ARTICLE

Assessment of the potential role of natural selection in type 2 diabetes and related traits across human continental ancestry groups: comparison of phenotypic with genotypic divergence

Robert L. Hanson[1] · Cristopher V. Van Hout[2] · Wen-Chi Hsueh[1] · Alan R. Shuldiner[2] · Sayuko Kobes[1] ·
Madhumita Sinha[1] · Leslie J. Baier[1] · Regeneron Genetics Center[2] · William C. Knowler[1]

Abstract

Aims/hypothesis Prevalence of type 2 diabetes differs among human ancestry groups, and many hypotheses invoke differential natural selection to account for these differences. We sought to assess the potential role of differential natural selection across major continental ancestry groups for diabetes and related traits, by comparison of genetic and phenotypic differences.

Methods This was a cross-sectional comparison among 734 individuals from an urban sample (none of whom was more closely related to another than third-degree relatives), including 83 African Americans, 523 American Indians and 128 European Americans. Participants were not recruited based on diabetes status or other traits. BMI was calculated, and diabetes was diagnosed by a 75 g oral glucose tolerance test. In those with normal glucose tolerance ($n = 434$), fasting insulin and 30 min post-load insulin, adjusted for 30 min glucose, were taken as measures of insulin resistance and secretion, respectively. Whole exome sequencing was performed, resulting in 97,388 common (minor allele frequency $\geq 5\%$) variants; the coancestry coefficient (F_{ST}) was calculated across all markers as a measure of genetic divergence among ancestry groups. The phenotypic divergence index (P_{ST}) was also calculated from the phenotypic differences and heritability (which was estimated from genetic relatedness calculated empirically across all markers in 761 American Indian participants prior to the exclusion of close relatives). Under evolutionary neutrality, the expectation is $P_{ST} = F_{ST}$, while for traits under differential selection P_{ST} is expected to be significantly greater than F_{ST}. A bootstrap procedure was used to test the hypothesis $P_{ST} = F_{ST}$.

Results With adjustment for age and sex, prevalence of type 2 diabetes was 34.0% in American Indians, 12.4% in African Americans and 10.4% in European Americans ($p = 2.9 \times 10^{-10}$ for difference among groups). Mean BMI was 36.3, 33.4 and 33.0 kg/m^2, respectively ($p = 1.9 \times 10^{-7}$). Mean fasting insulin was 63.8, 48.4 and 45.2 pmol/l ($p = 9.2 \times 10^{-5}$), while mean 30 min insulin was 559.8, 553.5 and 358.8 pmol/l, respectively ($p = 5.7 \times 10^{-8}$). F_{ST} across all markers was 0.130, while P_{ST} for liability to diabetes, adjusted for age and sex, was 0.149 ($p = 0.35$ for difference with F_{ST}). P_{ST} was 0.094 for BMI ($p = 0.54$), 0.095 for fasting insulin ($p = 0.54$) and 0.216 ($p = 0.18$) for 30 min insulin. For type 2 diabetes and BMI, the maximum divergence between populations was observed between American Indians and European Americans ($P_{ST-MAX} = 0.22$, $p = 0.37$, and $P_{ST-MAX} = 0.14$, $p = 0.61$), which suggests that a relatively modest 22% or 14% of the genetic variance, respectively, can potentially be explained by differential selection (assuming the absence of neutral drift).

✉ Robert L. Hanson
rhanson@phx.niddk.nih.gov

1 Phoenix Epidemiology and Clinical Research Branch, National Institute of Diabetes and Digestive and Kidney Diseases, Phoenix, AZ, USA

2 Regeneron Genetics Center, Tarrytown, NY, USA

Research in context

What is already known about this subject?

- Prevalence of type 2 diabetes and obesity differs across human continental ancestry groups

- In the USA, prevalence of diabetes and obesity is particularly high in American Indians, whereas it is low in European Americans and intermediate in African Americans

- Several hypotheses have invoked differences in natural selection across ancestry groups to explain these differences

What is the key question?

- Does phenotypic divergence for type 2 diabetes and related traits among American Indians, African Americans and European Americans exceed that expected from genotypic divergence, at markers throughout the genome, and trait heritability?

What are the new findings?

- Prevalence of type 2 diabetes differed significantly among ancestry groups, as did levels of BMI, fasting insulin and insulin secretion, consistent with previous epidemiologic findings

- Phenotypic divergence did not significantly exceed evolutionarily neutral expectations, based on genetic distance and heritability, for any of these traits

- This suggests that differences in natural selection are not necessary to explain the phenotypic differences among groups

How might this impact on clinical practice in the foreseeable future?

- Clinical investigations of differences in diabetes prevalence among these populations would do well to consider explanations other than differential natural selection

Conclusions/interpretation These analyses suggest that while type 2 diabetes and related traits differ significantly among continental ancestry groups, the differences are consistent with neutral expectations based on heritability and genetic distances. While these analyses do not exclude a modest role for natural selection, they do not support the hypothesis that differential natural selection is necessary to explain the phenotypic differences among these ancestry groups.

Abbreviations

F_{ST}	Coancestry coefficient
NHANES	National Health and Nutrition Examination Survey
P_{ST}	Phenotypic divergence index
Q_{ST}	Quantitative genetic divergence index
SWAI	Southwest American Indian

Introduction

Prevalence of type 2 diabetes varies among human continental ancestry groups, as does obesity, which is a strong risk factor for diabetes. In the USA, prevalence of diabetes and obesity is particularly high in American Indians, whereas prevalence is low in European Americans and intermediate in African Americans [1, 2]. Both type 2 diabetes and obesity are highly heritable [3, 4], and several hypotheses have invoked differences in natural selection across ancestry groups to explain differences in prevalence [5–13]. Recent genome-wide association studies have identified many variants reproducibly associated with both type 2 diabetes and obesity [14, 15]. Several investigators have analysed these established susceptibility loci for evidence of natural selection. Such studies have generally involved assessment of genetic signatures of recent selection or comparison of allele frequencies among ancestry groups [16–22]. Results of these studies are largely equivocal; however, both approaches are limited in their ability to detect selection on polygenic traits. An alternative approach involves comparison of genetic components of variance for the trait, among and within ancestry groups, with corresponding genotypic variance components across representative genomic markers [23, 24]. These variance components

methods are well suited for detection of polygenic selection that differs in magnitude across groups, and they do not require knowledge of specific susceptibility variants. They do require comparably measured phenotypic data, along with genotypic data, across diverse ancestry groups. In the present study, we compare phenotypic divergence for type 2 diabetes and related traits with genotypic divergence in a cohort that includes African American, American Indian and European American individuals who had undergone whole exome sequencing.

Methods

Participants and measures Participants were derived from a multiethnic study, conducted in urban Phoenix, Arizona, designed to identify determinants of diabetes and related traits; the methods have been previously described [25]. In brief, individuals were ≥18 years old, of any ethnicity, and participants were not recruited based on diabetes or other conditions. A large proportion of participants were American Indian, primarily from tribes of the Southwestern United States. The study was approved by the institutional review boards of the National Institute of Diabetes and Digestive and Kidney Diseases and the Phoenix Area Indian Health Service, and all participants gave informed consent. The present cross-sectional sample was derived from 1389 participants, examined in 2011–2016, who had relevant phenotypic data available, and data from whole exome sequencing. After exclusion of ten individuals who did not cluster with their primary self-reported ancestry group in principal components analyses, there were 88 individuals who were full-heritage African American by self-report, 761 individuals who were full-heritage American Indian and 129 individuals who were full-heritage European American. Since some analyses may be influenced by presence of closely related individuals, genetic relatedness was calculated between pairs of individuals using PREST (version 3.02) [26], and a set of 'unrelated' individuals was selected by randomly excluding one member of each pair in whom the observed proportion of alleles shared identical by descent was >0.14. (This excludes individuals who are second-degree relatives or closer to another individual in the sample.) This resulted in 83 African Americans, 523 American Indians and 128 European Americans. Characteristics of individuals are shown in electronic supplementary material (ESM) Table 1, and a principal components plot is shown in ESM Fig. 1.

Fasting plasma glucose and HbA$_{1c}$ were measured, and a 75 g oral glucose tolerance test was administered to those without a previous diagnosis of diabetes, with glucose concentrations measured 30 min and 2 h after the oral glucose load. Individuals were classified as having diabetes if they had a previous diagnosis by self-report, fasting plasma glucose ≥7.0 mmol/l, 2 h plasma glucose ≥11.1 mmol/l or HbA$_{1c}$ ≥ 6.5% (48 mmol/l) [27]. Serum insulin concentrations were

measured by immunoassay (Tosoh Bioscience, Tokyo, Japan); fasting serum insulin and 30 min serum insulin, adjusted for 30 min glucose level, were taken as measures of insulin resistance and insulin secretion, respectively. Analyses of insulin measures were restricted to those with normal glucose tolerance (nondiabetic, and 2 h glucose <7.8 mmol/l), constituting 59 African Americans, 281 American Indians and 94 European Americans. Height and weight were measured for calculation of BMI. The maximum weight and contemporaneous height were also obtained by self-report. Analyses of BMI are generally shown based on self-reported maximum weight, as this was more strongly associated with diabetes.

Genotypes Whole exome sequencing in DNA derived from peripheral blood was conducted at Regeneron Genetics Center, as previously described [28, 29]. Sequencing was conducted using a Hi-Seq 2500 sequencer (Illumina, San Diego, CA, USA). Sequencing was part of a larger project involving 8137 individuals, 43 of whom were excluded for low-quality sequence data. In 98% of samples, at least 90% of the exome achieved at least 20x coverage. Analysis was restricted to variants with <10% missing genotype calls, that were within Hardy–Weinberg equilibrium ($p > 0.0001$ in full-heritage American Indians), that had concordance rates >97.5% in 100 duplicate samples and for which average minor allele frequency across ancestry groups was ≥5%. This resulted in 97,388 autosomal markers. Missing genotypes were imputed from phased haplotypic data in each ancestry group using BEAGLE (version 3.2.2) [30].

Measures of divergence The coancestry coefficient (F_{ST}) was calculated as a measure of genotypic divergence among ancestry groups. F_{ST} represents the proportion of variance in allele frequency in the total population that is explained by group membership. For each marker, we calculated F_{ST} by the method of moments [31, 32]. Across r ancestry groups, the mean squares among and within groups for a given allele u are, respectively:

$$MSA = \frac{1}{r-1} \sum_{i=1}^{r} 2n_i \left(p_{iu} - \overline{p}_u \right)^2$$

$$MSW = \frac{1}{\sum_{i=1}^{r}(2n_i - 1)} \sum_{i=1}^{r} 2n_i p_{iu}(1 - p_{iu})$$

where $2n_i$ is the total number of alleles measured in the ith group (twice the number of individuals), p_{iu} is the frequency of allele u in the ith group and $\overline{p}_u$ is the mean frequency of the u allele across groups. F_{ST} for a single marker with m alleles is:

$$F_{ST-M} = \frac{\sum_{u=1}^{m}(MSA_u - MSW_u)}{\sum_{u=1}^{m}[MSA_u + (2n_c - 1)MSW_u]}$$

where $2n_c = \frac{1}{(r-1)} \sum_{i=1}^{r} 2n_{ic}$, and $2n_{ic} = 2n_i - 4n_i^2 / \sum_{i=1}^{r} 2n_i$. The mean value of F_{ST-M} over all markers was taken as the overall F_{ST}. This mean marker-wise F_{ST} is comparable to the

phenotypic divergence measures described below and, thus, represents the expected value under neutrality [24, 33]. However, it tends to modestly underestimate the evolutionary distance, so we also report F_{ST} calculated by the 'ratio of averages' method, which provides a better estimate of this distance [34]. F_{ST} calculated from exome sequence data is generally comparable to that calculated from whole genome sequence data [35].

The quantitative genetic divergence index (Q_{ST}) is a measure of phenotypic divergence that is analogous to F_{ST} [23, 24]. For diploid organisms Q_{ST} is calculated as:

$$Q_{ST=} \frac{\sigma^2_{Ga}}{\sigma^2_{Ga} + 2\sigma^2_{Gw}}$$

where σ^2_{Ga} is the variance among ancestry groups attributable to additive genetic effects and σ^2_{Gw} is the genetic variance within groups. Under evolutionary neutrality, the expectation is that $Q_{ST} = F_{ST}$, whereas with diversifying selection (when differences in direction or magnitude of natural selection across groups drive phenotypic divergence), the expectation is $Q_{ST} > F_{ST}$ [24, 33]. With stabilising selection (when selection is of similar direction and magnitude across groups), then $Q_{ST} < F_{ST}$. Variance components for calculation of Q_{ST} are typically estimated by 'common garden' controlled breeding experiments.

In humans and other natural populations where controlled breeding experiments are not feasible, Q_{ST} can be approximated by the phenotypic divergence index (P_{ST}). This uses the total phenotypic components of variance among and within ancestry groups, σ^2_{Pa} and σ^2_{Pw}, respectively, rather than the genetic variance components. A general formula for P_{ST} is:

$$P_{ST=} \frac{c\sigma^2_{Pa}}{c\sigma^2_{Pa} + 2h^2\sigma^2_{Pw}} \tag{1}$$

where h^2 represents the proportion of the within-group phenotypic variance due to additive genetic effects (i.e., heritability) and c represents the proportion of the among-group variance due to genetic factors [36]. When h^2 and c are known from representative populations, then P_{ST}, calculated from eq. 1, is an unbiased estimate of Q_{ST}.

In the present study we estimate h^2 in genetically related individuals (i.e., in pedigree data without exclusion of close relatives), but c is unknown, as is often the case. In this situation, there are two widely used formulae for P_{ST}, which make different assumptions about c. The formula of Leinonen et al. is [37]:

$$P_{ST=} \frac{\sigma^2_{Pa}}{\sigma^2_{Pa} + 2h^2\sigma^2_{Pw}} \tag{2}$$

This assumes that $c = 1$, i.e., that all phenotypic differences among ancestry groups are due to genetic factors, and this

estimate represents the maximum possible value of Q_{ST} for a given h^2. This can be justified by the notion that P_{ST} is a screen for identifying traits potentially under differential natural selection. Other investigators, however, consider it more prudent to assume that $c = h^2$ [38], and this leads to:

$$P_{ST=} \frac{\sigma^2_{Pa}}{\sigma^2_{Pa} + 2\sigma^2_{Pw}} \tag{3}$$

This is more stringent in that it gives lower values of P_{ST} than eq. 2 (unless $h^2 = 1$). We calculate P_{ST} under both equations, and we present analyses under eq. 2 as the primary results with the recognition that these represent maximal estimates of P_{ST}. Results calculated under eq. 3 are presented in ESM Tables 5 and 8, and we conduct sensitivity analyses across a range of values for h^2 and c (including situations with $c < h^2$) to evaluate effects on the conclusions.

Statistical analyses Analyses were conducted in SAS (version 9.4; SAS Institute, Cary, NC, USA). Kernel density estimation (PROC KDE in SAS) was used to estimate nonparametric density functions. Phenotypic differences for continuous traits among ancestry groups were assessed using linear regression models with control for age and sex (and 30 min glucose, for analyses of 30 min insulin). A logistic regression model was used for analyses of diabetes. Heritability was assessed in the 761 American Indian participants (without exclusion of close relatives) using a linear mixed model. The total phenotypic variance was modelled as:

$$\sigma^2_P = \Phi\sigma^2_G + I\sigma^2_E$$

where σ^2_G is the variance potentially attributable to genetic factors, σ^2_E is the variance attributable to individual-level environmental factors, Φ is a matrix of the proportion of alleles shared identical by descent between pairs of individuals (estimated by PREST [26]) and I is an identity matrix. Heritability was calculated as $h^2 = \sigma^2_G/\sigma^2_P$. These analyses were conducted in SOLAR (version 8.1.1) with adjustment for age, sex and the first genetic principal component in American Indians (to account for potential population stratification), and a probit model was used to analyse liability to diabetes [39]. Confidence intervals were calculated with a likelihood-based method [40].

For continuous traits, variance components for calculation of P_{ST} were taken from the mean squares among and within ancestry groups, derived from the regression model with ancestry group as a fixed effect [38, 41]. Thus,

$$\sigma^2_{Pw} = MSW = SSE / \left[\sum_{i=1}^{r} n_i - (r-1) - 1 \right]$$

$$\sigma^2_{Pa} = \frac{MSA - MSW}{n_0}$$

where SSE is the sums of squares error from the regression, and n_i is the number of individuals in the ith ancestry group; if μ_i is the trait mean in the ith ancestry group, predicted by the regression model, and $\overline{\mu}$ is the total sample mean:

$$n_0 = \frac{1}{(r-1)} \left(\sum_{i=1}^r n_i - \frac{\sum_{i=1}^r n_i^2}{\sum_{i=1}^r n_i} \right)$$

$$MSA = \frac{1}{(r-1)} \sum_{i=1}^r n_i \left(\mu_i - \overline{\mu} \right)^2$$

For liability to diabetes, parameters were inferred from estimates of variance components derived from a probit mixed model in which ancestry group was a random effect (fit with PROC GLIMMIX in SAS). If s_{grp}^2 is the variance attributed to ancestry group and s_{rsd}^2 is the residual variance (from the Pearson χ^2 fit of the model), then $SSE = s_{rsd}^2 \sum_{i=1}^r n_i$ and $MSA = \frac{s_{grp}^2}{(r-1)} \sum_{i=1}^r n_i$. Similar approaches have been used elsewhere [42].

Parameters were estimated using a bootstrap procedure with 1000 iterations; 90% CIs were generated from centiles of the bootstrap distribution. Estimates of F_{ST} and P_{ST} depend on sample sizes, and their interpretation as measures of population divergence is most straightforward if sample sizes are equal in each ancestry group [32]. Therefore, in each bootstrap iteration, a sample size equal to that of the ancestry group with the smallest sample size was selected for each group. Although the mean marker-wise F_{ST} represents the expectation of P_{ST} under neutrality, there is substantial biological variability; despite polygenicity, the distribution is approximately that of F_{ST} for individual markers [24, 33]. Thus, comparison of P_{ST} with F_{ST} is most appropriately made across the single-marker distribution. Following Guo et al. [43], we generated this comparison from the distribution of F_{ST-M} shown in ESM Fig. 2. The proportion of markers for which the difference with the mean ($F_{ST-M} - F_{ST}$) was greater than the observed value of $P_{ST} - F_{ST}$ was taken as the empirical one-sided p value for the null hypothesis $P_{ST} = F_{ST}$ against the alternative $P_{ST} > F_{ST}$. The distribution of F_{ST-M} had a thick lower tail, with ~5% of markers having a value <0.01; thus, we did not test for the alternative $P_{ST} < F_{ST}$. Diversifying selection may act primarily on one ancestry group, and, in this situation, it may be most powerful to consider the maximum divergence of P_{ST} from F_{ST} across all pairwise comparisons; we report this value as P_{ST-MAX} (and assess its p value with correction for three pairwise comparisons). We also conduct a multivariable test of the null hypothesis of $P_{ST} = F_{ST}$ across any of the five major traits (diabetes, BMI, height, fasting insulin, 30 min insulin) using a method that combines p values with allowance for the correlations between traits [44].

F_{ST} **outlier analyses** For primary analyses, we used the 'robust' approach of comparing P_{ST} with F_{ST} taken across all available markers [45]. However, some of the markers themselves may have been subject to natural selection, and, to examine their potential influence, we conducted Bayesian outlier analysis of the F_{ST-M} distribution using BAYESCAN2 (version 2.01) [46]. This models the allele frequency differences among groups as a function of a population-specific F_{ST} component and a locus-specific F_{ST} component, subject to selection. Markers for which the locus-specific component is necessary are candidates for being under natural selection. As results depend on the specified prior odds of selection vs neutrality, we varied this parameter over a range of values and designated markers for which posterior odds were >1:1 as potentially under selection. Then, p values were calculated using the remaining, putatively neutral, markers.

Genetic admixture estimates We used ADMIXTURE (version 1.3.0) to obtain estimates of genetic admixture proportions for each individual, assuming three ancestral populations [47]. In these analyses we included data from individuals in the 1000 Genomes project to improve resolution (ftp://ftp.1000genomes.ebi.ac.uk/vol1/ftp/release/20130502/); data from the HapMap Yoruba in Ibadan, Nigeria, population and the Centre d'Etude du Polymorphism Humain Utah population were used as representative of African and European ancestry groups, respectively. To reduce the influence of linkage disequilibrium, we selected markers ~100 kb apart, after exclusion of markers that did not have consistent reference and alternative alleles between our exome sequence data and the 1000 Genomes data (and excluding A/T and C/G polymorphisms); this resulted in 14,672 markers.

Additional population data To assess applicability of our results in more general population data, we obtained data for individuals ≥18 years of age from a population-based study from a high-risk southwest American Indian (SWAI) population (4032 full-heritage American Indians) [48], and from the oral glucose tolerance subset of the 2005–2010 National Health and Nutrition Examination Survey (NHANES) (https://wwwn.cdc.gov/nchs/nhanes/Default.aspx), which is representative of the general US population. This included 1271 individuals of non-Hispanic black ancestry, taken as representative of African Americans, and 2905 individuals of non-Hispanic white ancestry, taken as representative of European Americans. We calculated P_{ST} across these three populations for diabetes and BMI. In these analyses, diabetes was diagnosed based on self-report, fasting plasma glucose or 2 h plasma glucose (as HbA_{1c} was not available in all participants). Exome sequencing data were available from 3435 SWAI participants, but, as such data were not available for NHANES participants, we could not directly compare P_{ST}
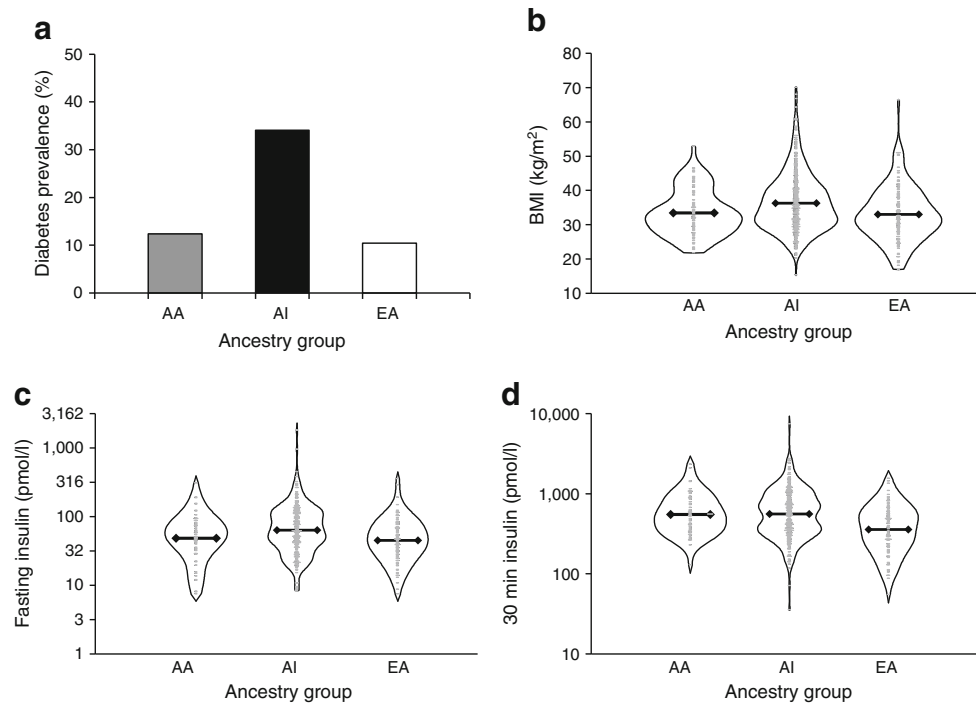
Fig. 1 (**a**) Age- and sex-adjusted prevalence of diabetes by ancestry group. Prevalence is adjusted to the mean age and sex distribution of the total sample. Prevalence was 12.4% in African Americans (AA), 34.0% in American Indians (AI) and 10.4% in European Americans (EA) ($p = 2.9 \times 10^{-10}$). (**b**) 'Bean' plot of the distribution of age- and sex-adjusted BMI by ancestry group. Symmetrical lines represent a nonparametric local density function, estimated with PROC KDE in SAS. Thick black horizonal lines represent the mean value. Grey horizontal lines represent individual data points, with the length of the line indicating the number of observations at each level. Mean BMI was 33.4 kg/m² in AA, 36.3 kg/m² in AI and 33.0 kg/m² in EA ($p = 1.9 \times 10^{-7}$). (**c**) 'Bean' plot of the distribution of age- and sex-adjusted fasting insulin level in individuals with normal glucose tolerance by ancestry group. Geometric mean fasting insulin was 48.4 pmol/l in AA, 63.8 pmol/l in AI and 45.2 pmol/l in EA ($p = 9.2 \times 10^{-5}$). (**d**) 'Bean' plot of the distribution of 30 min post-load insulin levels, adjusted for age, sex and 30 min glucose, in individuals with normal glucose tolerance by ancestry group. Geometric mean 30 min insulin was 553.5 pmol/l in AA, 559.8 pmol/l in AI and 358.8 pmol/l in EA ($p = 5.7 \times 10^{-8}$)

with F_{ST}. For an indirect comparison, we used genotypic data from the 1000 Genomes project, including the HapMap African Americans from the American Southwest population, as representative of African Americans, and the Centre d'Etude du Polymorphism Humain Utah population, as representative of European Americans. F_{ST} was calculated across these populations for 81,700 markers that had alleles called consistently between the exome sequence and the 1000 Genomes data.

Results

Phenotypic differences Phenotypic differences among ancestry groups are shown in Fig. 1. Age- and sex-adjusted prevalence of diabetes was highest in American Indians (34.0%) and lower in African Americans (12.4%) and European Americans (10.4%, $p = 2.9 \times 10^{-10}$ for difference among groups). Similarly, mean age- and sex-adjusted maximum BMI was 36.3 kg/m² in American Indians, 33.4 kg/m² in African Americans and 33.0 kg/m² in European Americans ($p = 1.9 \times 10^{-7}$). Height was also significantly different among

ancestry groups, with American Indians being shorter than African Americans and European Americans ($p = 1.9 \times 10^{-18}$, ESM Fig. 3). Among those with normal glucose tolerance, fasting serum insulin was higher in American Indians (geometric mean = 63.8 pmol/l adjusted for age and sex) than in African Americans (48.4 pmol/l) or European Americans (45.2 pmol/l, $p = 9.2 \times 10^{-5}$). The 30 min insulin, adjusted for age, sex and 30 min glucose, was lower in European Americans (358.8 pmol/l) than in African Americans (553.5 pmol/l) and American Indians (559.8 pmol/l, $p = 5.7 \times 10^{-8}$). With additional adjustment for BMI, differences in fasting insulin were largely attenuated ($p = 0.25$), while with additional adjustment for BMI and fasting insulin, differences in 30 min insulin remained statistically significant ($p = 2.3 \times 10^{-5}$).

Heritability In 761 American Indian participants, type 2 diabetes was highly familial; 77% of the liability was potentially due to genetic factors ($h^2 = 0.77$; 90% CI 0.34, 1.00; $p = 0.0017$). Similarly, significant familial aggregation was observed for maximum BMI ($h^2 = 0.36$; 0.19, 0.53; $p = 2.5 \times 10^{-4}$) and height ($h^2 = 0.71$; 0.54, 0.85; $p = 4.4 \times$

Table 1 Genotypic divergence (F_{ST}) and phenotypic divergence (P_{ST}) for type 2 diabetes and related traits among ancestry groups

Genotype/phenoptype	African American	American Indian	European American
Genotype F_{ST} = 0.1301 (0.1437)			
African American	–	0.1452	0.0949
American Indian	(0.1838)	–	0.1029
European American	(0.1142)	(0.1310)	–
Type 2 diabetes [h^2 = 0.77] P_{ST} = 0.1488 (0.3515), $P_{ST\text{-}MAX}$ = 0.2164 (0.3685)			
African American	–	0.1472	0.0038
American Indian	(0.4067)	–	0.2164[b]
European American	(0.8723)	(0.1421)	–
Maximum BMI[a] [h^2 = 0.36] P_{ST} = 0.0937 (0.5442), $P_{ST\text{-}MAX}$ = 0.1436 (0.6060)			
African American	–	0.1197	0.0078
American Indian	(0.4743)	–	0.1436[b]
European American	(0.8146)	(0.2669)	–
Height [h^2 = 0.71] P_{ST} = 0.1155 (0.4605), $P_{ST\text{-}MAX}$ = 0.1886 (0.4516)			
African American	–	0.1541	0.0026
American Indian	(0.3869)	–	0.1886[b]
European American	(0.8991)	(0.1815)	–
Fasting insulin [h^2 = 0.35] P_{ST} = 0.0946 (0.5388), $P_{ST\text{-}MAX}$ = 0.1305 (0.6556)			
African American	–	0.1285	0.0088
American Indian	(0.4450)	–	0.1305[b]
European American	(0.8039)	(0.2990)	–
30 min insulin [h^2 = 0.31] P_{ST} = 0.2156 (0.1754), $P_{ST\text{-}MAX}$ = 0.2801 (0.2292)			
African American	–	0.0157	0.2698
American Indian	(0.8146)	–	0.2801[b]
European American	(0.0730)	(0.0831)	–

Data represent F_{ST} or P_{ST} or p values for the null hypothesis $P_{ST} = F_{ST}$; data in each cell pertain to the comparisons between pairs of ancestry groups

For genotypes, values in parentheses represent F_{ST} calculated by the 'ratio of averages' method (see Methods/Measures of divergence)

For phenotypes, values in parentheses are p values for the null hypothesis $P_{ST} = F_{ST}$ vs the alternative hypothesis $P_{ST} > F_{ST}$. In the title rows, F_{ST}/P_{ST} values and p values pertain to the three-way comparisons. P_{ST} was calculated from eq. 2 (see Methods/Measures of divergence) using the specified value of h^2. $P_{ST\text{-}MAX}$ represents the P_{ST} value with the largest deviation from expected for any pair of ancestry groups; the p value is corrected for multiple comparisons

[a] From self-reported maximum weight

[b] The P_{ST} value corresponding to the maximum

10^{-11}). In 434 American Indian participants with normal glucose tolerance, significant heritability was observed for fasting insulin (h^2 = 0.35; 0.04, 0.65; p = 0.035) and 30 min insulin adjusted for 30 min glucose (h^2 = 0.31; 0.03, 0.66; p = 0.035). These estimates were made with adjustment for the first genetic principal component, which captures the major source of stratification in this population; additional sources of population stratification may be captured with additional principal components, but, given the small number of relative pairs, at the risk of model overspecification. To assess the robustness of the h^2 estimates, we repeated the analysis with adjustment for the first five genetic principal components. For most traits, the h^2 estimates were only modestly attenuated (ESM Table 2). The exception was 30 min insulin, for which h^2 approached 0 (where P_{ST} is of questionable meaning as a measure of selection), so the h^2 estimate for this trait is not robust.

Comparison of genotypic and phenotypic divergence Mean $F_{ST\text{-}M}$ among all three ancestry groups across all 97,388 markers was 0.130 (Table 1). Estimates of P_{ST} for each phenotype are shown in Table 1; none of these were significantly higher than F_{ST}. For type 2 diabetes P_{ST} = 0.149 (90% CI 0.038, 0.272; p = 0.35 for comparison with F_{ST}). For maximum BMI, P_{ST} = 0.094 (0.017, 0.184; p = 0.54), and for height P_{ST} = 0.116 (0.053, 0.189; p = 0.46). Among those with normal glucose tolerance, P_{ST} for fasting insulin was 0.095 (0.001, 0.214; p = 0.54), while for 30 min insulin P_{ST} = 0.216 (0.082, 0.358; p = 0.18). The multivariable test across all five traits was not significant (p = 0.46). The largest
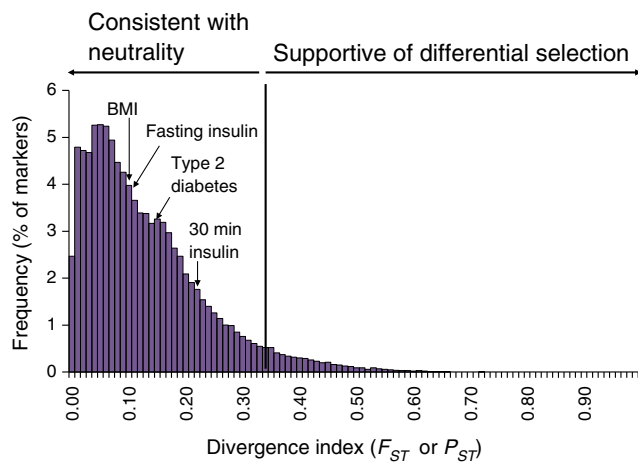
Fig. 2 Summary of results of primary analyses. The extent of divergence among ancestry groups (F_{ST} for genotypes, P_{ST} for phenotypes) is shown on the x-axis. The height of the bars on the y-axis represents the frequency (%) among 97,388 markers at each divergence level. The values of P_{ST} for each phenotype are indicated by arrows. Divergence levels above the critical value of 0.340 (indicated by the vertical line) are considered indicative of diversifying selection, while those below this value are more consistent with neutrality

departures from neutral expectations for pairs of ancestry groups generally occurred between American Indians and European Americans, but were not statistically significant; for diabetes $P_{ST-MAX} = 0.22$ ($p = 0.37$), while for BMI $P_{ST-MAX} = 0.14$ ($p = 0.61$). Similar results were obtained with directly measured BMI, with fasting insulin adjusted for BMI and with 30 min insulin adjusted for BMI and fasting insulin (ESM Table 3). Similar results were also obtained when men and women were analysed separately (ESM Fig. 4, ESM Table 4). When P_{ST} was calculated according to eq. 3, P_{ST} values tended to be lower than F_{ST} (ESM Table 5). A summary of the primary analyses is shown in Fig. 2.

In Bayesian outlier analyses, the number of markers potentially under selection ranged from 416 with prior odds for selection vs neutrality of 1:10, to 61,195 with prior odds of 4:3. Statistical significance levels were similar when restricted to putatively neutral markers regardless of prior odds (ESM Table 6). By analysis of individual admixture proportions, we estimated that, on average, 80% of the ancestry of African American participants derived from African sources, 96% of the ancestry of American Indian participants derived from Amerindian sources and 99% of the ancestry of European American participants derived from European sources. We repeated the divergence analyses with restriction to those whose genetic ancestry derived ≥85% from the continent corresponding to their stated ancestry group, constituting 31 African Americans, 475 American Indians and 126 European Americans. Similar results were obtained (ESM Table 7); for diabetes $P_{ST} = 0.143$ ($p = 0.41$), and, for BMI, $P_{ST} = 0.098$ ($p = 0.57$), while $F_{ST} = 0.144$.

Sensitivity analyses Results of sensitivity analyses, which calculate P_{ST} for different values of h^2 and c, are shown in Fig. 3. For diabetes, P_{ST} was generally less than the 95th centile of marker-wise F_{ST} (0.340), except when h^2 was low and c was high (e.g., $h^2 = 0.20$ and $c > 0.80$). For BMI, fasting insulin and 30 min insulin, P_{ST} did not exceed the critical value of 0.340 for any value of c, for any $h^2 \geq 0.20$.

Analyses of additional population data Results of analyses comparing African Americans and European Americans from

Fig. 3 (**a**) Sensitivity analyses showing values of P_{ST} for maximum BMI for various levels of h^2 and c. (**b**) Sensitivity analyses showing values of P_{ST} for type 2 diabetes for various levels of h^2 and c. (**c**) Sensitivity analyses showing values of P_{ST} for fasting insulin level for various levels of h^2 and c. (**d**) Sensitivity analyses showing values of P_{ST} for 30 min post-load insulin level for various levels of h^2 and c. Results are shown for $h^2 = 0.20$, the estimated value of h^2, its upper confidence limit and its lower confidence limit (if >0.20). Points shown as triangles are calculated under eq. 2, and are the same as those shown in Table 1, while points shown as circles are calculated under eq. 3 (ESM Table 5)
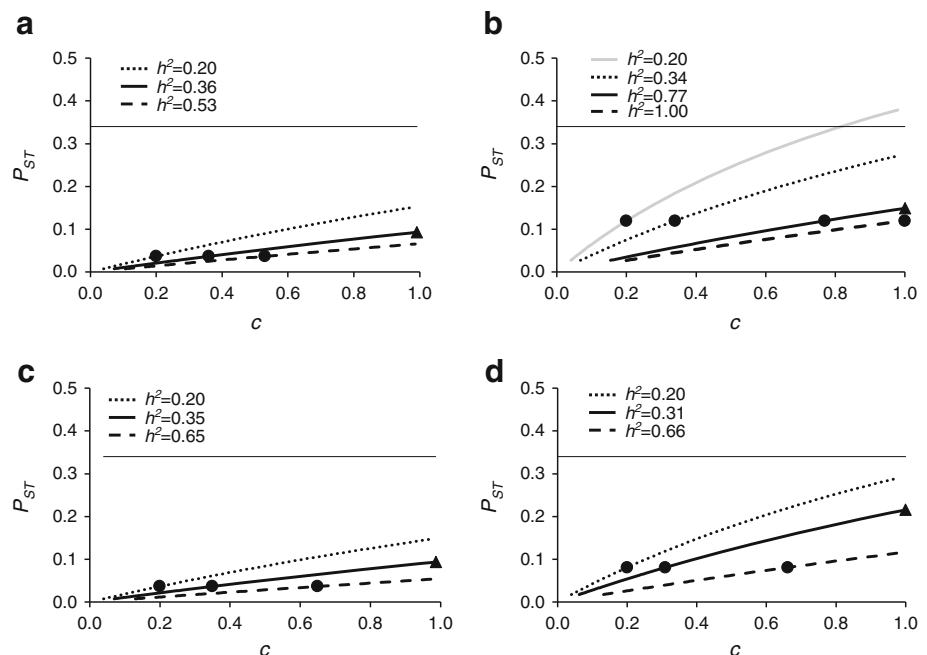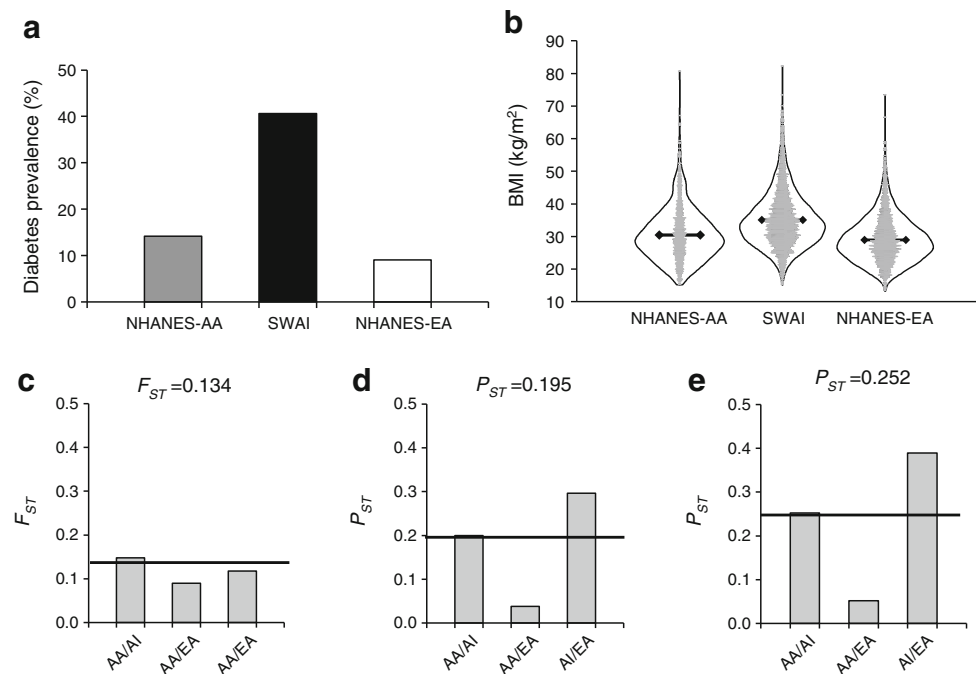
Fig. 4 (**a**) Age- and sex-adjusted prevalence of diabetes in African American (AA) and European American (EA) individuals from NHANES, and the SWAI cohort. Prevalence is adjusted to a mean age of 42.3 and a percentage of men of 54% for comparability with Fig. 1. For comparability with the NHANES data, the diabetes diagnosis for the SWAI cohort was based on data obtained at the last examination and included self-reported current or previous use of diabetes medication, as well as concurrent measures of glucose. Prevalence was 14.2% in AA, 40.6% in SWAI and 8.0% in EA ($p = 1.1 \times 10^{-214}$). (**b**) 'Bean' plot of the distribution of age- and sex-adjusted BMI in AA and EA individuals from NHANES, and the SWAI cohort. Mean BMI was 30.5 kg/m² in AA, 35.1 kg/m² in SWAI and 29.0 kg/m² in EA ($p = 2.1 \times 10^{-286}$). (**c**) Genetic divergence, measured by F_{ST} among AA and EA from the HapMap samples 1000 Genomes project and the SWAI cohort. Values for each pairwise comparison between groups are shown as bars, while the overall divergence is indicated by the horizontal line. (**d**) Divergence for diabetes prevalence, measured by P_{ST}, among AA and EA from NHANES and the SWAI cohort. (**e**) Divergence in BMI, measured by P_{ST}, among AA and EA from NHANES and the SWAI cohort. Heritability estimates for P_{ST} calculations are the same as those used for the Phoenix cohort in Table 1

NHANES, and SWAI, are shown in Fig. 4. Age- and sex-adjusted prevalence of diabetes was highest in SWAI (40.6%), and lower in African Americans (14.2%) and European Americans (8.0%, $p = 1.1 \times 10^{-214}$); similarly, mean BMI was highest in SWAI (35.1 kg/m²), and lower in African Americans (30.5 kg/m²) and European Americans (29.0 kg/m², $p = 2.1 \times 10^{-286}$). F_{ST} comparing SWAI with African Americans and European Americans from the HapMap populations was 0.134; the 95th centile of marker-wise F_{ST} was 0.354. Although P_{ST} values were modestly higher among these three populations than among the Phoenix cohort, they were well within the expected distribution of F_{ST-M}; for diabetes $P_{ST} = 0.195$ (90% CI 0.166, 0.224), while for BMI $P_{ST} = 0.252$ (0.223, 0.282) (ESM Table 8). When calculated by eq. 3, $P_{ST} = 0.157$ (0.133, 0.182) for diabetes and $P_{ST} = 0.109$ (0.094, 0.124) for BMI.

Discussion

Prevalence of type 2 diabetes and obesity differs across human continental ancestry groups, and there has been considerable speculation about the role of natural selection in these differences. The 'thrifty genotype' hypothesis posits that greater efficiency in using energy from food conferred a selective advantage in time of famine but that this predisposes to diabetes and obesity in modern environments [49, 50]. Differences in exposure and response to famine could, thus, have resulted in differences in prevalence of diabetes and obesity across ancestry groups [5–7]. Alternatively, it has been proposed that release from predation freed humans from selective pressure against obesity, and that high prevalence of obesity in humans is the result of neutral genetic drift [51]. Others have proposed that agriculture introduced a high load of carbohydrate into human diets, and that populations that adopted agriculture early, such as Europeans, have experienced greater selection for carbohydrate tolerance than other populations, resulting in protection from diabetes [8, 9]. Others have hypothesised that differences in diabetes and obesity across ancestry groups reflect adaptation to different climates or to different infectious diseases [10–13]. Such hypotheses have sometimes been discussed on the basis of the differences among ancestry groups, without consideration of whether these differences could arise neutrally.

There are few empirical genetic data supporting hypotheses that natural selection across ancestry groups contributes to risk of diabetes or obesity. Some studies have analysed established type 2 diabetes and obesity variants for molecular signatures of recent natural selection, such as extended haplotypic homozygosity. These studies have generally not found greater evidence for selection at established variants in comparison with suitably matched genomic variants [16–18], although one study did find a modest excess of evidence for selection at protective alleles for diabetes [22]. These methods are most powerful for detecting classic 'sweeps' where a previously rare allele at a single locus rapidly increases in frequency, and they are not well suited for detecting selection on polygenic traits. Others have used established variants to estimate a polygenic Q_{ST} analogue, and they found little evidence for differential selection for BMI or type 2 diabetes variants [52]. Other studies have compared allele frequencies for established variants across ancestry groups. Although these have found considerable differences in allele frequencies, the pattern of differences has not corresponded to epidemiologic risk [2, 19–21]; type 2 diabetes risk allele frequencies for established variants tend to be high in Africans but low in American Indians. However, the causal variants that contribute to diabetes and obesity are incompletely known and linkage disequilibrium patterns vary across populations, and this can introduce unpredictable biases into these comparisons [53]. The variance components methods used in the present study are designed to detect polygenic selection which differs across ancestry groups, and they do not require knowledge of specific causal variants. Our analyses show that phenotypic differences for type 2 diabetes and related traits among African Americans, American Indians and European Americans are consistent with expectations based on heritability and genetic divergence. Thus, strong diversifying selection is not necessary to explain the phenotypic differences.

Recent genetic admixture among groups could attenuate the phenotypic differences. However, we obtained similar results when analyses were restricted to those whose genetic ancestry derived ≥85% from the continent corresponding to their stated ancestry, and this suggests that our results are not unduly influenced by admixture. Phenotypic measures were made identically across the ancestry groups in our cohort. Although the Phoenix cohort is a relatively small convenience sample, we observed significant differences in diabetes and obesity risk among groups, which replicate known epidemiologic associations [1, 2]. Furthermore, P_{ST} estimates were only modestly higher when calculated among NHANES samples representative of the US African American and European American populations and an SWAI population with high prevalence of obesity and diabetes. Thus, results from these larger population samples are generally consistent with those from the Phoenix cohort.

Comparisons of phenotypic with genotypic divergence are optimally made using genetic components of variance for the trait (i.e., Q_{ST}), while we have used the total phenotypic components of variance (P_{ST}). The extent to which P_{ST} approximates Q_{ST} depends on heritability (h^2) and the proportion of the among-population phenotypic variance explained by genetic factors (c). Our estimates of h^2 are based on a relatively small sample of related individuals, but they are comparable to those reported in large meta-analyses of twin studies for BMI and diabetes [3, 4]. Although our primary analyses were conducted under the assumption that $c = 1$, environmental differences in determinants of these traits among populations would result in lower values of c, and would tend toward lower values of P_{ST}. Our sensitivity analyses suggest that our findings are consistent over a large range of values of h^2 and c. For BMI and diabetes, the largest pairwise differences in phenotypic and genotypic divergence were observed between American Indians and European Americans, and these differences suggest that a relatively modest 14% or 22% of the genetic variance in each trait, respectively, is potentially attributable to differential selection (this is under the assumption that there is no effect of neutral genetic drift, although the genotypic divergence suggests these values are well within neutral expectations).

With comparisons over only three ancestry groups representing a large portion of modern human genetic history, the present approach has limited power to detect modest degrees of diversifying selection—fairly strong or sustained selection is required for the phenotypic divergence of a single trait to exceed expectations based on genotypic divergence. We estimate by simulation that $P_{ST} \approx 0.40$ is required to detect $p < 0.05$ with 80% power (ESM Fig. 5; this corresponds to an average difference in phenotypic mean of ~0.9 SD between pairs of ancestry groups). This degree of divergence is somewhat less than observed for traits under established diversifying selection, such as skin pigmentation or craniofacial morphometry ([38, 43]; we did not directly measure skin pigmentation in the present study, but, for skin pigmentation predicted genetically [54], $P_{ST} = [Q_{ST}] = 0.831$, $p = 1.0 \times 10^{-5}$, ESM Fig. 6). The power of these variance component methods to detect diversifying selection depends on the number of ancestry groups included, as well as the number of individuals in each ancestry group [45]. While we obtained similar P_{ST} estimates among the larger population cohorts as in the Phoenix cohort with the same three ancestry groups, inclusion of additional ancestry groups may be required to detect more subtle selection. The present results do not exclude more complex models of selection, such as selection on a suite of complex traits, including some diabetes-related traits, with the overall differentiation constrained by pleiotropy, nor do they exclude modest diversifying selection on diabetes-related traits too weak to be detected by the present methods primarily affecting the American Indian group.

However, our results were obtained in major continental ancestry groups at diverse risk for diabetes and obesity, and they suggest that differences in natural selection across these groups are not necessary to explain the phenotypic differences. Investigations of the causes of differences in diabetes risk across these groups would do well to consider alternative explanations.

References

1. Cowie CC, Casagrande SS, Geiss LS (2018) Prevalence and incidence of type 2 diabetes and prediabetes. In: Cowie CC, Casagrande SS, Menke A et al (eds) Diabetes in America, 3rd edn. National Institute of Diabetes and Digestive and Kidney Diseases, Bethesda, MD, NIH Publication, 17–1468, pp 3–1 to 3–32.

2. Hanson RL, Rong R, Kobes S et al (2015) Role of established type 2 diabetes-susceptibility genetic variants in a high prevalence American Indian population. Diabetes 64:2646–2657. https://doi.org/10.2337/db14-1715

3. Willemsen G, Ward KJ, Bell CG et al (2015) The concordance and heritability of type 2 diabetes in 34,166 twin pairs from international twin registers: The discordant twin (DISCOTWIN) consortium. Twin Res Hum Genet 18:762–771. https://doi.org/10.1017/thg.2015.83

4. Elks CE, den Hoed M, Zhao JH et al (2012) Variability in the heritability of body mass index: A systematic review and meta-regression. Front Endocrinol (Lausanne) 3:29. https://doi.org/10.3389/fendo.2012.00029

5. Joffe B, Zimmet P (1998) The thrifty genotype in type 2 diabetes: An unfinished symphony moving to its finale? Endocrine 9:139–141. https://doi.org/10.1385/ENDO:9:2:139

6. Wendorf M, Goldfine ID (1991) Archaeology of NIDDM. Excavation of the "thrifty" genotype. Diabetes 40:161–165. https://doi.org/10.2337/diab.40.2.161

7. Gerstein HC, Waltman L (2006) Why don't pigs get diabetes? Explanations for variations in diabetes susceptibility in human populations living in a diabetogenic environment. CMAJ 174:25–26. https://doi.org/10.1503/cmaj.050649

8. Miller JC, Colagiuri S (1994) The carnivore connection: Dietary carbohydrate in the evolution of NIDDM. Diabetologia 37:1280–1286. https://doi.org/10.1007/BF00399803

9. Corbett SJ, McMichael AJ, Prentice AM (2009) Type 2 diabetes, cardiovascular disease, and the evolutionary paradox of the polycystic ovary syndrome: A fertility first hypothesis. Am J Hum Biol 21:587–598. https://doi.org/10.1002/ajhb.20937

10. Sellayah D, Cagampang FR, Cox RD (2014) On the evolutionary origins of obesity: a new hypothesis. Endocrinology 155:1573–1588. https://doi.org/10.1210/en.2013-2103

11. Fridlyand LE, Philipson LH (2006) Cold climate genes and the prevalence of type 2 diabetes mellitus. Med Hypotheses 67:1034–1041. https://doi.org/10.1016/j.mehy.2006.04.057

12. Dayaratne DA (2010) Impact of ecology on development of NIDDM. Med Hypotheses 74:986–988. https://doi.org/10.1016/j.mehy.2009.12.017

13. Wells JC (2009) Ethnic variability in adiposity and cardiovascular risk: the variable disease selection hypothesis. Int J Epidemiol 38:63–71. https://doi.org/10.1093/ije/dyn183

14. Mahajan A, Taliun D, Thurner M et al (2018) Fine-mapping type 2 diabetes loci to single-variant resolution using high-density imputation and islet-specific epigenome maps. Nat Genet 50:1505–1513. https://doi.org/10.1038/s41588-018-0241-6

15. Locke AE, Kahali B, Berndt SI et al (2015) Genetic studies of body mass index yield new insights for obesity biology. Nature 518:197–206. https://doi.org/10.1038/nature14177

16. Southam L, Soranzo N, Montgomery SB et al (2009) Is the thrifty genotype hypothesis supported by evidence based on confirmed type 2 diabetes- and obesity-susceptibility variants? Diabetologia 52:1846–1851. https://doi.org/10.1007/s00125-009-1419-3

17. Ayub Q, Moutsianas L, Chen Y et al (2014) Revisiting the thrifty gene hypothesis via 65 loci associated with susceptibility to type 2 diabetes. Am J Hum Genet 94:176–185. https://doi.org/10.1016/j.ajhg.2013.12.010

18. Wang G, Speakman JR (2016) Analysis of positive selection at single nucleotide polymorphisms associated with body mass index does not support the "thrifty gene" hypothesis. Cell Metab 24:531–541. https://doi.org/10.1016/j.cmet.2016.08.014

19. Klimentidis YC, Abrams M, Wang J, Fernandez JR, Allison DB (2011) 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. Hum Genet 129:407–418. https://doi.org/10.1007/s00439-010-0935-z

20. Chen R, Corona E, Sikora M et al (2012) Type 2 diabetes risk alleles demonstrate extreme directional differentiation among human populations, compared to other diseases. PLoS Genet 8:e1002621. https://doi.org/10.1371/journal.pgen.1002621

21. Corona E, Chen R, Sikora M et al (2013) Analysis of the genetic basis of disease in the context of worldwide human relationships and migration. PLoS Genet 9:e1003447. https://doi.org/10.1371/journal.pgen.1003447

22. Segurel L, Austerlitz F, Toupance B et al (2013) Positive selection of protective variants for type 2 diabetes from the Neolithic onward: a case study in Central Asia. Eur J Hum Genet 21:1146–1151. https://doi.org/10.1038/ejhg.2012.295

23. Spitze K (1993) Population structure in *Daphnia obtusa*: quantitative genetic and allozymic variation. Genetics 135: 367–374

24. Leinonen T, McCairns RJ, O'Hara RB, Merila J (2013) Q(ST)-F(ST) comparisons: evolutionary and ecological insights from genomic heterogeneity. Nat Rev Genet 14:179–190

25. Olaiya MT, Hanson RL, Kavena KG et al (2019) Use of graded Semmes Weinstein monofilament testing for ascertaining peripheral neuropathy in people with and without diabetes. Diabetes Res Clin Pract 151:1–10. https://doi.org/10.1016/j.diabres.2019.03.029

26. Sun L, Wilder K, McPeek MS (2002) Enhanced pedigree error detection. Hum Hered 54:99–110. https://doi.org/10.1159/000067666

27. American Diabetes Association (2010) Diagnosis and classification of diabetes. Diabetes Care 33(Suppl 1):S62–S69. https://doi.org/10.2337/dc10-S062

28. Dewey FE, Murray MF, Overton JD et al (2016) Distribution and clinical impact of functional variants in 50,726 whole-exome sequences from the DiscovEHR study. Science 354:aaf6814. https://doi.org/10.1126/science.aaf6814

29. Kim HI, Ye B, Gosalia N et al (2020) Characterization of exome variants and their metabolic impact in 6,716 American Indians from southwest US. Am J Hum Genet 107:251–264. https://doi.org/10.1016/j.ajhg.2020.06.009

30. Browning SR, Browning BL (2007) Rapid and accurate haplotype phasing and missing-data inference for whole-genome association studies by use of localized haplotype clustering. Am J Hum Genet 81:1084–1097. https://doi.org/10.1086/521987

31. Weir BS, Cockerham CC (1984) Estimating F-statistics for the analysis of population structure. Evolution 38:1358–1370. https://doi.org/10.1111/j.1558-5646.1984.tb05657.x

32. Weir BS, Hill WG (2002) Estimating F-statistics. Annu Rev Genet 36:721–750. https://doi.org/10.1146/annurev.genet.36.050802.093940

33. Whitlock MC (2008) Evolutionary inference from Q_{ST}. Mol Ecol 17:1885–1896. https://doi.org/10.1111/j.1365-294X.2008.03712.x

34. Bhatia G, Patterson N, Sankararaman S, Price AL (2013) Estimating and interpreting F_{ST}: The impact of rare variants. Genome Res 23:1514–1521. https://doi.org/10.1101/gr.154831.113

35. Maroti Z, Boldogkoi Z, Tombacz D, Snyder M, Kalmar T (2018) Evaluation of whole exome sequencing as an alternative to BeadChip and whole genome sequencing in human population genetic analysis. BMC Genomics 19:778

36. Brommer JE (2011) Whither P_{ST}? The approximation of Q_{ST} by P_{ST} in evolutionary and conservation biology. J Evol Biol 24:1160–1168. https://doi.org/10.1111/j.1420-9101.2011.02268.x

37. Leinonen T, Cano JM, Makinen H, Merila J (2006) Contrasting patterns of body shape and neutral genetic divergence in marine and lake populations of threespine sticklebacks. J Evol Biol 19: 1803–1812. https://doi.org/10.1111/j.1420-9101.2006.01182.x

38. Zaidi AA, Mattern BC, Claes P, McEvoy B, Hughes C, Shriver MD (2017) Investigating the case of human nose shape and climate adaptation. PLoS Genet 13:e1006616. https://doi.org/10.1371/journal.pgen.1006616

39. Almasy L, Blangero J (1998) Multipoint quantitative-trait linkage analysis in general pedigrees. Am J Hum Genet 62:1198–1211. https://doi.org/10.1086/301844

40. Neale MC, Miller MB (1997) The use of likelihood-based confidence intervals in genetic models. Behav Genet 27:113–120. https://doi.org/10.1023/a:1025681223921

41. Sokal RR, Rohlf FJ (1969) Single-classification analysis of variance. In: Emerson R, Kennedy D, Park RB (eds) Biometry: The principles and practice of statistics in biological research. W.H. Freeman and Company, San Francisco, pp 204–252

42. Hangartner S, Laurila A, Rasanen K (2012) Adaptive divergence in moor frog (*Rana arvalis*) populations along an acidification gradient: Inferences from Q_{ST}–F_{ST} correlations. Evolution 66:867–881. https://doi.org/10.1111/j.1558-5646.2011.01472.x

43. Guo J, Tan J, Yang Y et al (2014) Variation and signatures of selection on the human face. J Hum Evol 75:143–152. https://doi.org/10.1016/j.jhevol.2014.08.001

44. Kost JT, McDermott MP (2002) Combining dependent *P-value*s. Stat Prob Lett 60:183–190. https://doi.org/10.1016/S0167-7152(02)00310-3

45. Whitlock MC, Guillaume F (2009) Testing for spatially divergent selection: comparing Q_{ST} to F_{ST}. Genetics 183:1055–1063. https://doi.org/10.1534/genetics.108.099812

46. Fischer MC, Foll M, Excoffier L, Heckel G (2011) Enhanced AFLP genome scans detect local adaptation in high-altitude populations of a small rodent (*Microtus arvalis*). Mol Ecol 20:1450–1462. https://doi.org/10.1111/j.1365-294X.2011.05015.x

47. Alexander DH, Novembre J, Lange K (2009) Fast model-based estimation of ancestry in unrelated individuals. Genome Res 19: 1655–1664. https://doi.org/10.1101/gr.094052.109

48. Traurig M, Hanson RL, Marinelarena A et al (2016) Analysis of *SLC16A11* variants in 12,811 American Indians: genotype-obesity interaction for type 2 diabetes and an association with *RNASEK* expression. Diabetes 65:510–519. https://doi.org/10.2337/db15-0571

49. Neel JV (1962) Diabetes mellitus: a "thrifty" genotype rendered detrimental by "progress"? Am J Hum Genet 14:353–362

50. Neel JV, Weder AB, Julius S (1998) Type II diabetes, essential hypertension, and obesity as "syndromes of impaired genetic homeostasis": the "thrifty genotype" hypothesis enters the 21st century. Perspect Biol Med 42:44–74. https://doi.org/10.1353/pbm.1998.0060

51. Speakman JR (2008) Thrifty genes for obesity, an attractive but flawed idea, and an alternative perspective: the 'drifty gene' hypothesis. Int J Obes 32:1611–1617. https://doi.org/10.1038/ijo.2008.161

52. Berg JJ, Coop G (2014) A population genetic signal of polygenic adaptation. PLoS Genet 10:e1004412. https://doi.org/10.1371/journal.pgen.1004412

53. Martin AR, Gignoux CR, Walters RK et al (2017) Human demographic history impacts genetic risk prediction across diverse populations. Am J Hum Genet 100:635–649. https://doi.org/10.1016/j.ajhg.2017.03.004

54. Maroñas O, Phillips C, Söchtig J et al (2014) Development of a forensic skin colour predictive test. Forensic Sci Int Genet 13:34–44. https://doi.org/10.1016/j.fsigen.2014.06.017