\gamma} = \frac{Cov(\vec{p'}, \vec{\alpha Y'})}{Var(\vec{\alpha Y'})} \quad (30)$$

we can decompose this into a sum across loci

$$\hat{\gamma} = \frac{\frac{1}{L} \sum_\ell Cov(p'_\ell, \alpha_\ell Y')}{\frac{1}{L} \sum_\ell Var(\alpha_\ell Y')} = \frac{1}{\sum_\ell \alpha_\ell^2} \frac{\sum_\ell \alpha_\ell Cov(p'_\ell, Y')}{Var(Y')}. \quad (31)$$

As noted in (8), our transformed genetic values can be written as

$$X_m = 2 \sum_\ell \alpha_\ell p'_{m\ell} \quad (32)$$

and so the estimated slope ($\hat{\beta}$) of our regression ($X = \beta Y'$) is

$$\hat{\beta} = \frac{Cov(X, Y')}{Var(Y)} = \frac{\frac{2}{L} \sum_\ell \alpha_\ell Cov(p'_\ell, Y')}{Var(Y')} \quad (33)$$

Comparing these equations, the mean regression coefficient at the individual loci (31) and the regression coefficient of the estimated genetic values (33) are proportional to each other via a constant that is given by the sum over effect sizes (i.e. $\gamma = \frac{2}{\sum_\ell \alpha_\ell^2} \beta$), and thus the two approaches are equivalent.

HGDP data and imputation

We used imputed allele frequency data in the HGDP, where the imputation was performed as part of the phasing procedure of PICKRELL *et al.* (2009), as per the recommendations of CONRAD *et al.* (2006). We briefly recap their procedure here.

Phasing and imputation were done using fastPHASE (SCHEET and STEPHENS, 2006), with the settings that allow variation in the switch rate between subpopulations. The populations were grouped into subpopulations corresponding to the clusters identified in ROSENBERG *et al.* (2002). Haplotypes from the HapMap YRI and CEU populations were included as known, as they were phased in trios and are highly accurate. HapMap JPT and CHB genotypes were also included to help with the phasing.

Choosing null SNPs

Various components of our procedure involve sampling random sets of SNPs from across the genome. While we control for biases in our test statistics introduced by population structure through our \mathbf{F} matrix, we are also concerned that subtle ascertainment effects of the GWAS process could lead to biased test statistics, even under neutral conditions. We control for this possibility by sampling null SNPs so as to match the joint distribution of certain properties of the ascertained GWAS SNPs. Specifically, we were concerned that the minor allele frequency (MAF) in the ascertainment population, the imputation status of the allele in the HGDP datasets, and the background selection environment experienced at a

given locus, as measured by B value (MCVICKER *et al.*, 2009) might influence the distribution of allele frequencies across populations in ways that we could not predict.

We partitioned SNPs into a three way contingency table, with 25 bins for MAF (i.e. a bin size of 0.02), 2 bins for imputation (either imputed or not), and 10 bins for B value (B values range from 0 to 1, and thus our bin size was 0.1). As such, for each set of null genetic values, we sampled one null SNP from the same cell in the contingency table as each of the GWAS SNPs, and assigned this null SNP the effect size associated with the GWAS SNP it was sampled to match. While we do not assign effect sizes to sampled SNPs used to estimate the covariance matrix \mathbf{F} (instead simply scaling \mathbf{F} by a weighted sum of squared effect sizes, which is mathematically equivalent under our assumption that all SNPs have the same covariance matrix), we follow the same sampling procedure to ensure that \mathbf{F} describes the expected covariance structure of the GWAS SNPs.

For the skin pigmentation GWAS (BELEZA *et al.*, 2013) we do not have a good proxy present in the HGDP population, as the Cape Verdeans are an admixed population. Cape Verdeans are admixed with $\sim 59.53\%$ African ancestry, and 41.47% European ancestry in the sample obtained by BELEZA *et al.* (2013) (Beleza, pers. comm., April 8, 2013). As such, we estimated genome wide allele frequencies in Cape Verde by taking a weighted mean of the frequencies in the French and Yoruban populations of the HGDP, such that $p_{CV} = 0.5953p_Y + 0.4147p_F$. We then used these estimated frequencies to assign SNPs to frequency bins.

BELEZA *et al.* (2013) also used an admixture mapping strategy to map the genetic basis of skin pigmentation. However, if they had only mapped these loci in an admixture mapping setting we would have to condition our null model on having strong enough allele frequency differentiation between Africans and Europeans at the functional loci for admixture mapping to have power (REICH and PATTERSON, 2005). The fact that BELEZA *et al.* (2013) mapped these loci in a GWAS framework allows us to simply reproduce the strategy, and we ignore the results of the admixture mapping study (although we note that the loci and effect sizes estimated were similar). This highlights the need for a reasonably well defined ascertainment population for our approach, a point which we comment further on in the discussion.

References

- AKEY, J. M., 2009 Constructing genomic maps of positive selection in humans: Where do we go from here? *Genome Research*.
- ARENDT, J. and D. REZNICK, 2008 Convergence and parallelism reconsidered: what have we learned about the genetics of adaptation? *Trends in Ecology & Evolution*.
- ATWELL, S., Y. S. HUANG, B. J. VILHJ A LMSSON, G. WILLEMS, M. HORTON, Y. LI, D. MENG, A. PLATT, A. M. TARONE, T. T. HU, R. JIANG, N. W. MULIYATI, X. ZHANG, M. A. AMER, I. BAXTER, B. BRACHI, J. CHORY, C. DEAN, M. DEBIEU, J. DE MEAUX, J. R. ECKER, N. FAURE, J. M. KNISKERN, J. D. G. JONES, T. MICHAEL, A. NEMRI, F. ROUX, D. E. SALT, C. TANG, M. TODESCO, M. B. TRAW, D. WEIGEL, P. MARJORAM, J. O. BOREVITZ, J. BERGELSON, and M. NORDBORG, 2010 Genome-wide association study of 107 phenotypes in *Arabidopsis thaliana* inbred lines. *Nature*.
- BATESON, W., 1902 *Mendel's Principles of Heredity: A Defense*. Cambridge University Press.
- BEGUN, D. J. and C. F. AQUADRO, 1992 Levels of naturally occurring DNA polymorphism correlate with recombination rates in *D. melanogaster*. *Nature*.
- BELEZA, S., N. A. JOHNSON, S. I. CANDILLE, D. M. ABSHER, M. A. CORAM, J. LOPES, J. CAMPOS, I. I. ARAÚJO, T. M. ANDERSON, B. J. VILHJÁLMSSON, M. NORDBORG, A. CORREIA E SILVA, M. D. SHRIVER, J. ROCHA, G. S. BARSH, and H. TANG, 2013 Genetic Architecture of Skin and Eye Color in an African-European Admixed Population. *PLoS Genetics*.
- BERSAGLIERI, T., P. C. SABETI, and N. PATTERSON, 2004 ScienceDirect.com - The American Journal of Human Genetics - Genetic Signatures of Strong Recent Positive Selection at the Lactase Gene. . . . of Human Genetics.
- BLOOM, J. S., I. M. EHRENREICH, W. T. LOO, T.-L. V. LITE, and L. KRUGLYAK, 2013 Finding the sources of missing heritability in a yeast cross. *Nature*.
- BLOWS, M. W., 2007 A tale of two matrices: multivariate approaches in evolutionary biology. *Journal of Evolutionary Biology*.
- BONHOMME, M., C. CHEVALET, B. SERVIN, S. BOITARD, J. M. ABDALLAH, S. BLOTT, and M. SAN CRISTOBAL, 2010 Detecting Selection in Population Trees: The Lewontin and Krakauer Test Extended. *Genetics*.
- BRADBURD, G., P. RALPH, and G. COOP, 2013 Disentangling the effects of geographic and ecological isolation on genetic differentiation. *arXiv preprint arXiv:1302.3274*.
- CAGLIANI, R., U. POZZOLI, D. FORNI, A. CASSINOTTI, M. FUMAGALLI, M. GIANI, M. FICHERA, M. LOMBARDINI, S. ARDIZZONE, R. ASSELTA, R. DE FRANCHIS, S. RIVA, M. BIASIN, G. P. COMI, N. BRESOLIN, M. CLERICI, and M. SIRONI, 2013 Crohn's Disease Loci Are Common Targets of Protozoa-Driven Selection. *Molecular Biology and Evolution*.
- CAMPBELL, C. D., E. L. OGBURN, K. L. LUNETTA, H. N. LYON, M. L. FREEDMAN, L. C. GROOP, D. ALTSHULER, K. G. ARDLIE, and J. N. HIRSCHHORN, 2005 Demonstrating stratification in a European American population. *Nature Genetics*.
- CARTY, C. L., N. A. JOHNSON, C. M. HUTTER, A. P. REINER, U. PETERS, H. TANG, and C. KOOPERBERG, 2012 Genome-wide association study of body height in African Americans: the Women's Health Initiative SNP Health Association Resource (SHARe). *Human Molecular Genetics*.
- CASTO, A. M. and M. W. FELDMAN, 2011 Genome-Wide Association Study SNPs in the Human Genome Diversity Project Populations: Does Selection Affect Unlinked SNPs with Shared Trait Associations? *PLoS Genetics*.
- CAVALLI-SFORZA, L. L., I. BARRAI, and A. W. F. EDWARDS, 1964 Analysis of Human Evolution Under Random Genetic Drift. *Cold Spring Harbor Symposia on Quantitative Biology*.

- CHARLESWORTH, B., M. NORDBORG, and D. CHARLESWORTH, 1997 The effects of local selection, balanced polymorphism and background selection on equilibrium patterns of genetic diversity in subdivided populations. *Genetics Research*.
- CHEN, H., N. PATTERSON, and D. REICH, 2010 Population differentiation as a test for selective sweeps. *Genome Research*.
- CHENOWETH, S. F. and M. W. BLOWS, 2008 Q STMEETS THE G MATRIX: THE DIMENSIONALITY OF ADAPTIVE DIVERGENCE IN MULTIPLE CORRELATED QUANTITATIVE TRAITS. *Evolution*.
- CHO, Y.-S., C.-H. CHEN, C. HU, J. LONG, R. T.-H. ONG, X. SIM, F. TAKEUCHI, Y. WU, M.-J. GO, T. YAMAUCHI, Y.-C. CHANG, S. H. KWAK, R. C. W. MA, K. YAMAMOTO, L. S. ADAIR, T. AUNG, Q. CAI, L.-C. CHANG, Y.-T. CHEN, Y. GAO, F. B. HU, H.-L. KIM, S. KIM, Y. J. KIM, J. J.-M. LEE, N. R. LEE, Y. LI, J.-J. LIU, W. LU, J. NAKAMURA, E. NAKASHIMA, D. P.-K. NG, W. T. TAY, F.-J. TSAI, T.-Y. WONG, M. YOKOTA, W. ZHENG, R. ZHANG, C. WANG, W. Y. SO, K. OHNAKA, H. IKEGAMI, K. HARA, Y. M. CHO, N. H. CHO, T.-J. CHANG, Y. BAO, Å. K. HEDMAN, A. P. MORRIS, M. I. MCCARTHY, R. TAKAYANAGI, K. S. PARK, W. JIA, L.-M. CHUANG, J. C. N. CHAN, S. MAEDA, T. KADOWAKI, J.-Y. LEE, J.-Y. WU, Y.-Y. TEO, E. S. TAI, X.-O. SHU, K. L. MOHLKE, N. KATO, B.-G. HAN, and M. SEIELSTAD, 2011 Meta-analysis of genome-wide association studies identifies eight new loci for type 2 diabetes in east Asians. *Nature Publishing Group*.
- CHO, Y.-S., M.-J. GO, Y. J. KIM, J. Y. HEO, J. H. OH, H.-J. BAN, D. YOON, M. H. LEE, D.-J. KIM, M. PARK, S.-H. CHA, J.-W. KIM, B.-G. HAN, H. MIN, Y. AHN, M. S. PARK, H. R. HAN, H.-Y. JANG, E. Y. CHO, J.-E. LEE, N. H. CHO, C. SHIN, T. PARK, J. W. PARK, J.-K. LEE, L. CARDON, G. CLARKE, M. I. MCCARTHY, J.-Y. LEE, J.-K. LEE, B. OH, and H.-L. KIM, 2009 A large-scale genome-wide association study of Asian populations uncovers genetic factors influencing eight quantitative traits. *Nature Genetics*.
- CONOVER, D. O. and E. T. SCHULTZ, 1995 Phenotypic similarity and the evolutionary significance of countergradient variation. *Trends in Ecology & Evolution*.
- CONRAD, D. F., M. JAKOBSSON, G. COOP, X. WEN, J. D. WALL, N. A. ROSENBERG, and J. K. PRITCHARD, 2006 A worldwide survey of haplotype variation and linkage disequilibrium in the human genome. *Nature Genetics*.
- COOP, G., J. K. PICKRELL, J. NOVEMBRE, S. KUDARAVALLI, J. LI, D. ABSHER, R. M. MYERS, L. L. CAVALLI-SFORZA, M. W. FELDMAN, and J. K. PRITCHARD, 2009 The role of geography in human adaptation. *PLoS Genetics*.
- COOP, G., D. WITONSKY, A. DI RIENZO, and J. K. PRITCHARD, 2010 Using environmental correlations to identify loci underlying local adaptation. *Genetics*.
- CORONA, E., R. CHEN, M. SIKORA, A. A. MORGAN, C. J. PATEL, A. RAMESH, C. D. BUSTAMANTE, and A. J. BUTTE, 2013 Analysis of the Genetic Basis of Disease in the Context of Worldwide Human Relationships and Migration. *PLoS Genetics*.
- CROW, J. F., 2010 On epistasis: why it is unimportant in polygenic directional selection. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences*.
- DARBISHIRE, A. D., 1905 On the Supposed Antagonism of Mendelian to Biometric Theories of Heredity. *Manchester Memoirs*.
- DAVIES, G., A. TENESA, A. PAYTON, J. YANG, S. E. HARRIS, D. LIEWALD, X. KE, S. LE HELLARD, A. CHRISTOFOROU, M. LUCIANO, K. MCGHEE, L. LOPEZ, A. J. GOW, J. CORLEY, P. REDMOND, H. C. FOX, P. HAGGARTY, L. J. WHALLEY, G. MCNEILL, M. E. GODDARD, T. ESPESETH, A. J. LUNDERVOLD, I. REINVANG, A. PICKLES, V. M. STEEN, W. OLLIER, D. J. PORTEOUS, M. HORAN, J. M. STARR, N. PENDLETON, P. M. VISSCHER, and I. J. DEARY, 2011 Genome-wide association studies establish that human intelligence is highly heritable and polygenic. *Molecular Psychiatry*.

- DE LOS CAMPOS, G., D. GIANOLA, and D. B. ALLISON, 2010 PersPectives. Nature Publishing Group.
- DE LOS CAMPOS, G., Y. C. KLIMENTIDIS, A. I. VAZQUEZ, and D. B. ALLISON, 2012 Prediction of Expected Years of Life Using Whole-Genome Markers. PLoS ONE.
- DE LOS CAMPOS, G., A. I. VAZQUEZ, R. FERNANDO, Y. C. KLIMENTIDIS, and D. SORESENSEN, 2013 Prediction of Complex Human Traits Using the Genomic Best Linear Unbiased Predictor. PLoS Genetics.
- DIAO, L. and K. C. CHEN, 2012 Local Ancestry Corrects for Population Structure in *Saccharomyces cerevisiae* Genome-Wide Association Studies. Genetics.
- EDWARDS, M., A. BIGHAM, J. TAN, S. LI, A. GOZDZIK, K. ROSS, L. JIN, and E. J. PARRA, 2010 Association of the OCA2 Polymorphism His615Arg with Melanin Content in East Asian Populations: Further Evidence of Convergent Evolution of Skin Pigmentation. PLoS Genetics.
- FALCONER, D., 1960 *Introduction to Quantitative Genetics*. Ronald Press.
- FARIELLO, M. I., S. BOITARD, H. NAYA, M. SANCRISTOBAL, and B. SERVIN, 2013 Detecting Signatures of Selection Through Haplotype Differentiation Among Hierarchically Structured Populations. Genetics.
- FELSENSTEIN, J., 1982 How can we infer geography and history from gene frequencies? Journal of Theoretical Biology.
- FFRENCH-CONSTANT, R. H., N. ANTHONY, K. ARONSTEIN, T. ROCHELEAU, and G. STILWELL, 2000 Cyclodiene Insecticide Resistance: From Molecular to Population Genetics. Annual Review of Entomology.
- FISHER, R., 1930 *The Genetical Theory of Natural Selection*. Clarendon Press.
- FISHER, R. A., 1918 XV.—The Correlation between Relatives on the Supposition of Mendelian Inheritance. Transactions of the Royal Society of Edinburgh.
- FOURNIER-LEVEL, A., A. KORTE, M. D. COOPER, M. NORDBORG, J. SCHMITT, and A. M. WILCZEK, 2011 A Map of Local Adaptation in *Arabidopsis thaliana*. Science.
- FRANCO, M., U. BILAL, P. ORDUNEZ, M. BENET, A. MOREJON, B. CABALLERO, J. F. KENNELLY, and R. S. COOPER, 2013 Population-wide weight loss and regain in relation to diabetes burden and cardiovascular mortality in Cuba 1980-2010: repeated cross sectional surveys and ecological comparison of secular trends. BMJ.
- FREEDMAN, M. L., D. REICH, K. L. PENNEY, G. J. McDONALD, A. A. MIGNAULT, N. PATTERSON, S. B. GABRIEL, E. J. TOPOL, J. W. SMOLLER, C. N. PATO, M. T. PATO, T. L. PETRYSHEN, L. N. KOLONEL, E. S. LANDER, P. SKLAR, B. HENDERSON, J. N. HIRSCHHORN, and D. ALTSHULER, 2004 Assessing the impact of population stratification on genetic association studies. Nature Genetics.
- FUJIMOTO, A., R. KIMURA, J. OHASHI, K. OMI, R. YULIWULANDARI, L. BATUBARA, M. S. MUSTOFA, U. SAMAKKARN, W. SETTHEETHAM-ISHIDA, T. ISHIDA, Y. MORISHITA, T. FURUSAWA, M. NAKAZAWA, R. OHTSUKA, and K. TOKUNAGA, 2008 A scan for genetic determinants of human hair morphology: EDAR is associated with Asian hair thickness.
- GUILLOT, G., 2012 Detection of correlation between genotypes and environmental variables. A fast computational approach for genomewide studies. arXiv preprint arXiv:1206.0889.
- GÜNTHER, T. and G. COOP, 2012 Robust identification of local adaptation from allele frequencies. arXiv preprint arXiv:1009.3150.
- GUSTAFSSON, A. and P. LINDENFORS, 2004 Human size evolution: no evolutionary allometric relationship between male and female stature. Journal of Human Evolution.
- GUSTAFSSON, A. and P. LINDENFORS, 2009 Latitudinal patterns in human stature and sexual stature dimorphism. Annals of Human Biology.

- HANCOCK, A. M., D. B. WITONSKY, G. ALKORTA-ARANBURU, C. M. BEALL, A. GEBREMEDHIN, R. SUKERNIK, G. UTERMANN, J. K. PRITCHARD, G. COOP, and A. DI RIENZO, 2011 Adaptations to Climate-Mediated Selective Pressures in Humans. *PLoS Genetics*.
- HANCOCK, A. M., D. B. WITONSKY, E. EHLE, G. ALKORTA-ARANBURU, C. BEALL, A. GEBREMEDHIN, R. SUKERNIK, G. UTERMANN, J. PRITCHARD, and G. COOP, 2010 Human adaptations to diet, subsistence, and ecoregion are due to subtle shifts in allele frequency. *Proceedings of the National Academy of Sciences*.
- HANCOCK, A. M., D. B. WITONSKY, A. S. GORDON, G. ESHEL, J. K. PRITCHARD, G. COOP, and A. DI RIENZO, 2008 Adaptations to Climate in Candidate Genes for Common Metabolic Disorders. *PLoS Genetics*.
- HANSEN, R., J. M. THOMSON, E. M. EL-OMAR, and G. L. HOLD, 2010 The role of infection in the aetiology of inflammatory bowel disease. *Journal of Gastroenterology*.
- HAYES, B. J., P. J. BOWMAN, A. J. CHAMBERLAIN, and M. E. GODDARD, 2009 Invited review. *Journal of Dairy Science*.
- HELGASON, A., S. PÁLSSON, G. THORLEIFSSON, S. F. A. GRANT, V. EMILSSON, S. GUNNARSDOTTIR, A. ADEYEMO, Y. CHEN, G. CHEN, I. REYNISDOTTIR, R. BENEDIKTSSON, A. HINNEY, T. HANSEN, G. ANDERSEN, K. BORCH-JOHNSON, T. JØRGENSEN, H. SCHÄFER, M. FARUQUE, A. DOUMATEY, J. ZHOU, R. L. WILENSKY, M. P. REILLY, D. J. RADER, Y. BAGGER, C. CHRISTIANSEN, G. SIGURDSSON, J. HEBEBRAND, O. PEDERSEN, U. THORSTEINSDOTTIR, J. R. GULCHER, A. KONG, C. ROTIMI, and K. STEFANSSON, 2007 Refining the impact of TCF7L2 gene variants on type 2 diabetes and adaptive evolution. *Nature Genetics*.
- HERNANDEZ, R. D., J. L. KELLEY, E. ELYASHIV, S. C. MELTON, A. AUTON, G. MCVEAN, G. SELLA, and M. PRZEWORSKI, 2011 Classic selective sweeps were rare in recent human evolution. *Science* (New York, N.Y.).
- HILL, W. G., M. E. GODDARD, and P. M. VISSCHER, 2008 Data and Theory Point to Mainly Additive Genetic Variance for Complex Traits. *PLoS Genetics*.
- HUDSON, R. R., M. KREITMAN, and M. AGUADÉ, 1987 A test of neutral molecular evolution based on nucleotide data. *Genetics*.
- JABLONSKI, N. G. and G. CHAPLIN, 2000 The evolution of human skin coloration. *Journal of Human Evolution*.
- JABLONSKI, N. G. and G. CHAPLIN, 2010 Colloquium Paper: Human skin pigmentation as an adaptation to UV radiation. *Proceedings of the National Academy of Sciences of the United States of America*.
- JOSTINS, L., S. RIPKE, R. K. WEERSMA, R. H. DUERR, D. P. MCGOVERN, K. Y. HUI, J. C. LEE, L. P. SCHUMM, Y. SHARMA, and C. A. ANDERSON, 2012 Host-microbe interactions have shaped the genetic architecture of inflammatory bowel disease. *Nature*.
- KANG, H. M., J. H. SUL, S. K. SERVICE, N. A. ZAITLEN, S.-Y. KONG, N. B. FREIMER, C. SABATTI, and E. ESKIN, 2010 technical reports. *Nature Genetics*.
- KANG, H. M., N. A. ZAITLEN, C. M. WADE, A. KIRBY, D. HECKERMAN, M. J. DALY, and E. ESKIN, 2008 Efficient Control of Population Structure in Model Organism Association Mapping. *Genetics*.
- KAPLAN, N. L., R. R. HUDSON, and C. H. LANGLEY, 1989 The "hitchhiking effect" revisited. *Genetics*.
- KAWECKI, T. J. and D. EBERT, 2004 Conceptual issues in local adaptation. *Ecology Letters*.
- KINGSOLVER, J. G., H. E. HOEKSTRA, J. M. HOEKSTRA, D. BERRIGAN, S. N. VIGNIERI, C. E. HILL, A. HOANG, P. GIBERT, and P. BEERLI, 2001 The Strength of Phenotypic Selection in Natural Populations. *The American Naturalist*.

- KLIMENTIDIS, Y. C., M. ABRAMS, J. WANG, J. R. FERNANDEZ, and D. B. ALLISON, 2010 Natural selection at genomic regions associated with obesity and type-2 diabetes: East Asians and sub-Saharan Africans exhibit high levels of differentiation at type-2 diabetes regions. *Human Genetics*.
- KOONER, J. S., D. SALEHEEN, X. SIM, J. SEHMI, W. ZHANG, P. FROSSARD, L. F. BEEN, K.-S. CHIA, A. S. DIMAS, N. HASSANALI, T. JAFAR, J. B. M. JOWETT, X. LI, V. RADHA, S. D. REES, F. TAKEUCHI, R. YOUNG, T. AUNG, A. BASIT, M. CHIDAMBARAM, D. DAS, E. GRUNDBERG, Å. K. HEDMAN, Z. I. HYDRIE, M. ISLAM, C.-C. KHOR, S. KOWLESSUR, M. M. KRISTENSEN, S. LIJU, W.-Y. LIM, D. R. MATTHEWS, J. LIU, A. P. MORRIS, A. C. NICA, J. M. PINIDIYAPATHIRAGE, I. PROKOPENKO, A. RASHEED, M. SAMUEL, N. SHAH, A. S. SHERA, K. S. SMALL, C. SUO, A. R. WICKREMasinghe, T.-Y. WONG, M. YANG, F. ZHANG, G. R. ABECASIS, A. H. BARNETT, M. CAULFIELD, P. DELOUKAS, T. M. FRAYLING, P. FROGUEL, N. KATO, P. KATULANDA, M. A. KELLY, J. LIANG, V. MOHAN, D. K. SANGHERA, J. SCOTT, M. SEIELSTAD, P. Z. ZIMMET, P. ELLIOTT, Y.-Y. TEO, M. I. MCCARTHY, J. DANESH, E. S. TAI, and J. C. CHAMBERS, 2011 Genome-wide association study in individuals of South Asian ancestry identifies six new type 2 diabetes susceptibility loci. *Nature Publishing Group*.
- KREMER, A. and V. LE CORRE, 2011 Decoupling of differentiation between traits and their underlying genes in response to divergent selection. *Heredity*.
- KREMER, A., A. ZANETTO, and A. DUCOUSO, 1997 Multilocus and multitrait measures of differentiation for gene markers and phenotypic traits. *Genetics*.
- LANDE, R., 1992 Neutral theory of quantitative genetic variance in an island model with local extinction and colonization. *Evolution*.
- LANGO ALLEN, H., K. ESTRADA, G. LETTRE, S. I. BERNDT, M. N. WEEDON, F. RIVADENEIRA, C. J. WILLER, A. U. JACKSON, S. VEDANTAM, S. RAYCHAUDHURI, T. FERREIRA, A. R. WOOD, R. J. WEYANT, A. V. SEGRE, E. K. SPELIOTES, E. WHEELER, N. SORANZO, J.-H. PARK, J. YANG, D. GUDBJARTSSON, N. L. HEARD-COSTA, J. C. RANDALL, L. QI, A. VERNON SMITH, R. MAGI, T. PASTINEN, L. LIANG, I. M. HEID, J. LUAN, G. THORLEIFSSON, T. W. WINKLER, M. E. GODDARD, K. SIN LO, C. PALMER, T. WORKALEMAHU, Y. S. AULCHENKO, Å. JOHANSSON, M. C. ZILLIKENS, M. F. FEITOSA, T. O. N. ESKO, T. JOHNSON, S. KETKAR, P. KRAFT, M. MANGINO, I. PROKOPENKO, D. ABSHER, E. ALBRECHT, F. ERNST, N. L. GLAZER, C. HAYWARD, J.-J. HOTTENGA, K. B. JACOBS, J. W. KNOWLES, Z. A. N. KUTALIK, K. L. MONDA, O. POLASEK, M. PREUSS, N. W. RAYNER, N. R. ROBERTSON, V. STEINTHORSDDOTTIR, J. P. TYRER, B. F. VOIGHT, F. WIKLUND, J. XU, J. H. ZHAO, D. R. NYHOLT, N. PELLIKKA, M. PEROLA, J. R. B. PERRY, I. SURAKKA, M.-L. TAMMESOO, E. L. ALTMAIER, N. AMIN, T. ASPELUND, T. BHANGALE, G. BOUCHER, D. I. CHASMAN, C. CHEN, L. COIN, M. N. COOPER, A. L. DIXON, Q. GIBSON, E. GRUNDBERG, K. HAO, M. JUHANI JUNTILA, L. M. KAPLAN, J. KETTUNEN, I. R. KONIG, T. KWAN, R. W. LAWRENCE, D. F. LEVINSON, M. LORENTZON, B. MCKNIGHT, A. P. MORRIS, M. MULLER, J. SUH NGWA, S. PURCELL, S. RAFELT, R. M. SALEM, E. SALVI, S. SANNA, J. SHI, U. SOVIO, J. R. THOMPSON, M. C. TURCHIN, L. VANDENPUT, D. J. VERLAAN, V. VITART, C. C. WHITE, A. ZIEGLER, P. ALMGREN, A. J. BALMFORTH, H. CAMPBELL, L. CITTERIO, A. DE GRANDI, A. DOMINICZAK, J. DUAN, P. ELLIOTT, R. ELOSUA, J. G. ERIKSSON, N. B. FREIMER, E. J. C. GEUS, N. GLORIOSO, S. HAIQING, A.-L. HARTIKAINEN, A. S. HAVULINNA, A. A. HICKS, J. HUI, W. IGL, T. ILLIG, A. JULA, E. KAJANTIE, T. O. KILPELAINEN, M. KOIRANEN, I. KOLCIC, S. KOSKINEN, P. KOVACS, J. LAITINEN, J. LIU, M.-L. LOKKI, A. MARUSIC, A. MASCHIO, T. MEITINGER, A. MULAS, G. PARE, A. N. PARKER, J. F. PEDEN, A. PETERSMANN, I. PICHLER, K. H. PIETILAINEN, A. POUTA, M. RIDDERSTRAALE, J. I. ROTTER, J. G. SAMBROOK, A. R. SANDERS, C. O. SCHMIDT, J. SINISALO, J. H. SMIT, H. M. STRINGHAM, G. BRAGI WALTERS, E. WIDEN, S. H. WILD, G. WILLEMSSEN, L. ZAGATO, L. ZGAGA, P. ZITTING, H. ALAVERE, M. FARRALL, W. L. MCARDLE, M. NELIS, M. J. PETERS, S. RIPATTI, J. B. J. VAN MEURS, K. K. ABEN, K. G. ARDLIE, J. S. BECKMANN, J. P. BEILBY, R. N. BERGMAN, S. BERGMANN, F. S. COLLINS, D. CUSI, M. DEN HEIJER, G. EIRIKSDOTTIR, P. V. GEJMAN, A. S. HALL, A. HAMSTEN, H. V. HUIKURI, C. IRIBARREN, M. KAHONEN, J. KAPRIO, S. KATHIRESAN, L. KIEMENEY, T. KOCHER, L. J. LAUNER, T. LEHTIMAKI, O. MELANDER, T. H. MOSLEY, A. W. MUSK, M. S. NIEMINEN, C. J. O'DONNELL, C. OHLSSON, B. OOSTRA, L. J. PALMER, O. RAITAKARI, P. M. RIDKER, J. D. RIOUX,

- A. RISSANEN, C. RIVOLTA, H. SCHUNKERT, A. R. SHULDINER, D. S. SISCOVICK, M. STUMVOLL, A. T O NJES, J. TUOMILEHTO, G.-J. VAN OMMEN, J. VIHKARI, A. C. HEATH, N. G. MARTIN, G. W. MONTGOMERY, M. A. PROVINCE, M. KAYSER, A. M. ARNOLD, L. D. ATWOOD, E. BOERWINKLE, S. J. CHANOCK, P. DELOUKAS, C. GIEGER, H. GR O NBERG, P. HALL, and HATTE..., 2010 Hundreds of variants clustered in genomic loci and biological pathways affect human height. *Nature*.
- LAO, O., J. M. DE GRUIJTER, K. VAN DUIJN, A. NAVARRO, and M. KAYSER, 2007 Signatures of Positive Selection in Genes Associated with Human Skin Pigmentation as Revealed from Analyses of Single Nucleotide Polymorphisms. *Annals of Human Genetics*.
- LATTA, R. G., 1998 Differentiation of allelic frequencies at quantitative trait loci affecting locally adaptive traits. *American Naturalist*.
- LATTA, R. G., 2003 Gene flow, adaptive population divergence and comparative population structure across loci. *New Phytologist*.
- LE CORRE, V. and A. KREMER, 2003 Genetic variability at neutral markers, quantitative trait loci and trait in a subdivided population under selection. *Genetics*.
- LE CORRE, V. and A. KREMER, 2012 The genetic differentiation at quantitative trait loci under local adaptation. *Molecular Ecology*.
- LEE, S. H., T. R. DECANDIA, S. RIPKE, J. YANG, P. F. SULLIVAN, M. E. GODDARD, M. C. KELLER, P. M. VISSCHER, and N. R. WRAY, 2012 Estimating the proportion of variation in susceptibility to schizophrenia captured by common SNPs. *Nature Genetics*.
- LEWONTIN, R. C. and J. KRAKAUER, 1973 Distribution of gene frequency as a test of the theory of the selective neutrality of polymorphisms. *Genetics*.
- LI, J. Z., D. M. ABSHER, H. TANG, A. M. SOUTHWICK, and A. M. CASTO, 2008 Worldwide Human Relationships Inferred from Genome-Wide Patterns of Variation. *Science*.
- LIU, L., D. ZHANG, H. LIU, and C. ARENDT, 2013 Robust methods for population stratification in genome wide association studies. *BMC bioinformatics*.
- LIU, X., R. T.-H. ONG, E. N. PILLAI, A. M. ELZEIN, K. S. SMALL, T. G. CLARK, D. P. KWIATKOWSKI, and Y.-Y. TEO, 2013 Detecting and Characterizing Genomic Signatures of Positive Selection in Global Populations. *American journal of human genetics*.
- MACKAY, T. F. C., S. RICHARDS, E. A. STONE, A. BARBADILLA, J. F. AYROLES, D. ZHU, S. CASILLAS, Y. HAN, M. M. MAGWIRE, J. M. CRIDLAND, M. F. RICHARDSON, R. R. H. ANHOLT, M. BARRÓN, C. BESS, K. P. BLANKENBURG, M. A. CARBONE, D. CASTELLANO, L. CHABOUB, L. DUNCAN, Z. HARRIS, M. JAVAID, J. C. JAYASEELAN, S. N. JHANGIANI, K. W. JORDAN, F. LARA, F. LAWRENCE, S. L. LEE, P. LIBRADO, R. S. LINHEIRO, R. F. LYMAN, A. J. MACKEY, M. MUNIDASA, D. M. MUZNY, L. NAZARETH, I. NEWSHAM, L. PERALES, L.-L. PU, C. QU, M. RÂMIA, J. G. REID, S. M. ROLLMANN, J. ROZAS, N. SAADA, L. TURLAPATI, K. C. WORLEY, Y.-Q. WU, A. YAMAMOTO, Y. ZHU, C. M. BERGMAN, K. R. THORNTON, D. MITTELMAN, and R. A. GIBBS, 2012 The *Drosophila melanogaster* Genetic Reference Panel. *Nature*.
- MANOLIO, T. A., F. S. COLLINS, N. J. COX, D. B. GOLDSTEIN, L. A. HINDORFF, D. J. HUNTER, M. I. MCCARTHY, E. M. RAMOS, L. R. CARDON, A. CHAKRAVARTI, J. H. CHO, A. E. GUTTMACHER, A. KONG, L. KRUGLYAK, E. MARDIS, C. N. ROTIMI, M. SLATKIN, D. VALLE, A. S. WHITTEMORE, M. BOEHNKE, A. G. CLARK, E. E. EICHLER, G. GIBSON, J. L. HAINES, T. F. C. MACKAY, S. A. MCCARROLL, and P. M. VISSCHER, 2009 Finding the missing heritability of complex diseases. *Nature*.
- MARIGORTA, U. M. and A. NAVARRO, 2013 High Trans-ethnic Replicability of GWAS Results Implies Common Causal Variants. *PLoS Genetics*.
- MAYNARD SMITH, J. and J. HAIGH, 1974 The hitch-hiking effect of a favourable gene. *Genetics Research*.

- MCDONALD, J. H. and M. KREITMAN, 1991 Adaptive protein evolution at the Adh locus in *Drosophila*. *Nature*.
- McVICKER, G., D. GORDON, C. DAVIS, and P. GREEN, 2009 Widespread Genomic Signatures of Natural Selection in Hominid Evolution. *PLoS Genetics*.
- MEUWISSEN, T., B. HAYES, and M. GODDARD, 2013 Accelerating Improvement of Livestock with Genomic Selection. *Annual Review of Animal Biosciences*.
- MEUWISSEN, T. H., B. J. HAYES, and M. E. GODDARD, 2001 Prediction of total genetic value using genome-wide dense marker maps. *Genetics*.
- MILLER, C. T., S. BELEZA, A. A. POLLEN, D. SCHLUTER, R. A. KITTLES, M. D. SHRIVER, and D. M. KINGSLEY, 2007 cis-Regulatory Changes in Kit Ligand Expression and Parallel Evolution of Pigmentation in Sticklebacks and Humans. *Cell*.
- MORRIS, A. P., B. F. VOIGHT, T. M. TESLOVICH, T. FERREIRA, A. V. SEGRÈ, V. STEINTHORSDDOTTIR, R. J. STRAWBRIDGE, H. KHAN, H. GRALLERT, A. MAHAJAN, I. PROKOPENKO, H. M. KANG, C. DINA, T. ESKO, R. M. FRASER, S. KANONI, A. KUMAR, V. LAGOU, C. LANGENBERG, J. LUAN, C. M. LINDGREN, M. MÜLLER-NURASYID, S. PECHLIVANIS, N. W. RAYNER, L. J. SCOTT, S. WILTSHIRE, L. YENGO, L. KINNUNEN, E. J. ROSSIN, S. RAYCHAUDHURI, A. D. JOHNSON, A. S. DIMAS, R. J. F. LOOS, S. VEDANTAM, H. CHEN, J. C. FLOREZ, C. FOX, C.-T. LIU, D. RYBIN, D. J. COUPER, W. H. L. KAO, M. LI, M. C. CORNELIS, P. KRAFT, Q. SUN, R. M. VAN DAM, H. M. STRINGHAM, P. S. CHINES, K. FISCHER, P. FONTANILLAS, O. L. HOLMEN, S. E. HUNT, A. U. JACKSON, A. KONG, R. LAWRENCE, J. MEYER, J. R. B. PERRY, C. G. P. PLATOU, S. POTTER, E. REHNBERG, N. ROBERTSON, S. SIVAPALARATNAM, A. STANČÁKOVÁ, K. STIRRUPS, G. THORLEIFSSON, E. TIKKANEN, A. R. WOOD, P. ALMGREN, M. ATALAY, R. BENEDIKTSSON, L. L. BONNYCASTLE, N. BURTT, J. CAREY, G. CHARPENTIER, A. T. CRENSHAW, A. S. F. DONEY, M. DORKHAN, S. EDKINS, V. EMILSSON, E. EURY, T. FORSEN, K. GERTOW, B. GIGANTE, G. B. GRANT, C. J. GROVES, C. GUIDUCCI, C. HERDER, A. B. HREIDARSSON, J. HUI, A. JAMES, A. JONSSON, W. RATHMANN, N. KLOPP, J. KRAVIC, K. KRJUTŠKOV, C. LANGFORD, K. LEANDER, E. LINDHOLM, S. LOBBENS, S. MÄNNISTÖ, G. MIRZA, T. W. MÜHLEISEN, B. MUSK, M. PARKIN, L. RALLIDIS, J. SARAMIES, B. SENNB�AD, S. SHAH, G. SIGURDSSON, A. SILVEIRA, G. STEINBACH, B. THORAND, J. TRAKALO, F. VEGLIA, R. WENNAUER, W. WINCKLER, D. ZABANEH, H. CAMPBELL, C. VAN DUIJN, A. G. UITTERLINDEN, A. HOFMAN, E. SIJBRANDS, G. R. ABECASIS, K. R. OWEN, E. ZEGGINI, M. D. TRIP, N. G. FOROUHI, A.-C. SYVÄNEN, J. G. ERIKSSON, L. PELTONEN, M. M. NÖTHEN, B. BALKAU, C. N. A. PALMER, V. LYSSENKO, T. TUOMI, B. ISOMAA, D. J. HUNTER, L. QI, WELLCOME TRUST CASE CONTROL CONSORTIUM, META-ANALYSES OF GLUCOSE AND INSULIN-RELATED TRAITS CONSORTIUM (MAGIC) INVESTIGATORS, GENETIC INVESTIGATION OF ANTHROPOMETRIC TRAITS (GIANT) CONSORTIUM, ASIAN GENETIC EPIDEMIOLOGY NETWORK-TYPE 2 DIABETES (AGENT2D) CONSORTIUM, SOUTH ASIAN TYPE 2 DIABETES (SAT2D) CONSORTIUM, A. R. SHULDINER, M. RODEN, I. BARROSO, T. WILSGAARD, J. BEILBY, K. HOVINGH, J. F. PRICE, J. F. WILSON, R. RAURAMAA, T. A. LAKKA, L. LIND, G. DEDOUSSIS, I. NJØLSTAD, N. L. PEDERSEN, K.-T. KHAW, N. J. WAREHAM, S. M. KEINANEN-KIUKAANNIEMI, T. E. SAARISTO, E. KORPI-HYÖVÄLTI, J. SALTEVO, M. LAAKSO, J. KUUSISTO, A. METSPALU, F. S. COLLINS, K. L. MOHLKE, R. N. BERGMAN, J. TUOMILEHTO, B. O. BOEHM, C. GIEGER, K. HVEEM, S. CAUCHI, P. FROGUEL, D. BALDASSARRE, E. TREMOLI, S. E. HUMPHRIES, D. SALEHEEN, J. DANESH, E. INGELSSON, S. RIPATTI, V. SALOMAA, R. ERBEL, K.-H. JÖCKEL, S. MOEBUS, A. PETERS, T. ILLIG, U. DE FAIRE, A. HAMSTEN, A. D. MORRIS, P. J. DONNELLY, T. M. FRAYLING, A. T. HATTERSLEY, E. BOERWINKLE, O. MELANDER, S. KATHIRESAN, P. M. NILSSON, P. DELOUKAS, U. THORSTEINSDOTTIR, L. C. GROOP, K. STEFANSSON, F. HU, J. S. PANKOW, J. DUPUIS, J. B. MEIGS, D. ALTSHULER, M. BOEHNKE, M. I. MCCARTHY, and DIABETES GENETICS REPLICATION AND META-ANALYSIS (DIAGRAM) CONSORTIUM, 2012 Large-scale association analysis provides insights into the genetic architecture and pathophysiology of type 2 diabetes. *Nature Publishing Group*.
- MYLES, S., D. DAVISON, J. BARRETT, M. STONEKING, and N. TIMPSON, 2008 Worldwide population differentiation at disease-associated SNPs. *BMC Medical Genomics*.

- N'DIAYE, A., G. K. CHEN, C. D. PALMER, B. GE, and B. TAYO, 2011 PLoS Genetics: Identification, Replication, and Fine-Mapping of Loci Associated with Adult Height in Individuals of African Ancestry. PLoS Genetics.
- NEEL, J. V., 1962 Diabetes Mellitus: A "Thrifty" Genotype Rendered Detrimental by "Progress"? American journal of human genetics.
- NEI, M. and T. MARUYAMA, 1975 Lewontin-Krakauertest for neutral genes. Genetics.
- NICHOLSON, G., A. V. SMITH, F. JONSSON, O. GUSTAFSSON, K. STEFANSSON, and P. DONNELLY, 2002 Assessing population differentiation and isolation from single-nucleotide polymorphism data. Journal of the Royal Statistical Society: Series B (Statistical Methodology).
- NIELSEN, R., I. HELLMANN, M. HUBISZ, C. BUSTAMANTE, and A. G. CLARK, 2007 Recent and ongoing selection in the human genome. Nature Publishing Group.
- NIELSEN, R., S. WILLIAMSON, Y. KIM, M. J. HUBISZ, A. G. CLARK, and C. BUSTAMANTE, 2005 Genomic scans for selective sweeps using SNP data. Genome Research.
- NORTON, H. L., R. A. KITTLES, E. PARRA, P. MCKEIGUE, X. MAO, K. CHENG, V. A. CANFIELD, D. G. BRADLEY, B. MCEVOY, and M. D. SHRIVER, 2006 Genetic Evidence for the Convergent Evolution of Light Skin in Europeans and East Asians. Molecular Biology and Evolution.
- NOVEMBRE, J. and A. DI RIENZO, 2009 Spatial patterns of variation due to natural selection in humans. Nature Publishing Group.
- ORR, H. A. and J. A. COYNE, 1992 The genetics of adaptation: a reassessment. American Naturalist.
- OVASKAINEN, O., M. KARHUNEN, C. ZHENG, J. M. C. ARIAS, and J. MERILÄ, 2011 A New Method to Uncover Signatures of Divergent and Stabilizing Selection in Quantitative Traits. Genetics.
- PARRA, E. J., R. A. KITTLES, and M. D. SHRIVER, 2004 Implications of correlations between skin color and genetic ancestry for biomedical research. Nature Genetics.
- PEARCE, R. J., H. POTA, M.-S. B. EVEHE, E.-H. BÂ, G. MOMBO-NGOMA, A. L. MALISA, R. ORD, W. INOJOSA, A. MATONDO, D. A. DIALLO, W. MBACHAM, I. V. VAN DEN BROEK, T. D. SWARTHOUT, A. GETACHEW, S. DEJENE, M. P. GROBUSCH, F. NJIE, S. DUNYO, M. KWEKU, S. OWUSU-AGYEI, D. CHANDRAMOHAN, M. BONNET, J.-P. GUTHMANN, S. CLARKE, K. I. BARNES, E. STREAT, S. T. KATOKELE, P. UUSIKU, C. O. AGBOGHROMA, O. Y. ELEGBA, B. CISSÉ, I. E. A-ELBASIT, H. A. GIHA, S. P. KACHUR, C. LYNCH, J. B. RWAKIMARI, P. CHANDA, M. HAWELA, B. SHARP, I. NAIDOO, and C. ROPER, 2009 Multiple Origins and Regional Dispersal of Resistant dhps in African Plasmodium falciparum Malaria. PLoS Medicine.
- PEARSON, K., 1897 Mathematical contributions to the theory of evolution. On the law of ancestral heredity. Proceedings of the Royal Society of London.
- PEARSON, K., 1903 Mathematical Contributions to the Theory of Evolution. XII.—On a Generalised Theory of Alternative Inheritance, with Special Reference to Mendel's Laws. Proceedings of the Royal Society of London.
- PICKRELL, J. K., G. COOP, J. NOVEMBRE, S. KUDARAVALLI, J. Z. LI, D. ABSHER, B. S. SRINIVASAN, G. S. BARSH, R. M. MYERS, M. W. FELDMAN, and J. K. PRITCHARD, 2009 Signals of recent positive selection in a worldwide sample of human populations. Genome Research.
- PRICE, A. L., N. J. PATTERSON, R. M. PLENGE, M. E. WEINBLATT, N. A. SHADICK, and D. REICH, 2006 Principal components analysis corrects for stratification in genome-wide association studies. Nature Genetics.
- PRICE, A. L., N. A. ZAITLEN, D. REICH, and N. PATTERSON, 2010 New approaches to population stratification in genome-wide association studies. Nature Publishing Group.

- PRITCHARD, J. K., J. K. PICKRELL, and G. COOP, 2010 The Genetics of Human Adaptation: Hard Sweeps, Soft Sweeps, and Polygenic Adaptation. *Current Biology*.
- PROUT, T. and F. BARKER, 1993 F Statistics in *Drosophila buzzatii*: Selection, Population Size and Inbreed. *Genetics*.
- PROVINE, W. B., 2001 *The Origins of Theoretical Population Genetics*. University Of Chicago Press.
- RALPH, P. and G. COOP, 2010 Parallel adaptation: one or many waves of advance of an advantageous allele? *Genetics*.
- REICH, D. and N. PATTERSON, 2005 Will admixture mapping work to find disease genes? *Philosophical transactions of the Royal Society of London. Series B, Biological sciences*.
- RISCH, N. and K. MERIKANGAS, 1996 The future of genetic studies of complex human diseases. *Science*.
- ROBERTSON, A., 1975 GENE FREQUENCY DISTRIBUTIONS AS A TEST OF SELECTIVE NEUTRALITY. *Genetics*.
- ROCKMAN, M. V., 2012 The QTN program and the alleles that matter for evolution: all that's gold does not glitter. *Evolution; international journal of organic evolution*.
- ROGERS, A. R. and H. C. HARPENDING, 1983 Population structure and quantitative characters. *Genetics*.
- ROSENBERG, N. A., J. K. PRITCHARD, J. L. WEBER, H. M. CANN, K. K. KIDD, L. A. ZHIVOTOVSKY, and M. W. FELDMAN, 2002 Genetic structure of human populations. *Science (New York, N.Y.)*.
- SABETI, P. C., D. E. REICH, J. M. HIGGINS, H. Z. P. LEVINE, D. J. RICHTER, S. F. SCHAFFNER, S. B. GABRIEL, J. V. PLATKO, N. J. PATTERSON, G. J. McDONALD, H. C. ACKERMAN, S. J. CAMPBELL, D. ALTSHULER, R. COOPER, D. KWIATKOWSKI, R. WARD, and E. S. LANDER, 2002 Detecting recent positive selection in the human genome from haplotype structure. *Nature*.
- SABETI, P. C., S. F. SCHAFFNER, B. FRY, J. LOHMEUELLER, P. VARILLY, O. SHAMOVSKY, A. PALMA, T. S. MIKKELSEN, D. ALTSHULER, and E. S. LANDER, 2006 Positive natural selection in the human lineage. *Science (New York, N.Y.)*.
- SABETI, P. C., P. VARILLY, B. FRY, J. LOHMEUELLER, and E. HOSTETTER, 2007 Genome-wide detection and characterization of positive selection in human populations : Article : *Nature*. *Nature*.
- SCHEET, P. and M. STEPHENS, 2006 A fast and flexible statistical model for large-scale population genotype data: applications to inferring missing genotypes and haplotypic phase. *American journal of human genetics*.
- SLATE, J., 2004 INVITED REVIEW: Quantitative trait locus mapping in natural populations: progress, caveats and future directions. *Molecular Ecology*.
- SPELIOTES, E. K., C. J. WILLER, S. I. BERNDT, K. L. MONDA, G. THORLEIFSSON, A. U. JACKSON, H. L. ALLEN, C. M. LINDGREN, J. LUAN, R. M. A GI, J. C. RANDALL, S. VEDANTAM, T. W. WINKLER, L. QI, T. WORKALEMAHU, I. M. HEID, V. STEINTHORSDDOTTIR, H. M. STRINGHAM, M. N. WEEDON, E. WHEELER, A. R. WOOD, T. FERREIRA, R. J. WEYANT, A. V. SEGR E, K. ESTRADA, L. LIANG, J. NEMESH, J.-H. PARK, S. GUSTAFSSON, T. O. KILPEL A INEN, J. YANG, N. BOUATIA-NAJI, T. O. N. ESKO, M. F. FEITOSA, Z. A. N. KUTALIK, M. MANGINO, S. RAY-CHAUDHURI, A. SCHERAG, A. V. SMITH, R. WELCH, J. H. ZHAO, K. K. ABEN, D. M. ABSHER, N. AMIN, A. L. DIXON, E. FISHER, N. L. GLAZER, M. E. GODDARD, N. L. HEARD-COSTA, V. HOESEL, J.-J. HOTTENGA, Å. JOHANSSON, T. JOHNSON, S. KETKAR, C. LAMINA, S. LI, M. F. MOFFATT, R. H. MYERS, N. NARISU, J. R. B. PERRY, M. J. PETERS, M. PREUSS, S. RIPATTI, F. RIVADENEIRA, C. SANDHOLT, L. J. SCOTT, N. J. TIMPSON, J. P. TYRER, S. VAN WINGERDEN, R. M. WATANABE, C. C. WHITE, F. WIKLUND, C. BARLASSINA, D. I. CHASMAN, M. N. COOPER, J.-O. JANSSON, R. W. LAWRENCE, N. PELLIKKA, I. PROKOPENKO, J. SHI, E. THIERING, H. ALAVERE, M. T. S. ALIBRANDI, P. ALMGREN, A. M. ARNOLD, T. ASPELUND, L. D. ATWOOD,

- B. BALKAU, A. J. BALMFORTH, A. J. BENNETT, Y. BEN-SHLOMO, R. N. BERGMAN, S. BERGMANN, H. BIEBERMANN, A. I. F. BLAKEMORE, T. BOES, L. L. BONNYCASTLE, S. R. BORNSTEIN, M. J. BROWN, T. A. BUCHANAN, F. BUSONERO, H. CAMPBELL, F. P. CAPPUCCIO, C. CAVALCANTI-PROENÇA, Y.-D. I. CHEN, C.-M. CHEN, P. S. CHINES, R. CLARKE, L. COIN, J. CONNELL, I. N. M. DAY, M. D. HEIJER, J. DUAN, S. EBRAHIM, P. ELLIOTT, R. ELOSUA, G. EIRIKSDOTTIR, M. R. ERDOS, J. G. ERIKSSON, M. F. FACHERIS, S. B. FELIX, P. FISCHER-POSOVSZKY, A. R. FOLSOM, N. FRIEDRICH, N. B. FREIMER, M. FU, S. GAGET, P. V. GEJMAN, E. J. C. GEUS, C. GIEGER, A. P. GJESING, A. GOEL, P. GOYETTE, H. GRALLERT, J. U. R. GRASSLER, D. M. GREENAWALT, C. J. GROVES, V. GUDNASON, C. GUIDUCCI, A.-L. HARTIKAINEN, N. HASSANALI, A. S. HALL, A. S. HAVULINNA, C. HAYWARD, A. C. HEATH, C. HENGSTENBERG, A. A. HICKS, A. HINNEY, A. HOFMAN, G. HOMUTH, J. HUI, W. IGL, C. IRIBARREN, B. ISOMAA, K. B. JACOBS, I. JARICK, E. JEWELL, U. JOHN, T. JORGENSEN, P. JOUSILAHTI, A. JULA, M. KAAKINEN, E. KAJANTIE, L. M. KAPLAN, S. KATHIRESAN, J. KETTUNEN, L. KINNUNEN, J. W. KNOWLES, I. KOLCIC, I. R. KONIG, S. KOSKINEN, P. KOVACS, J. KUUSISTO, P. KRAFT, K. KVALØY, J. LAITINEN, O. LANTIERI, C. LANZANI, L. J. LAUNER, C. LECOEUR, T. LEHTIMÄKI, G. LETTRE, J. LIU, M.-L. LOKKI, M. LORENTZON, R. N. LUBEN, B. LUDWIG, P. MANUNTA, D. MAREK, M. MARRE, N. G. MARTIN, W. L. MCARDLE, A. MCCARTHY, B. MCKNIGHT, T. MEITINGER, O. MELANDER, D. MEYRE, K. MIDTHJELL, G. W. MONTGOMERY, M. A. MORKEN, A. P. MORRIS, R. MULIC, J. S. NGWA, M. NELIS, M. J. NEVILLE, D. R. NYHOLT, C. J. O'DONNELL, S. O'RAHILLY, K. K. ONG, B. OOSTRA, G. PARÈ, A. N. PARKER, M. PEROLA, I. PICHLER, K. H. PIETILÄINEN, C. G. P. PLATOU, O. POLASEK, A. POUTA, S. RAFELT, O. RAITAKARI, N. W. RAYNER, M. RIDDERSTRAALE, W. RIEF, A. RUOKONEN, N. R. ROBERTSON, P. RZEHA, V. SALOMAA, and S..., 2010 Association analyses of 249,796 individuals reveal 18 new loci associated with body mass index. *Nature Genetics*.
- SPENCER, C., E. HECHTER, D. VUKCEVIC, and P. DONNELLY, 2011 Quantifying the Underestimation of Relative Risks from Genome-Wide Association Studies. *PLoS Genetics*.
- SPITZE, K., 1993 Population structure in *Daphnia obtusa*: quantitative genetic and allozymic variation. *Genetics*.
- STURM, R. A., 2009 Molecular genetics of human pigmentation diversity. *Human Molecular Genetics*.
- TURCHIN, M. C., C. W. CHIANG, C. D. PALMER, S. SANKARARAMAN, D. REICH, and J. N. HIRSCHHORN, 2012 Evidence of widespread selection on standing variation in Europe at height-associated SNPs. *Nature Genetics*.
- TURELLI, M. and N. H. BARTON, 1990 Dynamics of polygenic characters under selection. *Theoretical Population Biology*.
- VATN, M. H., 2012 Environmental Factors in the Epidemiology of Inflammatory Bowel Disease. In *Crohn's Disease and Ulcerative Colitis: From Epidemiology and Immunobiology to a Rational Diagnostic and Therapeutic Approach*.
- VISSCHER, P. M., M. A. BROWN, M. I. MCCARTHY, and J. YANG, 2012 Five Years of GWAS Discovery. *The American Journal of Human Genetics*.
- VOIGHT, B. F., S. KUDARAVALLI, X. WEN, and J. K. PRITCHARD, 2006 A Map of Recent Positive Selection in the Human Genome. *PLoS Biology*.
- VOIGHT, B. F., L. J. SCOTT, V. STEINTHORSDDOTTIR, A. P. MORRIS, C. DINA, R. P. WELCH, E. ZEGGINI, C. HUTH, Y. S. AULCHENKO, G. THORLEIFSSON, L. J. MCCULLOCH, T. FERREIRA, H. GRALLERT, N. AMIN, G. WU, C. J. WILLER, S. RAYCHAUDHURI, S. A. MCCARROLL, C. LANGENBERG, O. M. HOFMANN, J. DUPUIS, L. QI, A. V. SEGRÈ, M. VAN HOEK, P. NAVARRO, K. ARDLIE, B. BALKAU, R. BENEDIKTSSON, A. J. BENNETT, R. BLAGIEVA, E. BOERWINKLE, L. L. BONNYCASTLE, K. B. BOSTRÖM, B. BRAVENBOER, S. BUMPSTEAD, N. P. BURTT, G. CHARPENTIER, P. S. CHINES, M. CORNELIS, D. J. COUPER, G. CRAWFORD, A. S. F. DONEY, K. S. ELLIOTT, A. L. ELLIOTT, M. R. ERDOS, C. S. FOX, C. S. FRANKLIN, M. GANSER, C. GIEGER, N. GRARUP, T. GREEN, S. GRIFFIN, C. J. GROVES, C. GUIDUCCI, S. HADJADJ, N. HASSANALI,

- C. HERDER, B. ISOMAA, A. U. JACKSON, P. R. V. JOHNSON, T. JØRGENSEN, W. H. L. KAO, N. KLOPP, A. KONG, P. KRAFT, J. KUUSISTO, T. LAURITZEN, M. LI, A. LIEVERSE, C. M. LINDGREN, V. LYSSENKO, M. MARRE, T. MEITINGER, K. MIDTHJELL, M. A. MORKEN, N. NARISU, P. NILSSON, K. R. OWEN, F. PAYNE, J. R. B. PERRY, A.-K. PETERSEN, C. PLATOU, C. PROENÇA, I. PROKOPENKO, W. RATHMANN, N. W. RAYNER, N. R. ROBERTSON, G. ROCHELEAU, M. RODEN, M. J. SAMPSON, R. SAXENA, B. M. SHIELDS, P. SHRADER, G. SIGURDSSON, T. SPARSØ, K. STRASSBURGER, H. M. STRINGHAM, Q. SUN, A. J. SWIFT, B. THORAND, J. TICHET, T. TUOMI, R. M. VAN DAM, T. W. VAN HAEFTEN, T. VAN HERPT, J. V. VAN VLIET-OSTAPTCHOUK, G. B. WALTERS, M. N. WEEDON, C. WIJMENG, J. WITTEMAN, R. N. BERGMAN, S. CAUCHI, F. S. COLLINS, A. L. GLOYN, U. GYLLENSTEN, T. HANSEN, W. A. HIDE, G. A. HITMAN, A. HOFMAN, D. J. HUNTER, K. HVEEM, M. LAAKSO, K. L. MOHLKE, A. D. MORRIS, C. N. A. PALMER, P. P. PRAMSTALLER, I. RUDAN, E. SIJBRANDS, L. D. STEIN, J. TUOMILEHTO, A. UITTERLINDEN, M. WALKER, N. J. WAREHAM, R. M. WATANABE, G. R. ABECASIS, B. O. BOEHM, H. CAMPBELL, M. J. DALY, A. T. HATTERSLEY, F. B. HU, J. B. MEIGS, J. S. PANKOW, O. PEDERSEN, H.-E. WICHMANN, I. BARROSO, J. C. FLOREZ, T. M. FRAYLING, L. GROOP, R. SLADEK, U. THORSTEINSDOTTIR, J. F. WILSON, T. ILLIG, P. FROGUEL, C. M. VAN DUIN, K. STEFANSSON, D. ALTSHULER, M. BOEHNE, and M. I. MCCARTHY, 2010 Twelve type 2 diabetes susceptibility loci identified through large-scale association analysis. Nature Publishing Group.
- WANG, E. T., G. KODAMA, P. BALDI, and R. K. MOYZIS, 2006 Global landscape of recent inferred Darwinian selection for *Homo sapiens*. Proceedings of the National Academy of Sciences.
- WEINBERG, W., 1909 Über Vererbungsgesetze beim Menschen. Molecular and General Genetics MGG.
- WEIR, B. S. and W. G. HILL, 2002 Estimating F-statistics. Annual review of genetics.
- WELDON, W. F. R., 1902 Mendel's laws of alternative inheritance in peas. Biometrika.
- WELDON, W. F. R., 1903 Mr Bateson's revisions of Mendel's theory of heredity. Biometrika.
- WHITLOCK, M. C., 1999 Neutral additive genetic variance in a metapopulation. Genetical Research.
- WHITLOCK, M. C., 2008 Evolutionary inference from QST. Molecular Ecology.
- WHITLOCK, M. C. and F. GUILLAUME, 2009 Testing for Spatially Divergent Selection: Comparing QST to FST. Genetics.
- WILLIAMSON, S. H., M. J. HUBISZ, A. G. CLARK, B. A. PAYSEUR, C. D. BUSTAMANTE, and R. NIELSEN, 2007 Localizing Recent Adaptive Evolution in the Human Genome. PLoS Genetics.
- WRIGHT, S., 1951 The Genetical Structure of Populations. Annals of Eugenics.
- YANG, J., B. BENYAMIN, B. P. McEVOY, S. GORDON, A. K. HENDERS, D. R. NYHOLT, P. A. MADDEN, A. C. HEATH, N. G. MARTIN, G. W. MONTGOMERY, M. E. GODDARD, and P. M. VISSCHER, 2010 Common SNPs explain a large proportion of the heritability for human height. Nature Genetics.
- YANG, J., T. A. MANOLIO, L. R. PASQUALE, E. BOERWINKLE, N. CAPORASO, J. M. CUNNINGHAM, M. DE ANDRADE, B. FEENSTRA, E. FEINGOLD, M. G. HAYES, W. G. HILL, M. T. LANDI, A. ALONSO, G. LETTRE, P. LIN, H. LING, W. LOWE, R. A. MATHIAS, M. MELBYE, E. PUGH, M. C. CORNELIS, B. S. WEIR, M. E. GODDARD, and P. M. VISSCHER, 2011 ng.823. Nature Publishing Group.
- YULE, G. U., 1902 Mendel's Laws and Their Probable Relations to Intra-Racial Heredity. The New phytologist.
- ZAITLEN, N., P. KRAFT, N. PATTERSON, B. PASANIUC, G. BHATIA, S. POLLACK, and A. L. PRICE, 2013 Using Extended Genealogy to Estimate Components of Heritability for 23 Quantitative and Dichotomous Traits. PLoS Genetics.

- ZHANG, X. S. and W. G. HILL, 2008 The Anomalous Effects of Biased Mutation Revisited: Mean-Optimum Deviation and Apparent Directional Selection Under Stabilizing Selection. *Genetics*.
- ZHOU, X., P. CARBONETTO, and M. STEPHENS, 2013 Polygenic modeling with bayesian sparse linear mixed models. *PLoS Genetics*.
- ZHOU, X. and M. STEPHENS, 2012 technical reports. *Nature Genetics*.
- ZOLLNER, S. and J. K. PRITCHARD, 2007 Overcoming the Winner's Curse: Estimating Penetrance Parameters from Case-Control Data. *The American Journal of Human Genetics*.

6 Supplementary materials

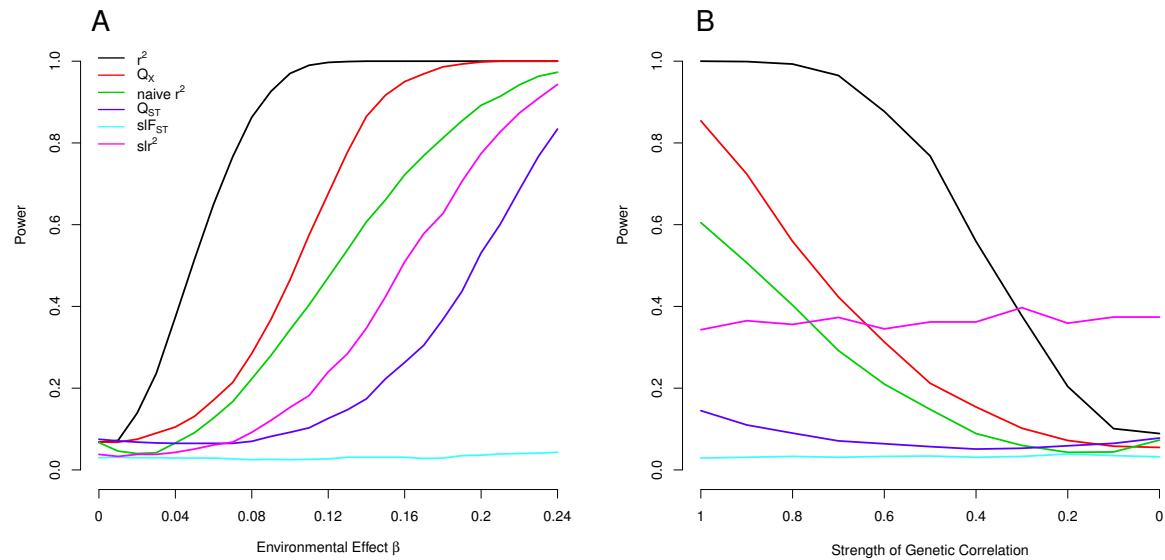


Figure S1: Power of tests described in the main text to detect a signal of selection on the mapped genetic basis of height (LANGO ALLEN *et al.*, 2010) as an increasing function of the strength of selection (A), and a decreasing function of the genetic correlation between height and the selected trait with the effect of selection held constant at $\beta = 0.14$ (B).

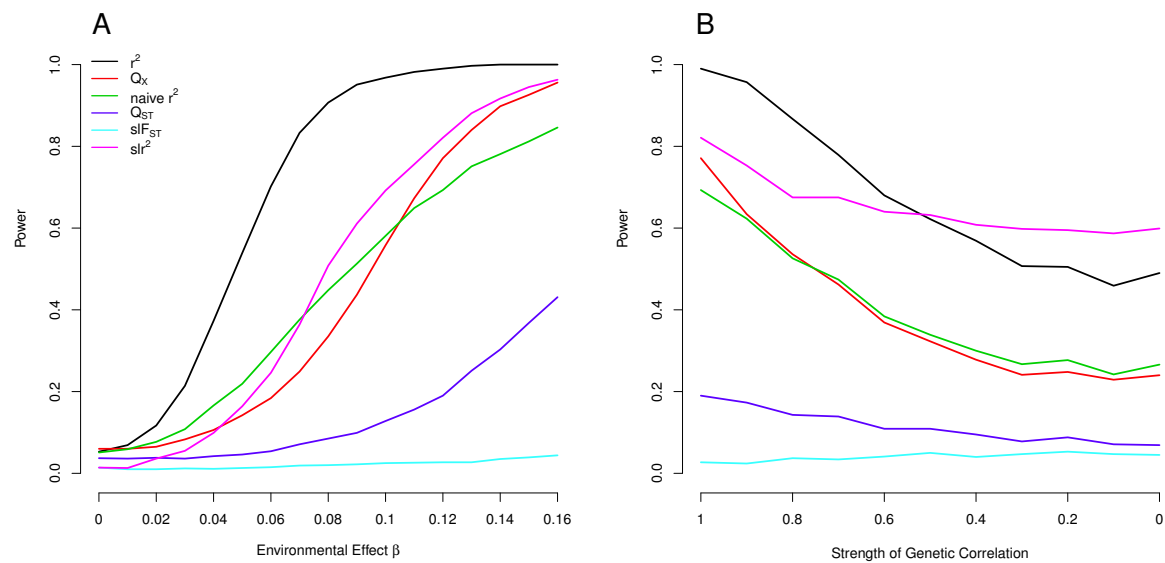


Figure S2: Power of tests described in the main text to detect a signal of selection on the mapped genetic basis of skin pigmentation (BELEZA *et al.*, 2013) as an increasing function of the strength of selection (A), and a decreasing function of the genetic correlation between skin pigmentation and the selected trait with the effect of selection held constant at $\beta = 0.12$ (B).

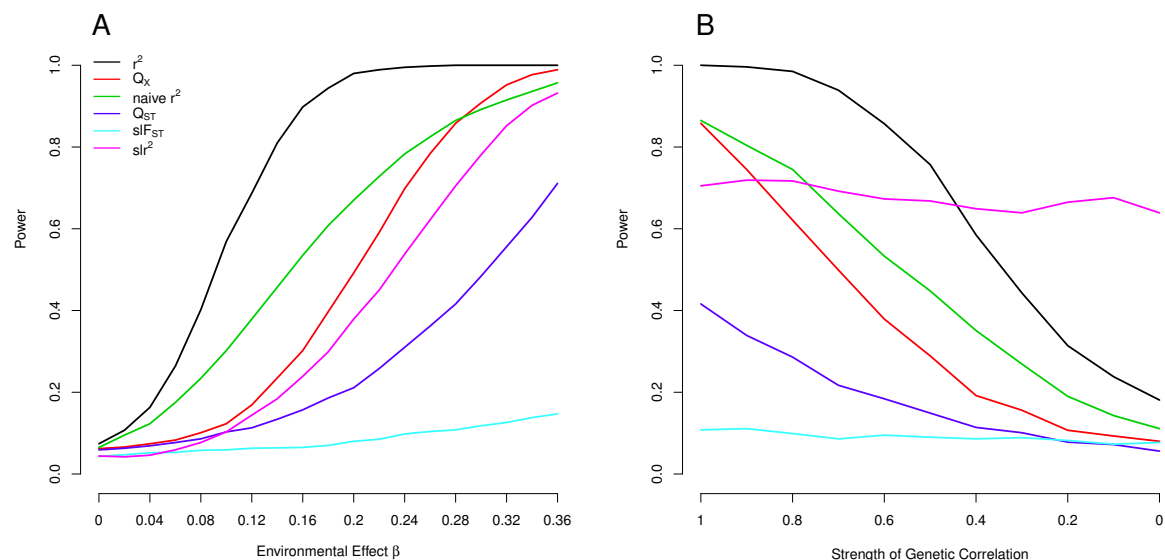


Figure S3: Power of tests described in the main text to detect a signal of selection on the mapped genetic basis of BMI (SPELIOTES *et al.*, 2010) as an increasing function of the strength of selection (A), and a decreasing function of the genetic correlation between BMI and the selected trait with the effect of selection held constant at $\beta = 0.28$ (B).

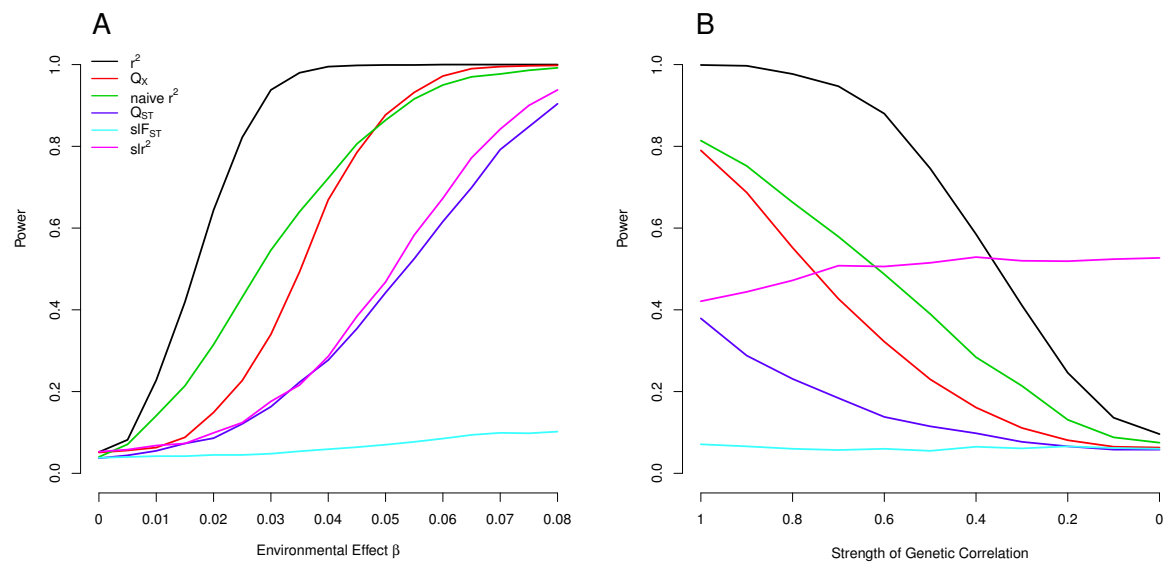


Figure S4: Power of tests described in the main text to detect a signal of selection on the mapped genetic basis of CD (JOSTINS *et al.*, 2012) as an increasing function of the strength of selection (A), and a decreasing function of the genetic correlation between CD and the selected trait with the effect of selection held constant at $\beta = 0.045$ (B).

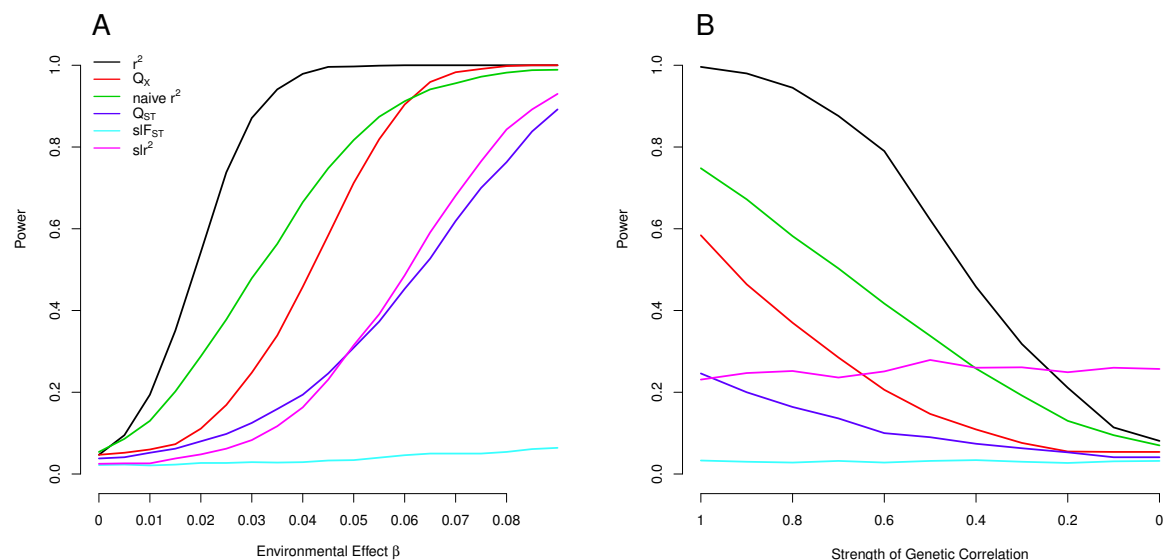


Figure S5: Power of tests described in the main text to detect a signal of selection on the mapped genetic basis of UC (JOSTINS *et al.*, 2012) as an increasing function of the strength of selection (A), and a decreasing function of the genetic correlation between UC and the selected trait with the effect of selection held constant at $\beta = 0.045$ (B).

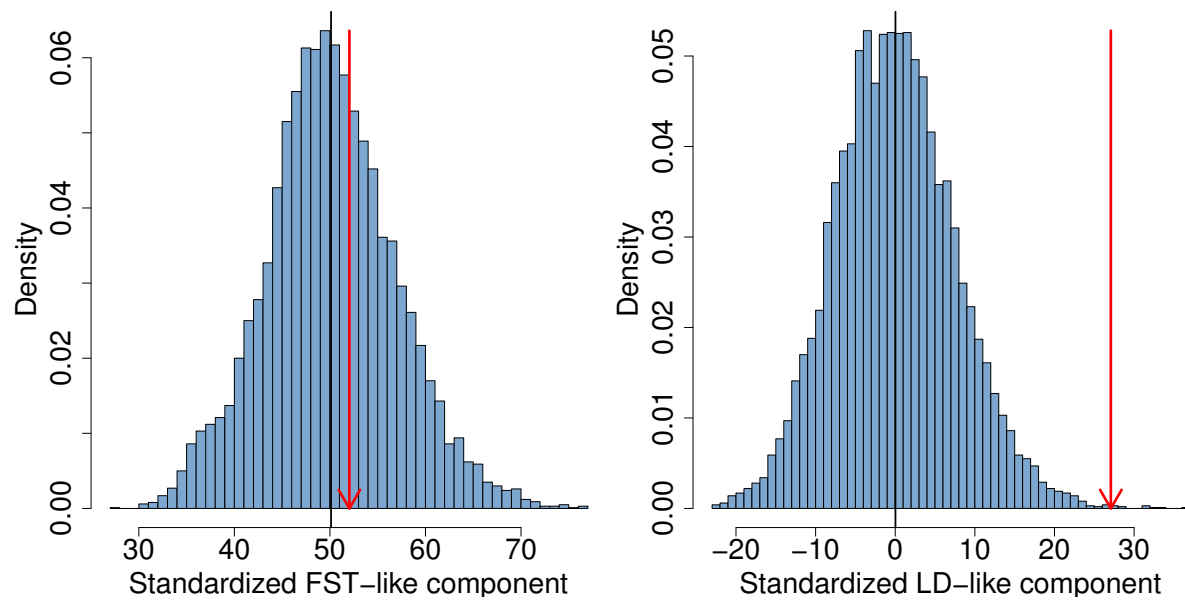


Figure S6: The two components of Q_X for the skin pigmentation dataset, as described by the left and right terms in (14). The null distribution of each component is shown as a histogram. The expected value is shown as a black bar, and the observed value as a red arrow.

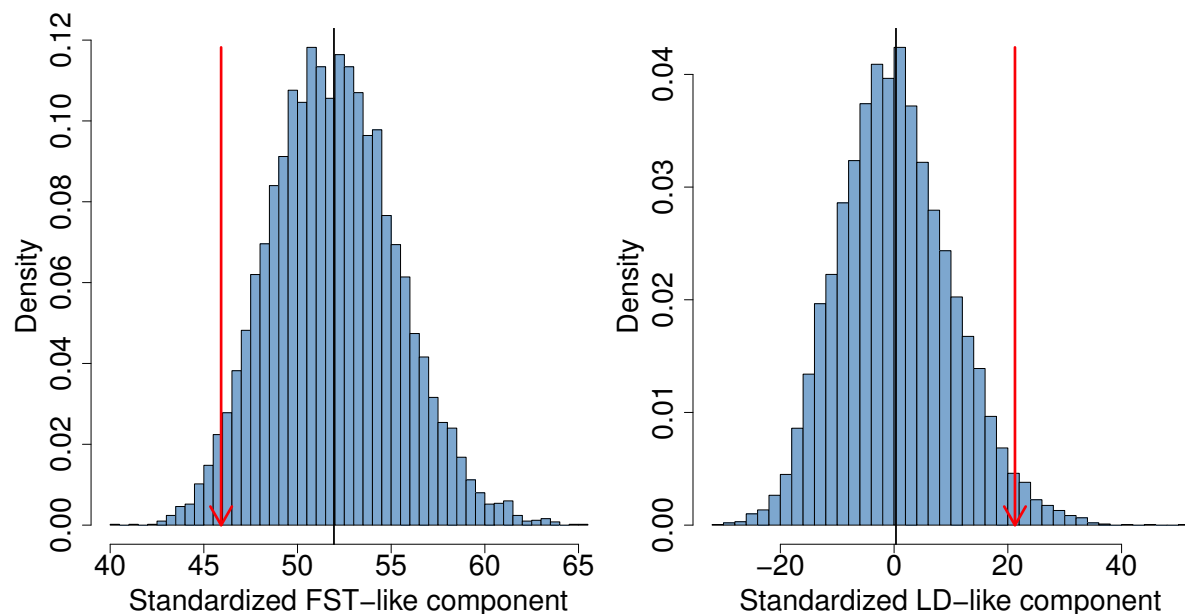


Figure S7: The two components of Q_X for the BMI dataset, as described by the left and right terms in (14). The null distribution of each component is shown as a histogram. The expected value is shown as a black bar, and the observed value as a red arrow.

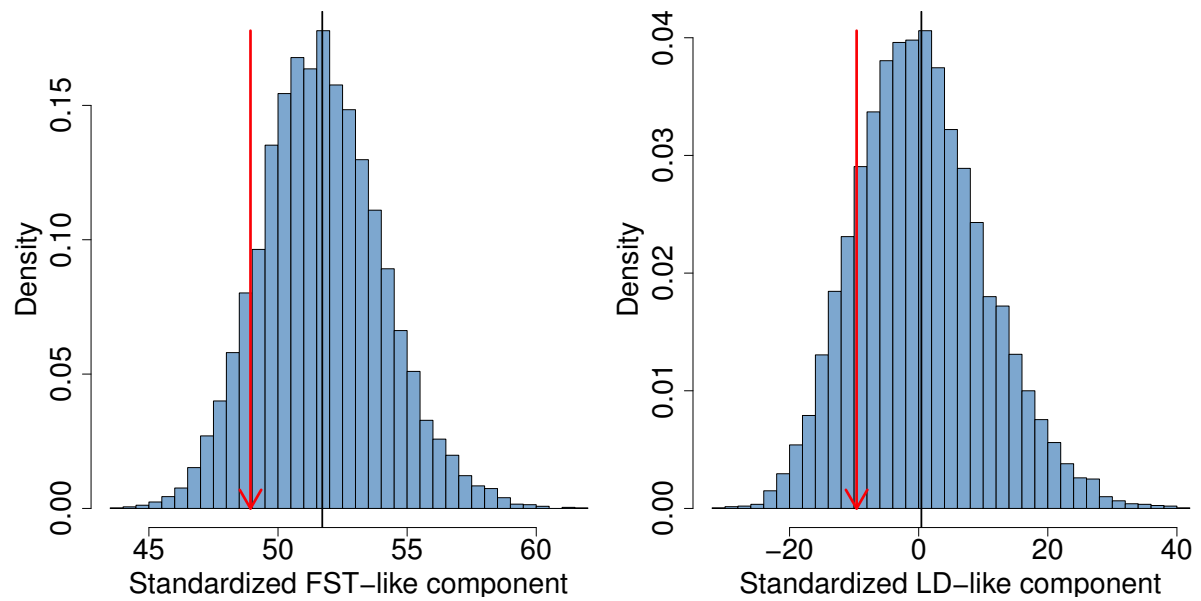


Figure S8: The two components of Q_X for the T2D dataset, as described by the left and right terms in (14). The null distribution of each component is shown as a histogram. The expected value is shown as a black bar, and the observed value as a red arrow.

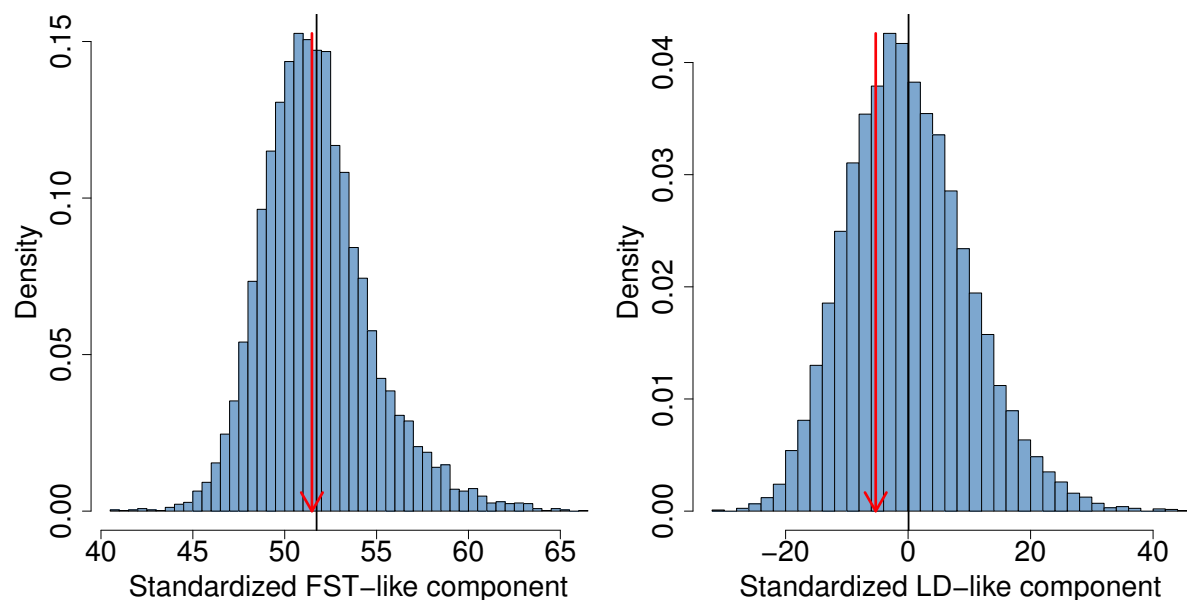


Figure S9: The two components of Q_X for the CD dataset, as described by the left and right terms in (14). The null distribution of each component is shown as a histogram. The expected value is shown as a black bar, and the observed value as a red arrow.

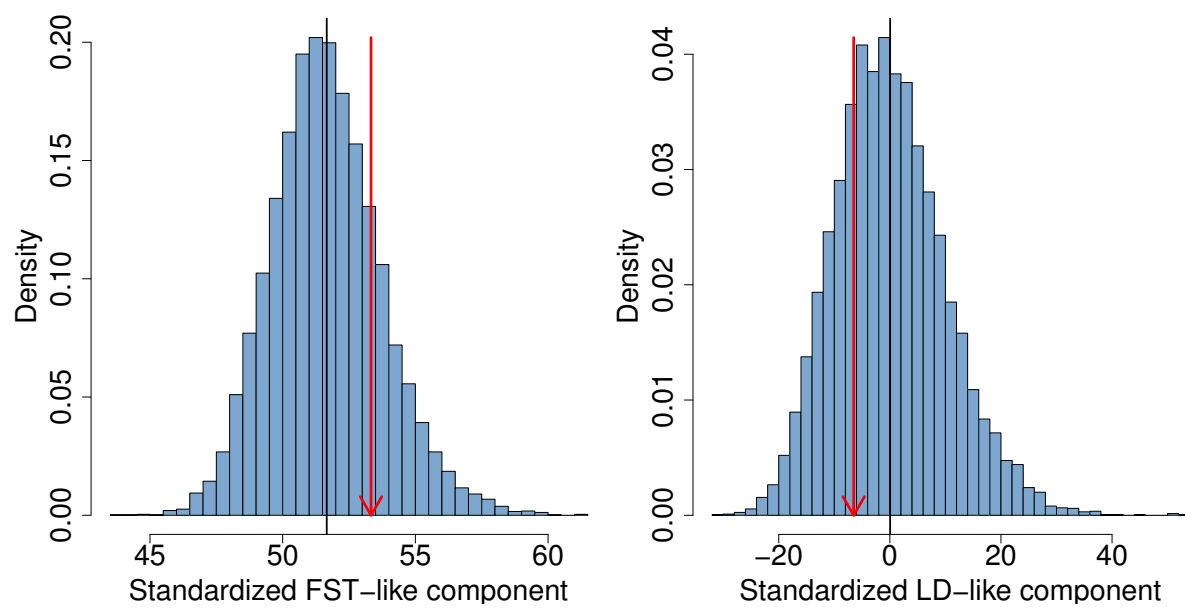


Figure S10: The two components of Q_X for the UC dataset, as described by the left and right terms in (14). The null distribution of each component is shown as a histogram. The expected value is shown as a black bar, and the observed value as a red arrow.

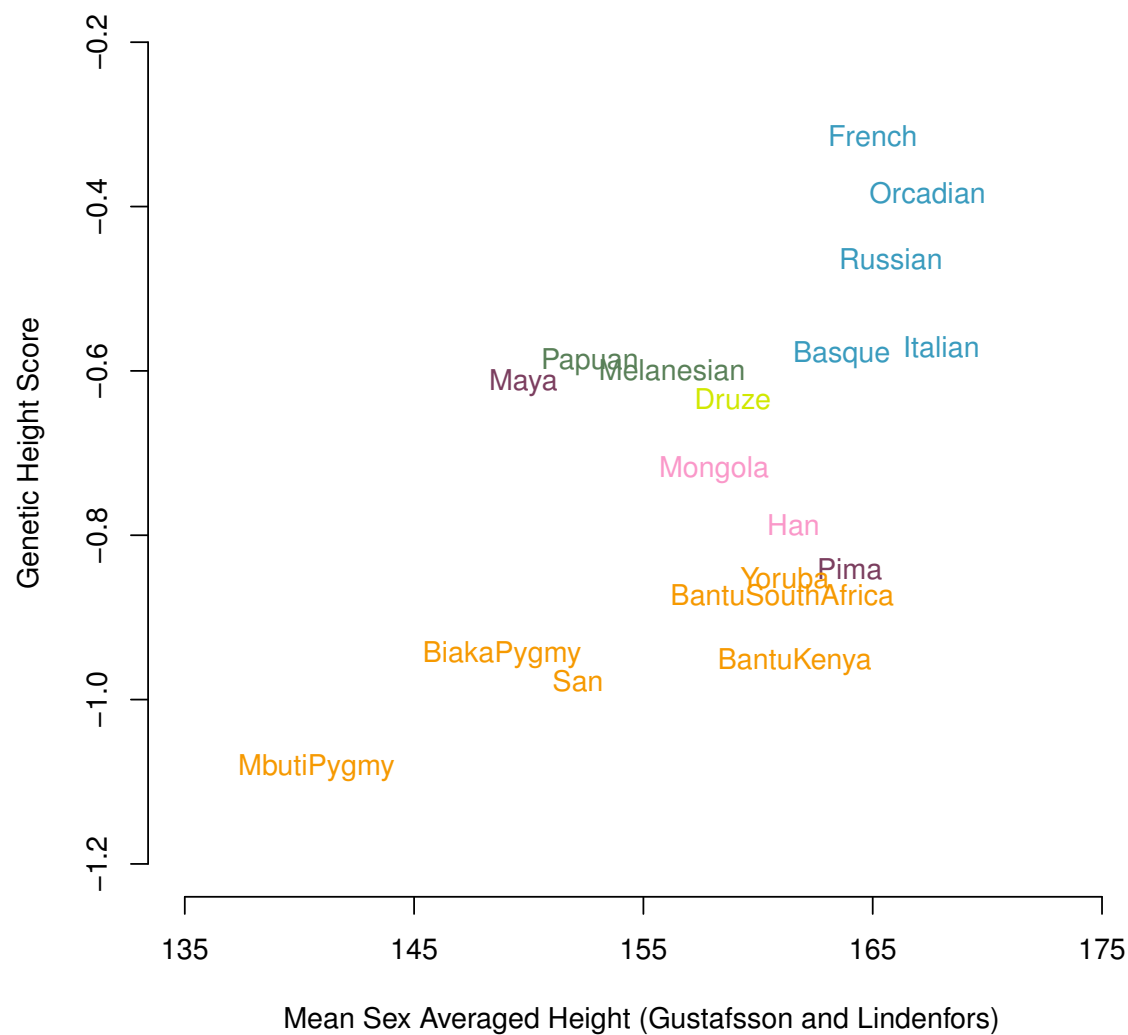
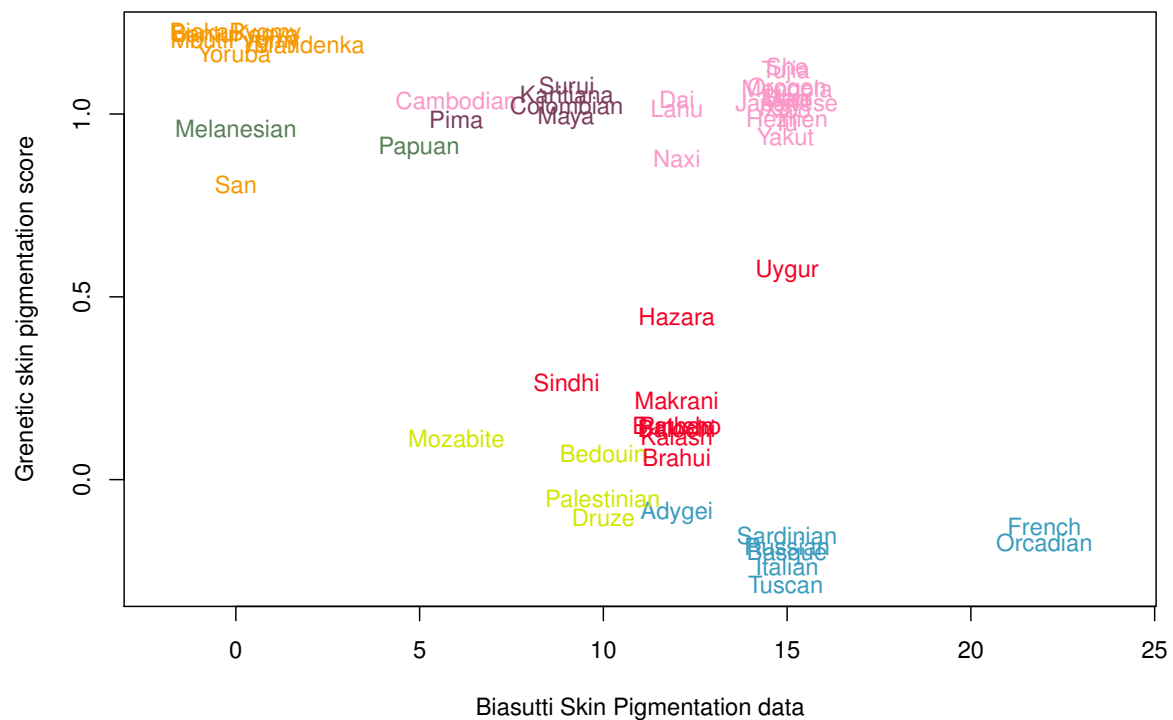


Figure S11: The genetic values for height in each HGDP population plotted against the measured sex averaged height taken from GUSTAFSSON and LINDENFORS (2004). Only the subset of populations with an appropriately close match in the named population in GUSTAFSSON and LINDENFORS's (2004) Appendix I are shown, values used are given in Supplementary table ??



Predicted skin pigmentation with OCA2 and KITLG

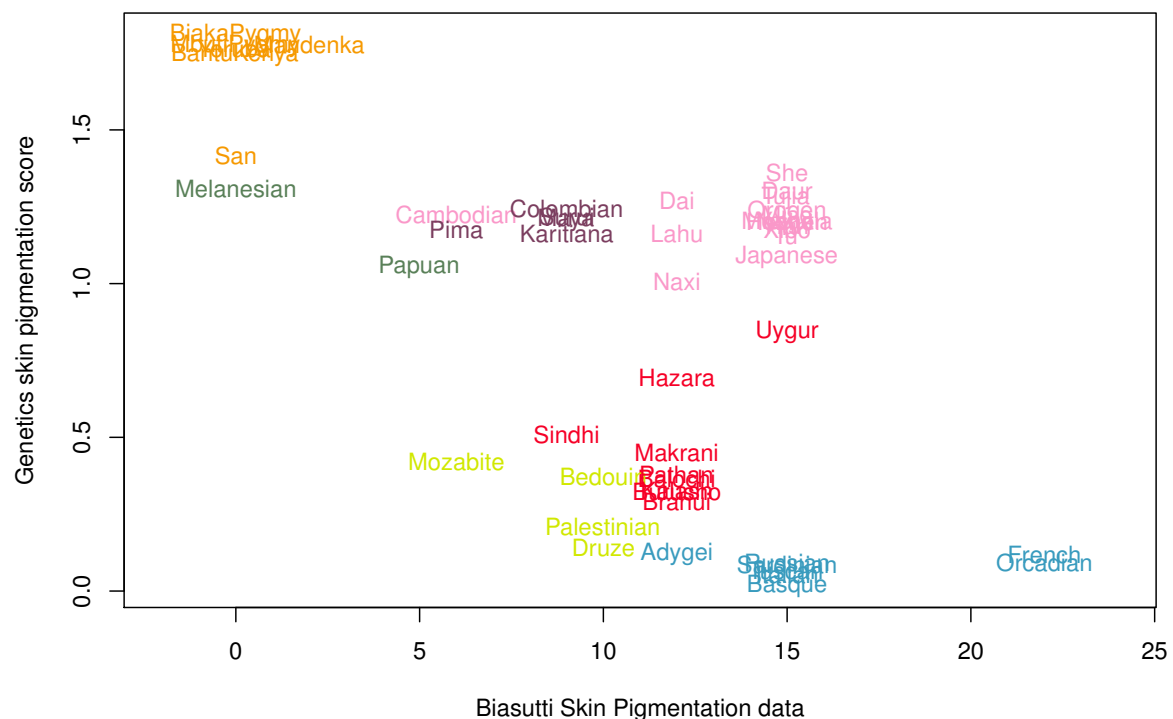
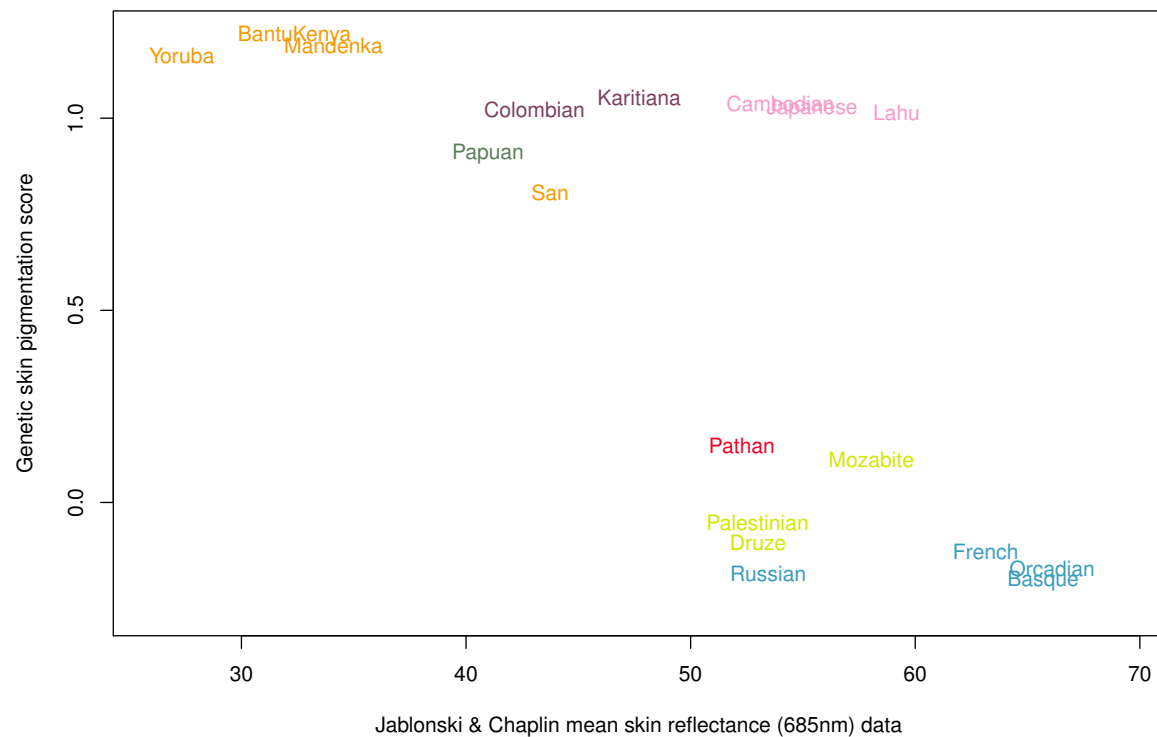


Figure S12: The genetic skin pigmentation score for a each HGDP population plotted against the HGDP populations values on the skin pigmentation index map of Biasutti 1959. Data obtained from Supplementary table of LAO *et al.* (2007). Note that Biasutti map is interpolated, and so values are known to be imperfect. Values used are given in Supplementary table S2



Predicted skin pigmentation with OCA2 and KITLG

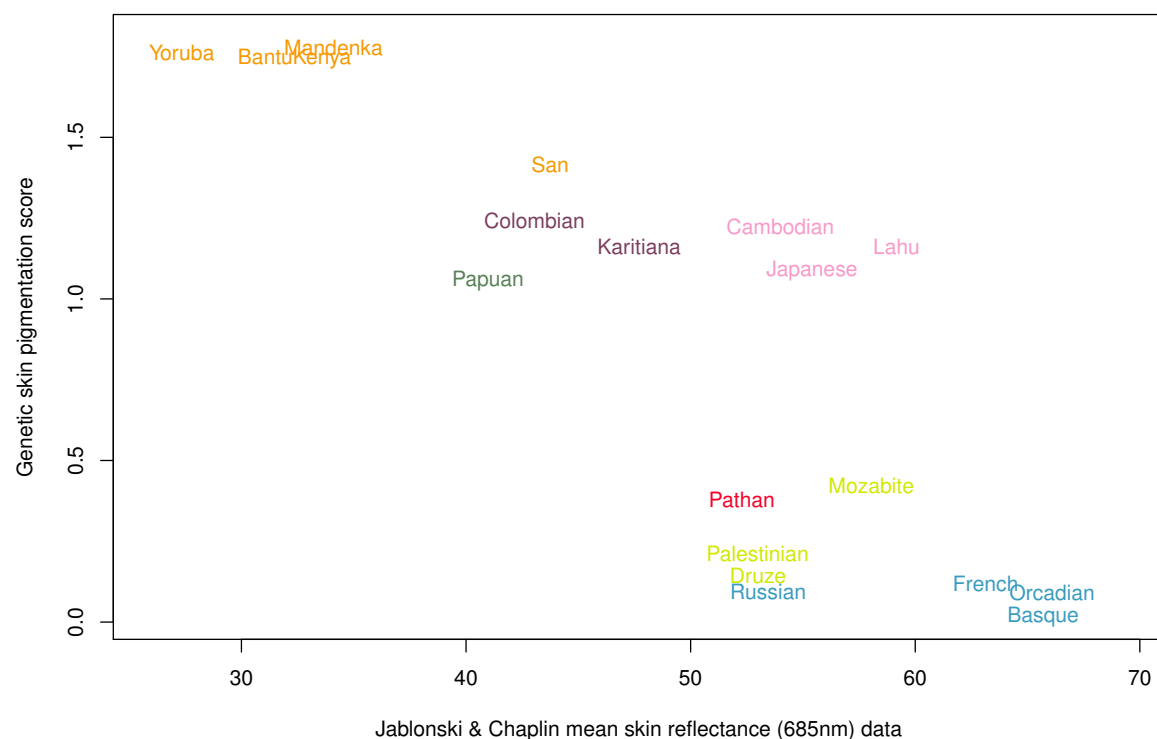


Figure S13: The genetic skin pigmentation score for a each HGDP population plotted against the HGDP populations values from the JABLONSKI and CHAPLIN (2000) mean skin reflectance (685nm) data (their Table 6). Only the subset of populations with an appropriately close match were used as in the Supplementary table of LAO *et al.* (2007). Values and populations used are given in Table S2

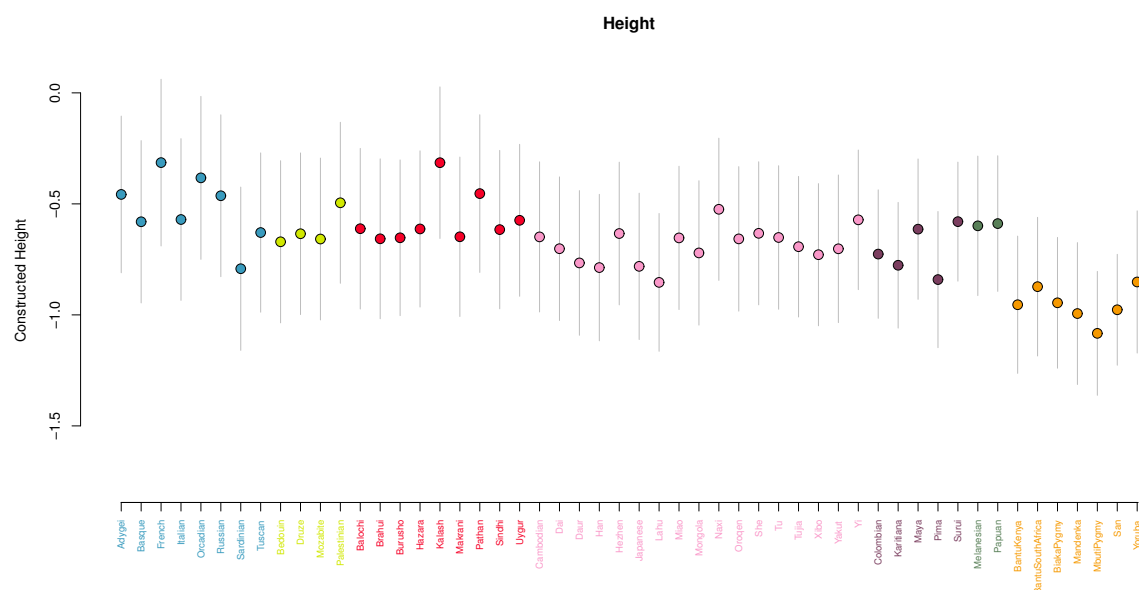


Figure S14: The distribution of genetic height score across all 52 HGDP populations. Grey bars represent the 95% confidence interval for the genetic height score of an individual randomly chosen from that population under Hardy-Weinberg assumptions

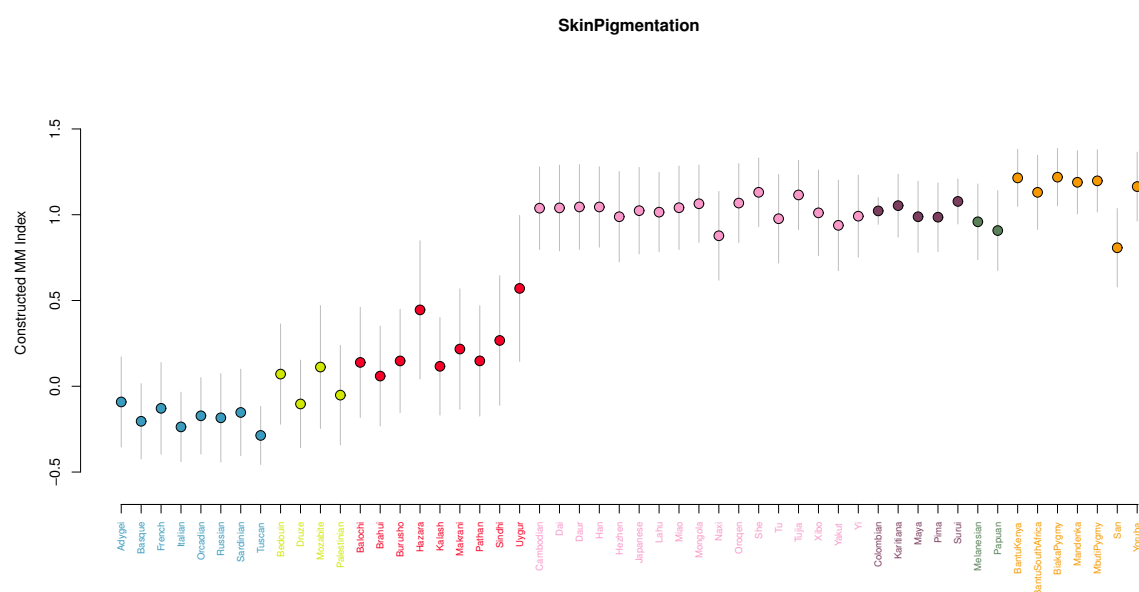


Figure S15: The distribution of genetic skin pigmentation score across all 52 HGDP populations. Grey bars represent the 95% confidence interval for the genetic skin pigmentation score of an individual randomly chosen from that population under Hardy-Weinberg assumptions

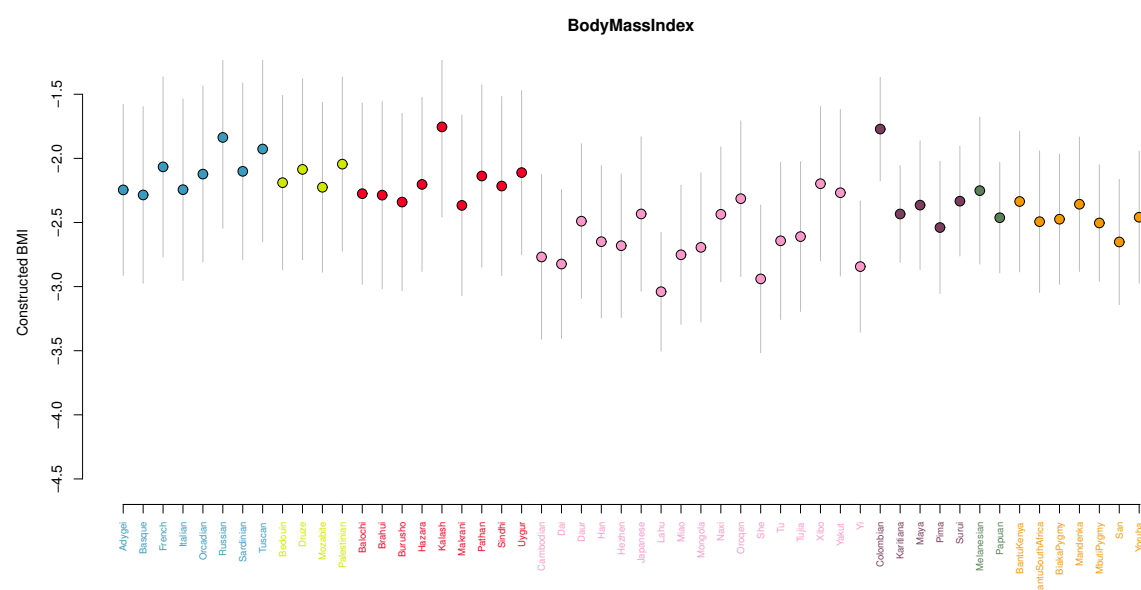
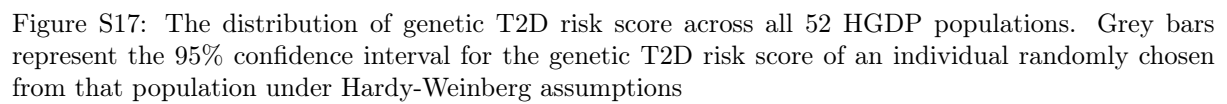


Figure S16: The distribution of genetic BMI score across all 52 HGDP populations. Grey bars represent the 95% confidence interval for the genetic BMI score of an individual randomly chosen from that population under Hardy-Weinberg assumptions



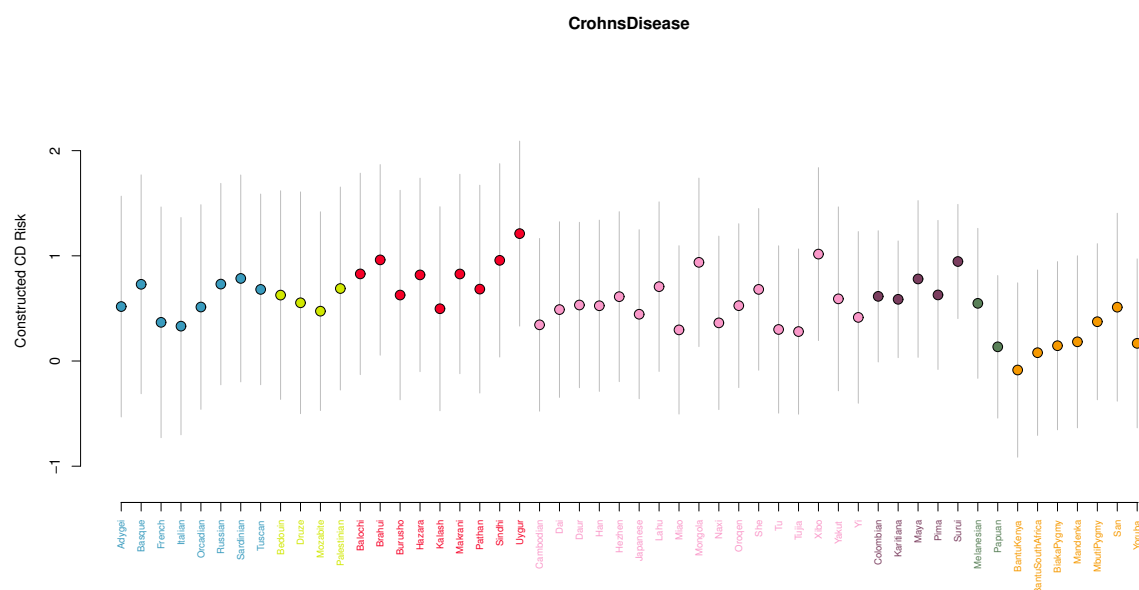


Figure S18: The distribution of genetic CD risk score across all 52 HGDP populations. Grey bars represent the 95% confidence interval for the genetic CD risk score of an individual randomly chosen from that population under Hardy-Weinberg assumptions

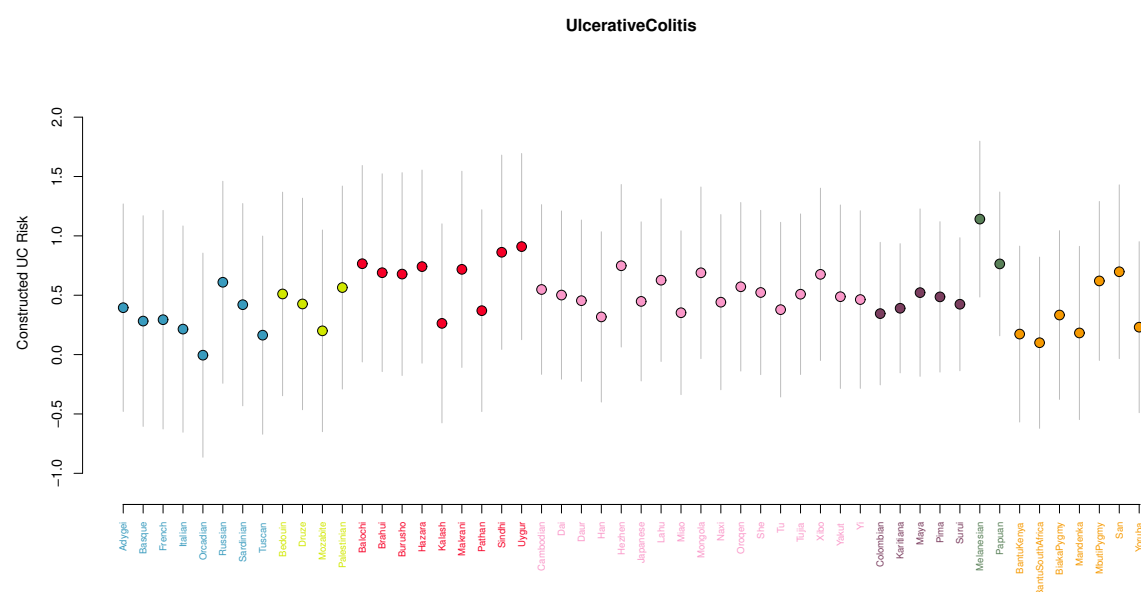


Figure S19: The distribution of genetic UC risk score across all 52 HGP populations. Grey bars represent the 95% confidence interval for the genetic UC risk score of an individual randomly chosen from that population under Hardy-Weinberg assumptions

Population	Genetic Height Score	Male Height (cm)	Female Height (cm)	Sex Averaged Height (cm)
BantuKenya	-0.95	167.30	156.00	161.65
BantuSouthAfrica	-0.87	166.10	156.00	161.05
Basque	-0.58	170.00	157.30	163.65
BiakaPygmy	-0.95	152.70	145.00	148.85
Druze	-0.63	165.60	152.20	158.90
French	-0.31	169.60	160.40	165.00
Han	-0.79	167.10	156.00	161.55
Italian	-0.57	174.00	162.00	168.00
Maya	-0.61	156.80	142.80	149.80
MbutiPygmy	-1.08	144.19	137.35	140.77
Melanesian	-0.60	162.10	150.40	156.25
Mongola	-0.72	164.83	151.33	158.08
Orcadian	-0.38	173.90	160.90	167.40
Papuan	-0.59	155.97	149.41	152.69
Pima	-0.84	171.00	157.00	164.00
Russian	-0.46	171.80	159.80	165.80
San	-0.98	157.70	146.60	152.15
Yoruba	-0.85	167.50	155.00	161.25

Table S1: Genetic height scores as compared to true heights for populations with a suitably close match in the dataset of GUSTAFSSON and LINDENFORS (2004). See Figure S11 for a plot of genetic height score against sex averaged height.

Population	Genetic Skin Pigmentation Score	Genetic Skin Pigment Score w/ KITLG+OCA2	Biasutti	JablonskiChaplin
Adygei	-0.09	0.12	12	
Balochi	0.14	0.36	12	
BantuKenya	1.21	1.74	0	32.40
Basque	-0.20	0.02	15	65.70
Bedouin	0.07	0.37	10	
BiakaPygmy	1.22	1.81	0	
Brahui	0.06	0.29	12	
Burusho	0.15	0.32	12	
Cambodian	1.04	1.22	6	54.00
Colombian	1.02	1.24	9	43.05
Dai	1.04	1.27	12	
Daur	1.05	1.30	15	
Druze	-0.10	0.14	10	53.00
French	-0.13	0.12	22	63.14
Han	1.05	1.19	15	
Hazara	0.45	0.69	12	
Hezhen	0.99	1.20	15	
Italian	-0.24	0.05	15	
Japanese	1.02	1.09	15	55.42
Kalash	0.12	0.33	12	
Karitiana	1.05	1.16	9	47.70
Lahu	1.01	1.16	12	59.17
Makrani	0.22	0.45	12	
Mandenka	1.19	1.78	2	34.10
Maya	0.99	1.21	9	
MbutiPygmy	1.20	1.77	0	
Melanesian	0.96	1.31	0	
Miao	1.04	1.22	15	
Mongola	1.06	1.20	15	
Mozabite	0.11	0.42	6	58.05
Naxi	0.88	1.01	12	
Orcadian	-0.17	0.09	22	66.10
Oroqen	1.07	1.23	15	
Palestinian	-0.05	0.21	10	53.00
Papuan	0.91	1.06	5	41.00
Pathan	0.15	0.38	12	52.30
Pima	0.99	1.18	6	
Russian	-0.18	0.09	15	53.45
San	0.81	1.42	0	43.75
Sardinian	-0.15	0.09	15	
She	1.13	1.36	15	
Sindhi	0.27	0.51	9	
Surui	1.08	1.22	9	
Tu	0.98	1.15	15	
Tujia	1.12	1.28	15	
Tuscan	-0.29	0.07	15	
Uygur	0.57	0.84	15	
Xibo	1.01	1.17	15	
Yakut	0.94	1.20	15	
Yi	0.99	1.19	15	
Yoruba	1.16	1.76	0	27.40

Table S2: Genetic skin pigmentation score as compared to values from Biasutti (PARRA *et al.*, 2004; LAO *et al.*, 2007) and JABLONSKI and CHAPLIN (2000). We also calculate a genetic skin pigmentation score including previously reported associations at KITLG and OCA2 for comparison. See also Figures S12 and S13.

	Observed	Expected	Variance	Z	p
Europe	-0.52	-0.56	0.0030	0.58	0.563573
Middle East	-0.61	-0.64	0.0025	0.50	0.620092
Central Asia	-0.57	-0.54	0.0025	-0.65	0.514631
East Asia	-0.69	-0.69	0.0068	0.01	0.990422
Americas	-0.71	-0.61	0.0234	-0.62	0.532068
Oceania	-0.59	-0.65	0.0312	0.31	0.757571
Africa	-0.95	-0.64	0.0238	-2.00	0.045035

Table S3: Conditional analysis at the regional level for the height dataset

	Observed	Expected	Variance	Z	p
Adygei	-0.46	-0.52	0.0058	0.81	0.420047
Balochi	-0.61	-0.59	0.0039	-0.34	0.732533
BantuKenya	-0.95	-0.88	0.0074	-0.89	0.370989
BantuSouthAfrica	-0.87	-0.93	0.0087	0.57	0.565944
Basque	-0.58	-0.49	0.0058	-1.14	0.253001
Bedouin	-0.67	-0.60	0.0040	-1.16	0.245140
BiakaPygmy	-0.95	-0.92	0.0089	-0.24	0.813609
Brahui	-0.66	-0.59	0.0041	-1.02	0.306021
Burusho	-0.65	-0.54	0.0049	-1.56	0.119344
Cambodian	-0.65	-0.72	0.0075	0.81	0.419708
Colombian	-0.73	-0.68	0.0149	-0.37	0.712734
Dai	-0.70	-0.72	0.0074	0.27	0.788165
Daur	-0.77	-0.71	0.0068	-0.73	0.466617
Druze	-0.63	-0.57	0.0039	-1.03	0.303758
French	-0.31	-0.55	0.0032	4.19	0.000028
Han	-0.79	-0.70	0.0018	-2.16	0.030910
Hazara	-0.61	-0.62	0.0040	0.09	0.932089
Hezhen	-0.63	-0.70	0.0070	0.79	0.431967
Italian	-0.57	-0.53	0.0067	-0.45	0.653982
Japanese	-0.78	-0.72	0.0036	-0.99	0.320879
Kalash	-0.31	-0.52	0.0110	1.92	0.054613
Karitiana	-0.78	-0.65	0.0205	-0.87	0.383564
Lahu	-0.85	-0.70	0.0113	-1.46	0.144720
Makrani	-0.65	-0.60	0.0038	-0.75	0.454499
Mandenka	-0.99	-0.88	0.0058	-1.45	0.146046
Maya	-0.61	-0.70	0.0070	1.05	0.295347
MbutiPygmy	-1.08	-0.95	0.0152	-1.12	0.262901
Melanesian	-0.60	-0.63	0.0199	0.24	0.810697
Miao	-0.65	-0.72	0.0065	0.84	0.398726
Mongola	-0.72	-0.69	0.0057	-0.43	0.664301
Mozabite	-0.66	-0.65	0.0071	-0.06	0.952256
Naxi	-0.52	-0.68	0.0088	1.67	0.095674
Orcadian	-0.38	-0.48	0.0069	1.20	0.228890
Oroqen	-0.66	-0.73	0.0075	0.85	0.397609
Palestinian	-0.50	-0.65	0.0031	2.82	0.004789
Papuan	-0.59	-0.59	0.0235	0.02	0.984319
Pathan	-0.45	-0.57	0.0042	1.86	0.063140
Pima	-0.84	-0.67	0.0152	-1.42	0.156243
Russian	-0.46	-0.47	0.0048	0.13	0.893316
San	-0.98	-0.95	0.0278	-0.16	0.873515
Sardinian	-0.79	-0.51	0.0053	-3.88	0.000105
She	-0.63	-0.72	0.0073	1.03	0.303926
Sindhi	-0.62	-0.57	0.0039	-0.69	0.487958
Surui	-0.58	-0.70	0.0253	0.78	0.433809
Tu	-0.65	-0.69	0.0062	0.48	0.630360
Tujia	-0.69	-0.72	0.0062	0.37	0.712043
Tuscan	-0.63	-0.52	0.0113	-1.00	0.318431
Uygur	-0.57	-0.59	0.0074	0.19	0.847180
Xibo	-0.73	-0.70	0.0068	-0.30	0.760787
Yakut	-0.70	-0.66	0.0059	-0.49	0.626665
Yi	-0.57	-0.71	0.0068	1.72	0.084613
Yoruba	-0.85	-0.96	0.0048	1.59	0.111605

Table S4: Conditional analysis at the individual population level for the height dataset

	Observed	Expected	Variance	Z	p
Europe	-0.18	-0.03	0.0030	-2.75	0.005935
Middle East	0.01	0.01	0.0030	-0.14	0.889313
Central Asia	0.23	0.25	0.0027	-0.24	0.813941
East Asia	1.02	0.82	0.0086	2.15	0.031684
Americas	1.03	0.68	0.0333	1.91	0.056218
Oceania	0.93	0.80	0.0392	0.68	0.493996
Africa	1.13	0.63	0.0506	2.21	0.027041

Table S5: Conditional analysis at the regional level for the skin pigmentation dataset

	Observed	Expected	Variance	Z	p
Adygei	-0.09	-0.04	0.0062	-0.64	0.525090
Balochi	0.14	0.13	0.0046	0.20	0.838807
BantuKenya	1.21	1.08	0.0128	1.16	0.247685
BantuSouthAfrica	1.13	1.17	0.0144	-0.32	0.746846
Basque	-0.20	-0.17	0.0063	-0.38	0.704915
Bedouin	0.07	-0.05	0.0049	1.74	0.081934
BiakaPygmy	1.22	1.11	0.0143	0.93	0.350280
Brahui	0.06	0.13	0.0050	-0.98	0.326220
Burusho	0.15	0.26	0.0057	-1.53	0.125199
Cambodian	1.04	0.92	0.0096	1.16	0.244370
Colombian	1.02	1.10	0.0188	-0.59	0.553594
Dai	1.04	1.04	0.0100	0.04	0.965427
Daur	1.05	0.99	0.0091	0.61	0.543435
Druze	-0.10	-0.09	0.0045	-0.20	0.844204
French	-0.13	-0.20	0.0039	1.22	0.222146
Han	1.05	1.06	0.0026	-0.33	0.744919
Hazara	0.45	0.49	0.0052	-0.64	0.519429
Hezhen	0.99	1.05	0.0100	-0.64	0.519743
Italian	-0.24	-0.12	0.0082	-1.26	0.207688
Japanese	1.02	1.06	0.0048	-0.54	0.586193
Kalash	0.12	0.09	0.0134	0.21	0.833512
Karitiana	1.05	1.07	0.0254	-0.10	0.919046
Lahu	1.02	1.03	0.0143	-0.09	0.924466
Makrani	0.22	0.14	0.0049	1.17	0.240751
Mandenka	1.19	1.12	0.0102	0.64	0.525346
Maya	0.99	0.85	0.0088	1.43	0.151325
MbutiPygmy	1.20	1.12	0.0252	0.50	0.616103
Melanesian	0.96	0.87	0.0284	0.51	0.611448
Miao	1.04	1.06	0.0087	-0.25	0.803925
Mongola	1.06	0.96	0.0079	1.17	0.240282
Mozabite	0.11	0.19	0.0096	-0.81	0.416292
Naxi	0.88	1.02	0.0112	-1.32	0.185939
Orcadian	-0.17	-0.12	0.0078	-0.60	0.548105
Oroqen	1.07	1.02	0.0098	0.49	0.626438
Palestinian	-0.05	0.02	0.0036	-1.16	0.247433
Papuan	0.91	0.89	0.0326	0.10	0.919411
Pathan	0.15	0.16	0.0048	-0.11	0.911081
Pima	0.99	0.97	0.0200	0.10	0.923266
Russian	-0.18	-0.05	0.0053	-1.89	0.058819
San	0.81	1.07	0.0405	-1.31	0.191477
Sardinian	-0.15	-0.20	0.0062	0.55	0.583092
She	1.13	1.04	0.0095	0.94	0.345316
Sindhi	0.27	0.22	0.0052	0.67	0.501326
Surui	1.08	0.99	0.0365	0.45	0.652715
Tu	0.98	0.99	0.0086	-0.12	0.905437
Tujia	1.12	1.07	0.0083	0.54	0.585802
Tuscan	-0.29	-0.13	0.0128	-1.36	0.174556
Uygur	0.57	0.50	0.0085	0.73	0.465493
Xibo	1.01	0.96	0.0086	0.56	0.577734
Yakut	0.94	0.86	0.0078	0.88	0.377402
Yi	0.99	1.02	0.0087	-0.32	0.745359
Yoruba	1.16	1.23	0.0086	-0.66	0.508211

Table S6: Conditional analysis at the individual population level for the skin pigmentation dataset

	Observed	Expected	Variance	Z	p
Europe	-2.10	-2.10	0.0092	-0.02	0.986166
Middle East	-2.14	-2.17	0.0080	0.41	0.683249
Central Asia	-2.19	-2.24	0.0077	0.63	0.528510
East Asia	-2.62	-2.24	0.0224	-2.53	0.011480
Americas	-2.29	-2.35	0.0788	0.21	0.832973
Oceania	-2.36	-2.51	0.1036	0.48	0.632885
Africa	-2.47	-2.37	0.0826	-0.33	0.742233

Table S7: Conditional analysis at the regional level for the BMI dataset

	Observed	Expected	Variance	Z	p
Adygei	-2.25	-2.09	0.0193	-1.13	0.259383
Balochi	-2.28	-2.26	0.0124	-0.12	0.906459
BantuKenya	-2.34	-2.45	0.0252	0.71	0.475069
BantuSouthAfrica	-2.49	-2.43	0.0297	-0.35	0.724172
Basque	-2.29	-2.07	0.0182	-1.59	0.110893
Bedouin	-2.19	-2.11	0.0128	-0.68	0.493665
BiakaPygmy	-2.47	-2.47	0.0290	-0.05	0.959030
Brahui	-2.29	-2.22	0.0143	-0.55	0.582714
Burusho	-2.34	-2.18	0.0147	-1.33	0.183764
Cambodian	-2.77	-2.73	0.0247	-0.27	0.788781
Colombian	-1.77	-2.44	0.0498	2.98	0.002842
Dai	-2.82	-2.73	0.0248	-0.60	0.545319
Daur	-2.49	-2.51	0.0237	0.11	0.913165
Druze	-2.09	-2.15	0.0128	0.58	0.561551
French	-2.07	-2.09	0.0104	0.20	0.844786
Han	-2.65	-2.65	0.0061	0.06	0.951026
Hazara	-2.20	-2.32	0.0124	1.04	0.300183
Hezhen	-2.68	-2.47	0.0241	-1.39	0.165463
Italian	-2.24	-2.09	0.0220	-1.06	0.287855
Japanese	-2.43	-2.59	0.0118	1.49	0.137107
Kalash	-1.75	-2.19	0.0364	2.28	0.022595
Karitiana	-2.43	-2.37	0.0680	-0.25	0.801202
Lahu	-3.04	-2.68	0.0378	-1.86	0.063279
Makrani	-2.37	-2.21	0.0122	-1.46	0.143943
Mandenka	-2.36	-2.44	0.0193	0.61	0.540442
Maya	-2.36	-2.20	0.0224	-1.12	0.264309
MbutiPygmy	-2.50	-2.49	0.0536	-0.07	0.947811
Melanesian	-2.25	-2.44	0.0676	0.74	0.459264
Miao	-2.75	-2.69	0.0218	-0.40	0.689780
Mongola	-2.69	-2.43	0.0192	-1.92	0.054866
Mozabite	-2.23	-2.15	0.0247	-0.49	0.625131
Naxi	-2.44	-2.70	0.0288	1.56	0.119884
Orcadian	-2.12	-2.07	0.0215	-0.37	0.709254
Oroqen	-2.31	-2.50	0.0241	1.18	0.237730
Palestinian	-2.04	-2.16	0.0097	1.13	0.258466
Papuan	-2.46	-2.37	0.0791	-0.33	0.743669
Pathan	-2.14	-2.21	0.0135	0.64	0.523054
Pima	-2.54	-2.30	0.0489	-1.10	0.271635
Russian	-1.84	-2.13	0.0151	2.37	0.017852
San	-2.65	-2.47	0.0894	-0.62	0.535534
Sardinian	-2.10	-2.12	0.0166	0.11	0.910561
She	-2.94	-2.70	0.0245	-1.54	0.123110
Sindhi	-2.22	-2.25	0.0135	0.33	0.741178
Surui	-2.33	-2.26	0.0834	-0.24	0.808288
Tu	-2.64	-2.57	0.0206	-0.52	0.604447
Tujia	-2.61	-2.63	0.0201	0.16	0.875686
Tuscan	-1.93	-2.13	0.0355	1.10	0.273049
Uygur	-2.11	-2.34	0.0228	1.54	0.124610
Xibo	-2.20	-2.50	0.0212	2.07	0.038059
Yakut	-2.27	-2.32	0.0192	0.40	0.690521
Yi	-2.84	-2.59	0.0219	-1.73	0.082850
Yoruba	-2.46	-2.38	0.0170	-0.61	0.541470

Table S8: Conditional analysis at the individual population level for the BMI dataset

	Observed	Expected	Variance	Z	p
Europe	0.72	1.00	0.0097	-2.79	0.005315
Middle East	1.02	0.81	0.0079	2.37	0.017671
Central Asia	0.92	0.88	0.0078	0.53	0.598600
East Asia	0.81	0.98	0.0216	-1.16	0.244114
Americas	0.85	0.79	0.0773	0.21	0.831384
Oceania	1.00	0.96	0.1013	0.11	0.912009
Africa	1.08	1.11	0.0813	-0.10	0.918661

Table S9: Conditional analysis at the regional level for the T2D dataset.

	Observed	Expected	Variance	Z	p
Adygei	0.84	0.88	0.0185	-0.30	0.763788
Balochi	0.81	0.93	0.0126	-1.02	0.306287
BantuKenya	1.21	1.05	0.0262	1.00	0.318499
BantuSouthAfrica	1.04	1.09	0.0300	-0.28	0.780539
Basque	0.76	0.72	0.0190	0.28	0.777993
Bedouin	0.94	0.99	0.0131	-0.39	0.694506
BiakaPygmy	0.94	1.09	0.0299	-0.82	0.409850
Brahui	0.89	0.92	0.0135	-0.27	0.784720
Burusho	0.91	0.93	0.0163	-0.16	0.875229
Cambodian	0.89	0.86	0.0254	0.22	0.829257
Colombian	1.03	0.88	0.0499	0.68	0.499559
Dai	0.89	0.82	0.0245	0.42	0.678026
Daur	0.72	0.79	0.0227	-0.49	0.627610
Druze	1.02	0.90	0.0130	1.06	0.291242
French	0.66	0.79	0.0110	-1.24	0.213762
Han	0.78	0.83	0.0063	-0.68	0.494587
Hazara	0.89	0.86	0.0134	0.25	0.799840
Hezhen	0.98	0.77	0.0256	1.37	0.172155
Italian	0.72	0.79	0.0223	-0.51	0.610084
Japanese	0.74	0.81	0.0124	-0.66	0.509728
Kalash	0.87	0.94	0.0380	-0.31	0.753570
Karitiana	0.66	0.87	0.0681	-0.81	0.419843
Lahu	0.81	0.83	0.0383	-0.11	0.909731
Makrani	0.76	0.94	0.0125	-1.66	0.097230
Mandenka	1.00	1.11	0.0203	-0.80	0.425733
Maya	0.98	0.90	0.0230	0.51	0.609012
MbutiPygmy	1.04	1.03	0.0518	0.04	0.971663
Melanesian	1.15	0.83	0.0695	1.24	0.215185
Miao	0.95	0.80	0.0225	0.99	0.322573
Mongola	0.74	0.79	0.0197	-0.42	0.676445
Mozabite	1.07	0.91	0.0239	1.08	0.279820
Naxi	0.94	0.80	0.0284	0.79	0.431243
Orcadian	0.80	0.73	0.0219	0.45	0.655908
Oroqen	0.82	0.79	0.0239	0.24	0.812587
Palestinian	1.05	0.89	0.0101	1.60	0.110698
Papuan	0.84	1.16	0.0802	-1.13	0.260430
Pathan	1.09	0.89	0.0138	1.74	0.081439
Pima	1.25	0.88	0.0543	1.57	0.116103
Russian	0.70	0.76	0.0158	-0.52	0.601197
San	1.21	0.98	0.0869	0.79	0.432263
Sardinian	0.69	0.82	0.0174	-0.99	0.322545
She	0.68	0.81	0.0246	-0.83	0.408294
Sindhi	1.10	0.92	0.0133	1.59	0.112277
Surui	0.36	0.95	0.0870	-2.01	0.044245
Tu	0.80	0.78	0.0216	0.15	0.877641
Tujia	0.88	0.81	0.0200	0.46	0.642326
Tuscan	0.63	0.85	0.0362	-1.11	0.265107
Uygur	0.99	0.84	0.0236	1.00	0.318441
Xibo	0.76	0.80	0.0223	-0.30	0.765435
Yakut	0.67	0.83	0.0208	-1.12	0.263127
Yi	0.75	0.83	0.0221	-0.56	0.573343
Yoruba	1.13	1.08	0.0168	0.33	0.740815

Table S10: Conditional analysis at the individual population level for the T2D dataset.

	Observed	Expected	Variance	Z	p
Europe	0.58	0.75	0.0188	-1.22	0.220764
Middle East	0.59	0.62	0.0166	-0.30	0.763396
Central Asia	0.82	0.55	0.0151	2.21	0.027100
East Asia	0.53	0.94	0.0459	-1.90	0.056975
Americas	0.71	0.67	0.1529	0.09	0.925298
Oceania	0.34	0.59	0.1959	-0.55	0.581088
Africa	0.20	0.52	0.1535	-0.83	0.409211

Table S11: Conditional analysis at the regional level for the CD dataset.

	Observed	Expected	Variance	Z	p
Adygei	0.52	0.69	0.0378	-0.89	0.370810
Balochi	0.83	0.79	0.0236	0.22	0.823291
BantuKenya	-0.09	0.20	0.0485	-1.31	0.191212
BantuSouthAfrica	0.08	0.21	0.0566	-0.56	0.575256
Basque	0.73	0.57	0.0381	0.80	0.423331
Bedouin	0.63	0.62	0.0254	0.05	0.963091
BiakaPygmy	0.15	0.28	0.0574	-0.54	0.587432
Brahui	0.96	0.80	0.0273	0.97	0.332305
Burusho	0.63	0.74	0.0315	-0.61	0.538809
Cambodian	0.34	0.51	0.0500	-0.72	0.470030
Colombian	0.61	0.76	0.0986	-0.46	0.647213
Dai	0.49	0.46	0.0475	0.14	0.892361
Daur	0.53	0.63	0.0457	-0.47	0.637409
Druze	0.55	0.68	0.0254	-0.79	0.430724
French	0.37	0.67	0.0219	-2.04	0.041575
Han	0.53	0.52	0.0123	0.05	0.957247
Hazara	0.82	0.68	0.0266	0.84	0.401314
Hezhen	0.61	0.58	0.0494	0.16	0.870215
Italian	0.33	0.65	0.0436	-1.51	0.132112
Japanese	0.44	0.59	0.0239	-0.94	0.349319
Kalash	0.50	0.76	0.0726	-0.97	0.332454
Karitiana	0.59	0.77	0.1389	-0.50	0.615762
Lahu	0.71	0.42	0.0765	1.02	0.305499
Makrani	0.83	0.77	0.0247	0.34	0.734710
Mandenka	0.18	0.11	0.0374	0.39	0.696885
Maya	0.78	0.64	0.0442	0.67	0.505123
MbutiPygmy	0.37	0.19	0.1039	0.56	0.572635
Melanesian	0.55	0.39	0.1321	0.44	0.656722
Miao	0.30	0.50	0.0450	-0.95	0.341732
Mongola	0.94	0.59	0.0392	1.73	0.082971
Mozabite	0.47	0.55	0.0468	-0.36	0.719113
Naxi	0.36	0.48	0.0548	-0.48	0.627989
Orcadian	0.51	0.62	0.0459	-0.51	0.609310
Oroqen	0.53	0.62	0.0500	-0.43	0.664401
Palestinian	0.69	0.60	0.0197	0.60	0.546529
Papuan	0.13	0.50	0.1611	-0.91	0.364157
Pathan	0.68	0.82	0.0275	-0.79	0.426714
Pima	0.63	0.73	0.0996	-0.33	0.743225
Russian	0.73	0.56	0.0319	0.96	0.338325
San	0.51	0.19	0.1691	0.79	0.431902
Sardinian	0.79	0.52	0.0342	1.44	0.150713
She	0.68	0.49	0.0498	0.86	0.391621
Sindhi	0.96	0.72	0.0254	1.47	0.142144
Surui	0.95	0.72	0.1718	0.54	0.586772
Tu	0.30	0.53	0.0432	-1.12	0.263329
Tujia	0.28	0.54	0.0411	-1.27	0.204893
Tuscan	0.68	0.62	0.0724	0.23	0.817115
Uygur	1.21	0.66	0.0466	2.57	0.010253
Xibo	1.02	0.59	0.0458	2.01	0.044766
Yakut	0.59	0.68	0.0401	-0.43	0.664595
Yi	0.41	0.53	0.0413	-0.56	0.577533
Yoruba	0.17	0.10	0.0329	0.39	0.695121

Table S12: Conditional analysis at the individual population level for the CD dataset.

	Observed	Expected	Variance	Z	p
Europe	0.30	0.55	0.0144	-2.08	0.037863
Middle East	0.43	0.40	0.0127	0.23	0.817361
Central Asia	0.67	0.47	0.0123	1.74	0.081000
East Asia	0.51	0.91	0.0343	-2.12	0.033839
Americas	0.43	0.58	0.1149	-0.43	0.667025
Oceania	0.95	0.55	0.1560	1.03	0.305336
Africa	0.33	0.55	0.1229	-0.62	0.532409

Table S13: Conditional analysis at the regional level for the UC dataset.

	Observed	Expected	Variance	Z	p
Adygei	0.40	0.52	0.0287	-0.75	0.454405
Balochi	0.77	0.65	0.0192	0.80	0.425065
BantuKenya	0.17	0.26	0.0381	-0.44	0.658187
BantuSouthAfrica	0.10	0.31	0.0459	-1.00	0.315538
Basque	0.28	0.36	0.0281	-0.44	0.658278
Bedouin	0.51	0.43	0.0203	0.58	0.561443
BiakaPygmy	0.33	0.37	0.0435	-0.19	0.850748
Brahui	0.69	0.70	0.0212	-0.04	0.966201
Burusho	0.68	0.61	0.0246	0.46	0.644084
Cambodian	0.55	0.47	0.0400	0.37	0.711734
Colombian	0.35	0.50	0.0769	-0.56	0.573665
Dai	0.50	0.46	0.0380	0.24	0.813796
Daur	0.45	0.59	0.0377	-0.69	0.492558
Druze	0.43	0.47	0.0196	-0.31	0.756796
French	0.29	0.38	0.0166	-0.63	0.527907
Han	0.32	0.53	0.0096	-2.13	0.033117
Hazara	0.74	0.56	0.0210	1.26	0.207136
Hezhen	0.75	0.47	0.0372	1.44	0.150826
Italian	0.21	0.36	0.0347	-0.79	0.431323
Japanese	0.45	0.51	0.0192	-0.46	0.642099
Kalash	0.26	0.62	0.0563	-1.52	0.127606
Karitiana	0.39	0.41	0.1076	-0.06	0.953921
Lahu	0.63	0.44	0.0600	0.77	0.441513
Makrani	0.72	0.64	0.0194	0.57	0.566336
Mandenka	0.18	0.25	0.0296	-0.37	0.714488
Maya	0.52	0.43	0.0360	0.46	0.646610
MbutiPygmy	0.62	0.40	0.0816	0.78	0.437484
Melanesian	1.14	0.68	0.1018	1.44	0.150504
Miao	0.35	0.49	0.0351	-0.73	0.465745
Mongola	0.69	0.51	0.0305	1.01	0.314222
Mozabite	0.20	0.39	0.0368	-0.98	0.329376
Naxi	0.44	0.51	0.0445	-0.30	0.762696
Orcadian	-0.00	0.39	0.0351	-2.12	0.034016
Oroqen	0.57	0.56	0.0373	0.06	0.950027
Palestinian	0.56	0.46	0.0148	0.84	0.399894
Papuan	0.76	0.96	0.1195	-0.55	0.579229
Pathan	0.37	0.68	0.0204	-2.15	0.031576
Pima	0.49	0.50	0.0803	-0.04	0.971033
Russian	0.61	0.31	0.0235	1.94	0.052946
San	0.70	0.40	0.1256	0.85	0.393139
Sardinian	0.42	0.29	0.0282	0.76	0.449837
She	0.52	0.46	0.0400	0.31	0.753611
Sindhi	0.86	0.59	0.0199	1.93	0.053454
Surui	0.42	0.53	0.1337	-0.29	0.770331
Tu	0.38	0.48	0.0328	-0.56	0.574730
Tujia	0.51	0.43	0.0324	0.46	0.649040
Tuscan	0.16	0.45	0.0582	-1.17	0.240321
Uygur	0.91	0.53	0.0358	2.00	0.045954
Xibo	0.68	0.49	0.0343	0.99	0.323788
Yakut	0.49	0.62	0.0319	-0.74	0.461639
Yi	0.46	0.47	0.0335	-0.05	0.957944
Yoruba	0.23	0.17	0.0250	0.37	0.708225

Table S14: Conditional analysis at the individual population level for the UC dataset

Table S15: Corresponding β statistics for all analyses presented in 2.

Phenotype	SUMPC1	SUMPC2	WINPC1	WINPC2	Latitude
Height	-0.22 (0.11)	0.003 (0.93)	-0.12 (0.37)	0.38 (0.008)	0.12 (0.34)
Skin Pigmentation	0.31(0.025)	0.07 (0.62)	0.27 (0.061)	-0.11 (0.38)	-0.36 (0.009)
Body Mass Index	-0.22 (0.15)	0.04 (0.74)	-0.17 (0.29)	0.24 (0.11)	0.21 (0.21)
Type 2 Diabetes	0.10 (0.36)	0.10 (0.42)	0.14 (0.25)	-0.07 (0.56)	-0.20 (0.093)
Crohn's Disease	0.25 (0.066)	-0.30 (0.033)	0.01 (0.86)	-0.28 (0.053)	0.08 (0.61)
Ulcerative Colitis	0.17 (0.22)	-0.29 (0.046)	0.06 (0.61)	-0.21 (0.15)	-0.11(0.40)

Table S16: Corresponding ρ statistics for all analyses presented in 2.

Phenotype	SUMPC1	SUMPC2	WINPC1	WINPC2	Latitude
Height	-0.15 (0.24)	-0.006 (0.89)	-0.13 (0.31)	0.32 (0.025)	0.10 (0.42)
Skin Pigmentation	0.28(0.045)	0.008 (0.98)	0.24 (0.097)	-0.11 (0.40)	-0.25 (0.087)
Body Mass Index	-0.24 (0.12)	0.07 (0.57)	-0.11 (0.57)	0.25 (0.097)	0.14 (0.47)
Type 2 Diabetes	0.12 (0.30)	0.11 (0.34)	0.09 (0.39)	-0.07 (0.56)	-0.21 (0.08)
Crohn's Disease	0.21 (0.12)	-0.29 (0.039)	0.01 (0.86)	-0.22 (0.12)	0.08 (0.63)
Ulcerative Colitis	0.19 (0.17)	-0.23 (0.12)	0.18 (0.65)	-0.15 (0.30)	-0.17(0.20)