nature genetics

Evidence of widespread selection on standing variation in Europe at height-associated SNPs

Michael C Turchin[1–5,8], Charleston WK Chiang[1–6,8], Cameron D Palmer[1–5], Sriram Sankararaman[5,6], David Reich[5,6], Genetic Investigation of ANthropometric Traits (GIANT) Consortium[7] & Joel N Hirschhorn[1–6]

Strong signatures of positive selection at newly arising genetic variants are well documented in humans[1–8], but this form of selection may not be widespread in recent human evolution[9]. Because many human traits are highly polygenic and partly determined by common, ancient genetic variation, an alternative model for rapid genetic adaptation has been proposed: weak selection acting on many pre-existing (standing) genetic variants, or polygenic adaptation[10–12]. By studying height, a classic polygenic trait, we demonstrate the first human signature of widespread selection on standing variation. We show that frequencies of alleles associated with increased height, both at known loci and genome wide, are systematically elevated in Northern Europeans compared with Southern Europeans ($P < 4.3 \times 10^{-4}$). This pattern mirrors intra-European height differences and is not confounded by ancestry or other ascertainment biases. The systematic frequency differences are consistent with the presence of widespread weak selection (selection coefficients ~10^{-3}–10^{-5} per allele) rather than genetic drift alone ($P < 10^{-15}$).

Positive selection on newly arising alleles produces a strong genetic signature: a long haplotype of unexpectedly high frequency[13]. In contrast, weak polygenic selection on standing variation acts on multiple haplotypes simultaneously[14–16]. As a result, the effects of polygenic adaptation on patterns of variation are generally modest and spread across many haplotypes at any one locus. To overcome the difficulties associated with identifying these individually weak signals, we have implemented an approach that combines evidence for selection across many loci. Specifically, we examined SNPs tested in genome-wide association studies (GWAS) to identify which allele at each SNP is a 'trait-increasing' allele, or one associated with increased trait values, and then tested these trait-increasing alleles as a group for systematic, directional differences in allele frequencies between populations. Under polygenic selection, we expect that the trait-increasing alleles will tend to be present in greater frequencies in the population with higher trait values[10,17].

We propose that adult height in Europe might provide an example of polygenic adaptation in humans. People of Northern-European descent are typically taller than people of Southern-European descent (**Supplementary Table 1**), and although nongenetic factors can produce phenotypic differences between groups[18,19], we suspected that the height differences between these closely related populations might be partially explained by genetic differences due to widespread selection on standing variation. We tested this hypothesis using recent GWAS data for height generated by the Genetic Investigation of ANthropometric Traits (GIANT) Consortium[20] and estimates of Northern- and Southern-European allele frequencies based on data sets from the Myocardial Infarction Genetics (MIGen) consortium[21] and the Population Reference Sample (POPRES)[22], expecting the height-increasing allele at height-associated loci to be more frequent in populations from Northern Europe than those from Southern Europe.

We first compared the allele frequencies in Northern and Southern Europeans of 139 variants that are known to be associated with height at genome-wide significance[20] and were directly genotyped in the MIGen study. We used 257 US individuals of Northern European ancestry and 254 Spanish individuals from the MIGen study, defined as the Northern- and Southern-European populations, respectively (**Supplementary Note** and **Supplementary Fig. 1**). We found that the height-increasing alleles were more likely to have higher frequencies in Northern than in Southern Europeans (85 out of 139, sign test $P = 0.011$; mean frequency difference = 0.012, t-test $P = 4.3 \times 10^{-4}$; **Table 1**). The difference in mean allele frequency was robust when compared to 10,000 sets of SNPs that were drawn at random from the genome and matched by average Northern- and Southern-European allele frequency to the known height-increasing SNPs on a per-SNP basis ($P = 0.0056$; **Fig. 1a**; see Online Methods). We observed similar results in an independent data set, POPRES (**Table 1**, **Supplementary Table 2** and **Supplementary Fig. 2a**). Thus, the group of height-increasing alleles at known associated variants is more common in Northern than in Southern Europe, indicating that the phenotypic difference between these two populations is at least partly due to genetic factors.

[1]Division of Genetics, Children's Hospital Boston, Boston, Massachusetts, USA. [2]Division of Endocrinology, Children's Hospital Boston, Boston, Massachusetts, USA. [3]Program in Genomics, Children's Hospital Boston, Boston, Massachusetts, USA. [4]Metabolism Initiative, Broad Institute, Cambridge, Massachusetts, USA. [5]Program in Medical and Population Genetics, Broad Institute, Cambridge, Massachusetts, USA. [6]Department of Genetics, Harvard Medical School, Boston, Massachusetts, USA. [7]A full list of members and institutions is provided in the **Supplementary Note**. [8]These authors contributed equally to this work. Correspondence should be addressed to J.N.H. (joelh@broadinstitute.org).

Received 2 March; accepted 5 July; published online 19 August 2012; doi:10.1038/ng.2368

Table 1 Comparisons of the mean allele frequency difference and the maximum likelihood estimate of selective parameter *s* in pairwise combinations of populations across Europe

Populations (data source)	Comparison	Sample size (*n*)	Mean allele frequency difference	*t*-test *P* value	*s* ($w = s\beta$)	LRT *P* value ($w = s\beta$ versus drift)	*s* ($w = s\beta$)	LRT *P* value ($w = s\beta$ versus drift)
					$T = 20$		$T = 500$	
US versus Spain (MIGen)	N versus S	257;254	0.0079	9.67×10^{-16}	0.138	9.57×10^{-16}	0.0055	9.65×10^{-16}
Sweden versus Spain (MIGen)	N versus S	58;58	0.0094	1.47×10^{-7}	0.183	5.48×10^{-8}	0.0073	5.44×10^{-8}
UK versus Italy (POPRES)	N versus S	208;208	0.016	1.06×10^{-33}	0.264	2.99×10^{-35}	0.0105	3.24×10^{-35}
UK versus Portugal (POPRES)	N versus S	125;125	0.012	1.72×10^{-20}	0.207	4.91×10^{-18}	0.0082	4.98×10^{-18}
UK versus Switzerland (French) (POPRES)	N versus C	208;208	0.0044	5.18×10^{-7}	0.076	1.52×10^{-5}	0.0030	1.52×10^{-5}
Switzerland (French) versus Italy (POPRES)	C versus S	208;208	0.011	1.73×10^{-25}	0.188	1.32×10^{-22}	0.0075	1.36×10^{-22}
Switzerland (French) versus Portugal (POPRES)	C versus S	125;125	0.0081	1.86×10^{-12}	0.139	1.07×10^{-9}	0.0055	1.08×10^{-9}

Populations are categorized as Northern (N), Central (C) or Southern (S) European. Results shown are for the set of ~1,400 independent SNPs (see main text and **Supplementary Methods** for exact numbers in each comparison), comparing the mean allele frequency difference between the more northern population and the more southern population, as well as the maximum likelihood estimate of the selection coefficients under a model in which the coefficients are proportional to the estimated effects on height ($w = s \times \beta$, where β is the estimated increase in height per allele in s.d. and w is the selective pressure per allele per generation). The *P* values shown for the mean allele frequency difference are assessed by *t*-test. The *P* values for the estimates of *s* are assessed by LRT, comparing a model of drift alone to a model of drift plus selection. Though too recent to be the realistic time frames for historical divergence (T, in generations) between the Northern- and Southern-European populations, results for $T = 20$ and $T = 500$ were included to account for the probable bidirectional migration between European populations, which would decrease the apparent time of divergence between the two populations. Note that our analysis is an estimate of the product of T and s. Because our estimates of T and s cannot be decoupled, the LRT statistics and *P* values are nearly identical across ranges of T (see **Supplementary Tables 4–10** for more detailed results across a full range of T). Accordingly, we are not estimating T but are instead estimating s under a range of values that are likely to span the actual (unknown) value of T.

We noted that the randomly matched SNPs used as a control in this analysis also showed a similar, if more subtle trend: the nominally height-increasing alleles of these SNPs (based on direction of effect in GIANT data) tended to be more common in Northern than in Southern Europeans (mean frequency difference across 10,000 matched SNP sets = 0.0035; **Fig. 1a**). In fact, throughout much of the genome, nominally height-increasing alleles are more likely to have higher frequencies in Northern than in Southern Europeans (**Fig. 1b** and **Supplementary Fig. 2b**). This observation suggested that, beyond the 180 known loci[20], many additional height-associated SNPs in the genome may reach genome-wide significance in GWAS as statistical power is improved (consistent with previous modeling)[20,23] and that the height-increasing alleles at these variants may further contribute to the difference in average height between these populations.

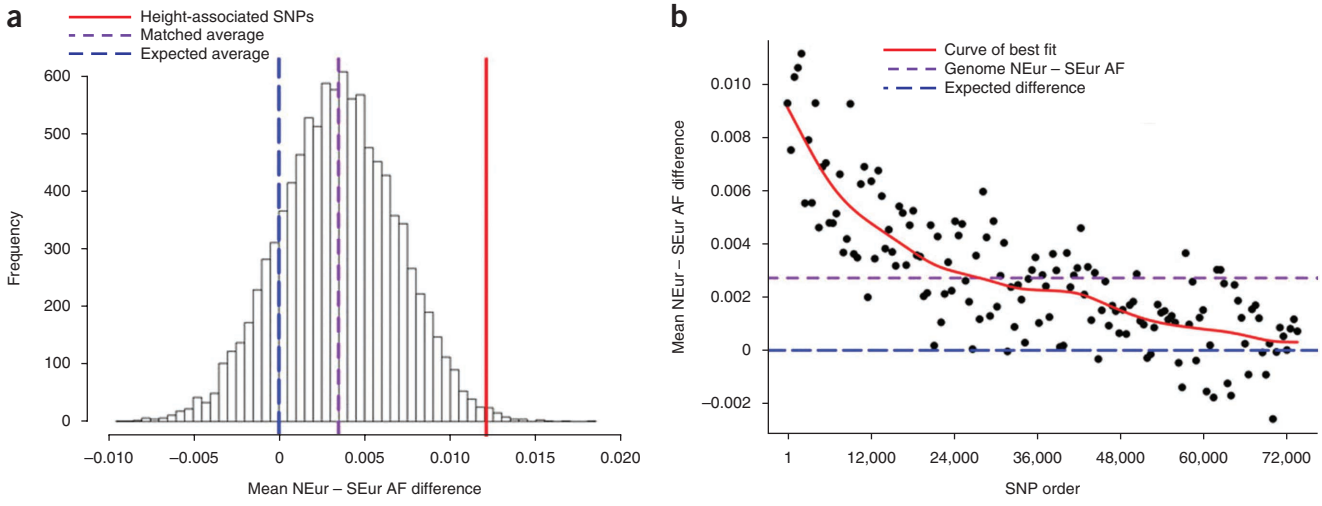

Figure 1 Mean allele frequency difference of height SNPs, matched SNPs and genome-wide SNPs between Northern- and Southern-European populations. (**a**) Mean frequency difference of the height-increasing alleles from 139 known height SNPs (solid red line) compared to that of 10,000 sets of randomly drawn SNPs matched by average Northern- and Southern-European allele frequencies to the known height SNPs on a per-SNP basis (purple dashed line). Blue dashed line, expected mean difference for matched SNPs ($x = 0$); NEur, Northern European; SEur, Southern European; AF, allele frequency. (**b**) Mean frequency difference of the height-increasing allele for sets of 500 independent ($r^2 < 0.1$) SNPs across the genome, sorted by GIANT height-association *P* value. Red line, curve of best fit; purple dashed line, genome-wide mean frequency difference; blue dashed line, expected mean difference ($y = 0$).

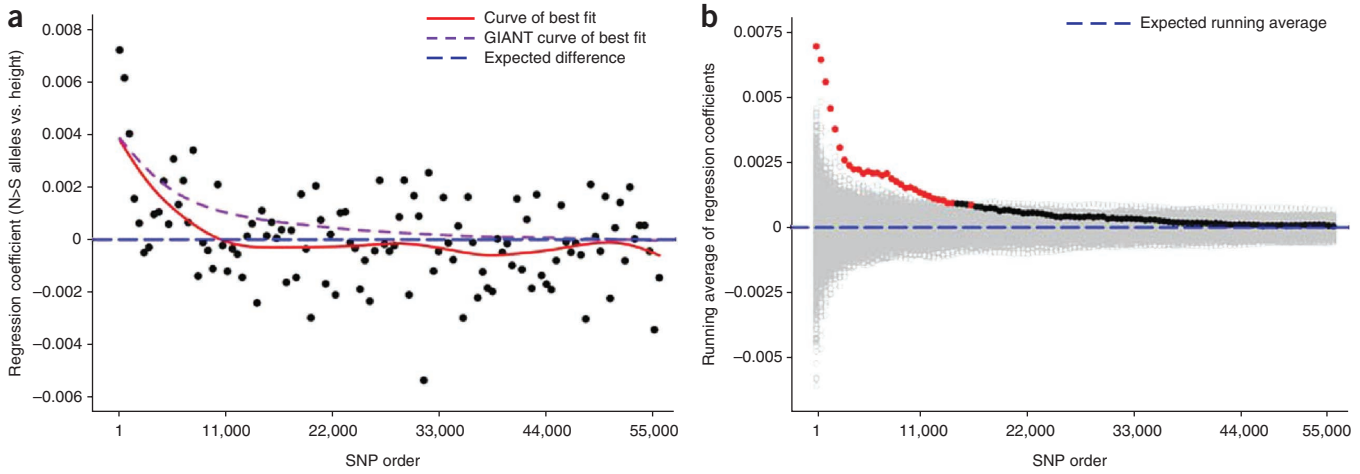

Figure 2 Within-family analyses of height and the Northern-predominant alleles across the genome. (**a**) Average regression coefficients of height versus number of Northern-predominant (N>S) alleles for sets of 500 SNPs sorted by GIANT height-association *P* value. Red line, curve of best fit; purple dashed line, curve of best fit for the GIANT effect sizes; blue dashed line, expected difference ($y = 0$). (**b**) Running averages of the regression coefficients (red and black circles; black indicates <0.01% of permuted values) sorted by GIANT height-association *P* value. Gray circles, running averages from 1,000 analyses where phenotypes were permuted within sibships; blue dashed line, expected running average ($y = 0$).

Although there seems to be a genome-wide trend for height-increasing alleles to be Northern-predominant (that is, more common in Northern than in Southern Europeans), we must also consider confounding by ancestry as a possible explanation for this observation[24–27]. The GIANT consortium took multiple steps to control for ancestry[20], but if these steps were not completely effective, then SNPs with an allele frequency difference between Northern and Southern Europeans would tend to be spuriously associated with height, with the Northern-predominant allele seeming to be the height-increasing allele.

We therefore estimated the effect sizes for the Northern-predominant alleles on height in a family-based cohort (the Framingham Heart Study (FHS)) using a sibship-based regression analysis that is immune to stratification (see Online Methods) and compared these estimates with those from GIANT. We observed that, for the most strongly associated ~1,400 SNPs, the estimated effects of the Northern-predominant alleles on height are indistinguishable between the sibship-based test and the GIANT data set (paired *t*-test $P = 0.36$; **Supplementary Fig. 3**). For the remaining SNPs, the average estimates of effect size from the family-based analysis fall toward 0 slightly faster than the GIANT estimates (**Fig. 2a** and **Supplementary Fig. 4a**). This faster decrease could be due to low power in the smaller family-based sample and/or residual stratification in the remaining GIANT data, although there is clearly a signal of true association beyond these ~1,400 SNPs (**Fig. 2** and **Supplementary Fig. 4**). To ensure that our conclusions are not confounded by stratification, we therefore focus our subsequent analyses on this set of ~1,400 independent SNPs. The allele frequency of these ~1,400 height-increasing alleles is significantly higher in Northern than in Southern Europeans, according to multiple different comparisons within both MIGen and POPRES (all *t*-test $P < 1.5 \times 10^{-7}$; **Table 1**). We also found that the frequencies in a Central-European population (Swiss-French from POPRES) fall between those of the Northern- and Southern-European POPRES populations (**Table 1**). Thus, the observation that many height-increasing alleles are more common in Northern than Southern Europeans is not explained by stratification. Rather, consistent with selection, the data indicate a small but systematic increase over time in the frequency of height-increasing alleles in Northern Europe and/or a decrease in frequency in Southern Europe.

Finally, we asked whether this systematic change in frequency of height-increasing alleles can be explained by genetic drift or is more consistent with a model that also incorporates selection (see Online Methods). In the absence of selection, the expected difference in allele frequency has a mean of 0 and a variance of $p(1-p)(2 \times F_{ST} + 1/N_1 + 1/N_2)$, where p is the estimated ancestral allele frequency, F_{ST} is estimated using the genome-wide data and N_i is the population sample size[28]. The expected effect of selection on allele frequency differences is estimated as

$$\Delta AF_{\mathrm{Sel}} \approx T \times \left(\frac{wp^2 + wp + p}{1 + 2wp} - p \right)$$

where T is the number of generations of differential selection and w is the selective pressure per allele per generation (see Online Methods). We used a likelihood ratio test (LRT) to compare models incorporating selection and drift to a model of drift alone; our simulations (**Supplementary Note**) verified that the LRT gave expected results under the null model of drift alone (**Supplementary Figs. 5** and **6**) and in models incorporating both drift and selection (**Supplementary Table 3**) and is robust to the choice of ancestral allele frequency p (data not shown).

By calculating the combined likelihood of the frequency data at the ~1,400 independent SNPs under each of the different models, we found that models incorporating both selection and drift were more consistent with the data than models of drift alone, with LRT *P* values ~10^{-16} over a range of values of T (**Table 1** and **Supplementary Tables 4–10**; see **Supplementary Tables 11** and **12** for results using a larger genome-wide set of SNPs). Given typical effect sizes of height-associated variants, which are generally ≤10^{-2}–10^{-3} s.d. (1 s.d. ≈ 6.5 cm), we estimate that, in a model where selection is proportional to effect size, the typical selective pressure on individual height-associated variants would be ~10^{-3}–10^{-5} per allele per generation. Thus, the data are much more consistent with the presence of widespread weak selection on standing variation than with a model of drift alone.

We also addressed several other factors that could confound our results. First, we considered whether demographic biases in GIANT

could have produced our results. Because GIANT consists largely of individuals of Northern-European ancestry, the consortium could have greater power to identify height-associated variants whose frequencies are closer to 0.5 in Northern Europeans. However, when we reordered the GIANT GWAS results on the basis of discovery power in Southern Europeans (**Supplementary Note**), our results were essentially unchanged (**Supplementary Table 2** and **Supplementary Figs. 7** and **8**). Second, the height-associated SNPs were limited to SNPs contained in HapMap, which itself ascertained the SNPs in part by sequencing in Northern- but not Southern-European samples. This ascertainment bias could influence the Northern- and Southern-European minor allele frequency distributions in HapMap SNPs and thus the height-associated SNPs. However, the minor allele frequency distribution of the ~1,400 height-associated SNPs is indistinguishable between Northern and Southern Europeans (Kolmagorov-Smirnov $P = 0.996$). Furthermore, we showed through simulations using an even more biased scheme of SNP ascertainment based on the 1000 Genomes Project[29] that such bias does not account for our results (**Supplementary Note**). Notably, our results show a directional shift in allele frequencies rather than an overall shift, so ascertainment biases in GIANT or HapMap would be potentially relevant only if height-increasing alleles were systematically biased toward being the major or minor allele. However, there is no statistically significant bias in either the known height-increasing alleles (70/138 major alleles in Northern Europeans; 71/139 major alleles in Southern Europeans) or the expanded set of ~1,400 SNPs (752/1,434 major alleles in Northern Europeans; 740/1,436 major alleles in Southern Europeans; all $P > 0.05$). Thus, our results cannot be explained by the ascertainment of height-associated SNPs largely in Northern Europeans.

Another important source of potential bias is our study of a phenotype (height) and pair of populations (Northern and Southern Europeans), where the phenotype was known to differ between the populations. As previously discussed by Orr[17], it may not be surprising that we observed more height-increasing alleles in the taller population, given that we selected a phenotype known to be differentiated. To determine whether height in Northern and Southern Europeans could be simply an extreme example of a neutrally evolving trait, we simulated 10,000 neutrally evolving traits having the same genetic architecture as height (**Supplementary Note**). We estimate that we would have had to ascertain height in Northern and Southern Europeans from $>10^{16}$ neutrally evolving trait-population pairs to obtain the degree of differentiation we observed in the actual data (**Supplementary Fig. 9**), indicating that our observations are not simply the extreme end of neutrally evolving traits but rather reflect the effects of selection.

In summary, we have provided an empirical example of widespread weak selection on standing variation. We observed genetic differences using multiple populations from across Europe, thereby showing that the adult height differences across populations of European descent are not due entirely to environmental differences but rather are, at least partly, genetic differences arising from selection. Height differences across populations of non-European ancestries may also be genetic in origin, but potential nongenetic factors, such as differences in timing of secular trends, mean that this inference would need to be directly tested with genetic data in additional populations. By aggregating evidence of directionally consistent intra-European frequency differences over many individual height-increasing alleles, none of which has a clear signal of selection on its own, we observed a combined signature of widespread weak selection. However, we were not able to determine whether this differential weak selection (either positive or negative) favored increased height in Northern Europe, decreased height in Southern Europe or both. One possibility is that sexual selection or assortative mating (sexual selection for partners in similar height percentiles) fueled the selective process. It is also possible that selection is not acting on height *per se* but on a phenotype closely correlated with height or a combination of phenotypes that includes height.

Our analysis is practicable because many variants have been reproducibly associated with height, and it also suggests that many more loci with small effects on height remain to be identified. As more GWAS data become available for human traits or diseases, this approach can be used to search for other examples of human polygenic adaptation, including traits or diseases associated with climate or other environmental factors that vary across otherwise closely related populations[8,30,31].

URLs. R 2.11, http://www.r-project.org/.

METHODS

Methods and any associated references are available in the online version of the paper.

Note: Supplementary information is available in the online version of the paper.

ACKNOWLEDGMENTS

The authors would like to thank E.L. Altmaier, K.E. Samocha, S.R. Grossman, G. Coop, other attendees of the Biology of Genomes 2011 conference, B.F. Voight, M. McCarthy, P. Visscher and other members of the Reich and Hirschhorn labs for their discussions and helpful comments. We gratefully thank the GIANT consortium and particularly the members of the height working group for making unpublished association data available. We thank the MIGen consortium for making allele frequency data available. This research was conducted using data and resources from the FHS of the National Heart, Lung, and Blood Institute of the US National Institutes of Health and Boston University School of Medicine based on analyses by FHS investigators participating in the SNP Health Association Resource project. This work was supported by the National Heart, Lung and Blood Institute's FHS (contract no. N01-HC-25195) and its contract with Affymetrix, Inc., for genotyping services (contract no. N02-HL-6-4278). A portion of this research used the Linux Cluster for Genetic Analysis (LinGA-II) funded by the Robert Dawson Evans Endowment of the Department of Medicine at Boston University School of Medicine and Boston Medical Center. This work was also supported by a graduate research fellowship from the National Science Foundation (to C.W.K.C.), the March of Dimes (6-FY09-507 to J.N.H.) and the National Institute of Diabetes and Digestive and Kidney Diseases (1R01DK075787 to J.N.H.).

AUTHOR CONTRIBUTIONS

M.C.T., C.W.K.C., C.D.P., S.S., D.R. and J.N.H. conceived of and designed the experiments; M.C.T. and C.D.P. performed the analyses; M.C.T., C.W.K.C. and J.N.H. interpreted the data; C.W.K.C., C.D.P., D.R. and the GIANT Consortium contributed materials; M.C.T., C.W.K.C. and J.N.H. wrote the paper with input from all coauthors.

COMPETING FINANCIAL INTERESTS

The authors declare no competing financial interests.

Published online at http://www.nature.com/doifinder/10.1038/ng.2368.
Reprints and permissions information is available online at http://www.nature.com/reprints/index.html.

1. Tishkoff, S.A. *et al.* Haplotype diversity and linkage disequilibrium at human *G6PD*: recent origin of alleles that confer malarial resistance. *Science* **293**, 455–462 (2001).
2. Hamblin, M.T. & Di Rienzo, A. Detection of the signature of natural selection in humans: evidence from the Duffy blood group locus. *Am. J. Hum. Genet.* **66**, 1669–1679 (2000).
3. Bersaglieri, T. *et al.* Genetic signatures of strong recent positive selection at the lactase gene. *Am. J. Hum. Genet.* **74**, 1111–1120 (2004).
4. The International HapMap Consortium. A second generation human haplotype map of over 3.1 million SNPs. *Nature* **449**, 851–861 (2007).
5. Sabeti, P.C. *et al.* Genome-wide detection and characterization of positive selection in human populations. *Nature* **449**, 913–918 (2007).
6. Voight, B.F., Kudaravalli, S., Wen, X. & Pritchard, J.K. A map of recent positive selection in the human genome. *PLoS Biol.* **4**, e72 (2006).

7. Williamson, S.H. *et al.* Localizing recent adaptive evolution in the human genome. *PLoS Genet.* **3**, e90 (2007).
8. Hancock, A.M. *et al.* Adaptations to climate-mediated selective pressures in humans. *PLoS Genet.* **7**, e1001375 (2011).
9. Hernandez, R.D. *et al.* Classic selective sweeps were rare in recent human evolution. *Science* **331**, 920–924 (2011).
10. Pritchard, J.K. & Di Rienzo, A. Adaptation—not by sweeps alone. *Nat. Rev. Genet.* **11**, 665–667 (2010).
11. Novembre, J. & Di Rienzo, A. Spatial patterns of variation due to natural selection in humans. *Nat. Rev. Genet.* **10**, 745–755 (2009).
12. Hermisson, J. & Pennings, P.S. Soft sweeps: molecular population genetics of adaptation from standing genetic variation. *Genetics* **169**, 2335–2352 (2005).
13. Sabeti, P.C. *et al.* Positive natural selection in the human lineage. *Science* **312**, 1614–1620 (2006).
14. Przeworski, M., Coop, G. & Wall, J.D. The signature of positive selection on standing genetic variation. *Evolution* **59**, 2312–2323 (2005).
15. Barrett, R.D. & Schluter, D. Adaptation from standing genetic variation. *Trends Ecol. Evol.* **23**, 38–44 (2008).
16. Pritchard, J.K., Pickrell, J.K. & Coop, G. The genetics of human adaptation: hard sweeps, soft sweeps, and polygenic adaptation. *Curr. Biol.* **20**, R208–R215 (2010).
17. Orr, H.A. Testing natural selection versus genetic drift in phenotypic evolution using quantitative trait locus data. *Genetics* **149**, 2099–2104 (1998).
18. Lewontin, R.C. Race and intelligence. *Bull. At. Sci.* **26**, 2–8 (1970).
19. Cavelaars, A.E. *et al.* Persistent variations in average height between countries and between socio-economic groups: an overview of 10 European countries. *Ann. Hum. Biol.* **27**, 407–421 (2000).
20. Lango Allen, H. *et al.* Hundreds of variants clustered in genomic loci and biological pathways affect human height. *Nature* **467**, 832–838 (2010).
21. Myocardial Infarction Genetics Consortium. Genome-wide association of early-onset myocardial infarction with single nucleotide polymorphisms and copy number variants. *Nat. Genet.* **41**, 334–341 (2009).
22. Nelson, M.R. *et al.* The Population Reference Sample, POPRES: a resource for population, disease, and pharmacological genetics research. *Am. J. Hum. Genet.* **83**, 347–358 (2008).
23. Yang, J. *et al.* Common SNPs explain a large proportion of the heritability for human height. *Nat. Genet.* **42**, 565–569 (2010).
24. Campbell, C.D. *et al.* Demonstrating stratification in a European American population. *Nat. Genet.* **37**, 868–872 (2005).
25. Hirschhorn, J.N. & Daly, M.J. Genome-wide association studies for common diseases and complex traits. *Nat. Rev. Genet.* **6**, 95–108 (2005).
26. Freedman, M.L. *et al.* Assessing the impact of population stratification on genetic association studies. *Nat. Genet.* **36**, 388–393 (2004).
27. Lander, E.S. & Schork, N.J. Genetic dissection of complex traits. *Science* **265**, 2037–2048 (1994).
28. Ayodo, G. *et al.* Combining evidence of natural selection with association analysis increases power to detect malaria-resistance variants. *Am. J. Hum. Genet.* **81**, 234–242 (2007).
29. The 1000 Genomes Project Consortium. A map of human genome variation from population-scale sequencing. *Nature* **467**, 1061–1073 (2010).
30. Yi, X. *et al.* Sequencing of 50 human exomes reveals adaptation to high altitude. *Science* **329**, 75–78 (2010).
31. Simonson, T.S. *et al.* Genetic evidence for high-altitude adaptation in Tibet. *Science* **329**, 72–75 (2010).

ONLINE METHODS

Study cohorts. We used a GWAS data set for height, generated from the GIANT consortium[20], as our source for per-SNP association statistics. The intra-European allele frequencies were obtained from MIGen[21] and POPRES[22]. Family-based analyses were conducted using the FHS[32]. Please see **Supplementary Note** for a detailed description of these cohorts.

Defining classes of height-associated SNPs for sign tests and mean allele frequency analyses. The height-increasing allele was defined as the allele that is associated (even if not significantly) with increased height in the GIANT data set. The GIANT data set, however, contained imputed genotypes. We were concerned that imputation using the HapMap CEU (Utah residents with ancestry from Northern and Western Europe) panel as the reference panel would bias our analyses, which focus on intra-European differences. Therefore, we examined only SNPs directly genotyped in MIGen or POPRES for our analysis. To determine whether the allele frequency of the height-increasing alleles is systematically increased or decreased in either the Northern- or Southern-European populations, we compared the Northern- and Southern-European allele frequencies for three different classes of SNPs: (i) the 180 known height-associated SNPs identified by GIANT[20]; (ii) sets of SNPs matched to the height-associated SNPs by frequency; and (iii) sets of independent SNPs genome wide. A fourth class, comprising ~1,400 independent SNPs most strongly associated with height and for which the effect-size estimates are similar between GIANT and a family-based analysis, was also defined and used for much of the later analyses presented in the manuscript (see definitions and descriptions below). Intra-European differences in allele frequencies were assessed using sign tests, to determine whether the proportion of SNPs whose height-increasing allele was more common in Northern versus Southern Europeans was significantly different from the expectation that 50% of the alleles would be more common in Northern Europeans, and paired *t*-tests, to determine whether the mean Northern- to Southern-European allele frequency differences were significantly different from zero. The analyses were performed using R 2.11.

For the 180 known height-associated SNPs, the allele frequencies for only 139 and 109 SNPs were analyzed in MIGen and POPRES, respectively, because we were restricted to using only directly genotyped SNPs. These groups of SNPs include 55 and 30 height-associated SNPs that were directly genotyped, and 84 and 79 proxies that were in high linkage disequilibrium (LD; $r^2 \geq 0.8$ in CEU) with an original height-associated SNP, in MIGen and POPRES, respectively. In cases where multiple proxy SNPs were available in CEU, we selected the SNP with the lowest *P* value for height association in GIANT. Our analysis showed similar patterns in the directly genotyped SNPs and proxies, and mean allele frequency differences remained significant for both subsets of SNPs (**Supplementary Table 13**).

For the sets of matched SNPs, randomly drawn SNPs were matched to the height-associated SNPs on ancestral European allele frequency (estimated as the average allele frequency of Northern- and Southern-European populations). The genome-wide data used have been pruned by clumping SNPs in high LD ($r^2 \geq 0.8$) into a single cluster to avoid drawing highly correlated SNPs. Clumping was done by first randomly choosing a SNP as the index SNP, then clustering all SNPs within 0.5 Mb of the index SNP that had a pairwise $r^2 \geq 0.8$ calculated on the basis of HapMap phase 2 CEU data. In total, 10,000 sets of matched SNPs were generated.

For the set of independent genome-wide SNPs, we calculated the mean Northern- to Southern-European allele frequency differences of the predicted height-increasing alleles in successive groups of 500 independent variants, sorted by their GIANT height-association *P* value starting from the most strongly associated SNP. Here, SNPs were clumped using the method described above but with an r^2 threshold of ≥0.1 to ensure that clumps of SNPs were nearly or completely independent from each other. In total, 73,657 SNPs and 54,542 SNPs genome wide were used from the MIGen and POPRES data sets, respectively, to estimate Northern- and Southern-European allele frequency. Curves of best fit were determined using a smoothing spline with spar parameter of 0.75 in R 2.11.

Within-sibship association test of Northern-predominant alleles and increased height. For each SNP, the allele that is more common in Northern Europeans than in Southern Europeans is defined as the Northern-predominant allele. To determine whether Northern-predominant alleles are associated with increased height in a family-based test that is immune to stratification, we conducted a within-sibship test using data from the family-based FHS. The numbers of SNPs genotyped in the FHS and used in these analyses (after clumping to remove correlated SNPs) were 55,927 and 52,680 for the MIGen and POPRES allele frequency data sets, respectively. For each individual within a sibship and for each independent SNP ($r^2 < 0.1$), we designated the genotype as the number of Northern-predominant alleles carried by that individual. Missing genotypes were skipped and treated as neither Northern- nor Southern-predominant alleles. We then adjusted the genotype at each SNP within each sibship by subtracting from the observed number of Northern-predominant alleles the average number of Northern-predominant alleles for that SNP in that sibship. Similarly, we adjusted the age- and sex-corrected height values within each sibship by subtracting the sibship mean. Then, across all individuals (each adjusted by the means in his/her own sibship), we regressed the sibship-adjusted height values against the sibship-adjusted genotypes, producing a pure family-based test immune to stratification. The family-based effect-size estimates (that is, the regression coefficients) were compared with the effect sizes estimated by the GIANT consortium. We note that the FHS was one of the cohorts included in the GIANT meta-analysis. We therefore removed the FHS results from the GIANT data and repeated the GIANT meta-analysis to generate new GIANT estimates that are completely independent of our family-based test.

From this comparison, we identified the ~1,400 most strongly associated and clearly independent SNPs for which the effect sizes are similar in GIANT and in our family-based test. The latter SNP set was determined by first clumping the above-mentioned genome-wide data sets according to the GIANT height-association *P* value using an $r^2 \geq 0.1$. The top 5,000 SNPs from this list were then further pruned by elimination of any SNP pairs occupying the same 1-Mb window, preferentially keeping SNPs more strongly associated with height. This yielded 1,437 SNPs in the MIGen data set and 1,429 in the POPRES data set. These SNPs have comparable effect sizes between our FHS within-sibship regression coefficients and GIANT effect sizes ($P = 0.36$ and 0.89 for MIGen and POPRES, respectively, by paired *t*-test; **Supplementary Fig. 3**). Thus, the height effect-size estimates for this set of ~1,400 SNPs are not inflated by stratification. In subsequent analyses, we used these sets of ~1,400 SNPs and the genome-wide data from which the ~1,400 SNPs were selected.

For within-sibship analyses using genome-wide sets of SNPs, running averages of regression coefficients were also determined by successively calculating regression coefficients for each group of 500 SNPs and then calculating a running average of all regression coefficients up to and including that group of SNPs. To determine the significance of these running averages, we ran simulations in which height values within each sibship were randomly redistributed 1,000 times and calculated the regression coefficients and running averages for each simulation. The observed running average regression coefficients were considered significant if none of the simulations had as large a running average regression coefficient at that point in the genome as the observed values.

Modeling genetic drift and selection. To calculate the relative likelihoods that the observed Northern- and Southern-European allele frequency data for height-increasing alleles are more consistent with a model based on genetic drift alone or models that incorporate selection, we used an LRT and modeled drift according to the methods previously outlined[28].

To model the effects of drift alone, the allele frequency difference between two populations was estimated as a random normal variable with mean = 0 and variance = $p(1 - p)(c + 1/N_1 + 1/N_2)$, where p is the ancestral allele frequency (the average of the two populations), c is a genetic drift parameter equal to $2 \times F_{ST}$ (F_{ST} is determined using the genome-wide data; $F_{ST} = 0.0019$ for MIGen and 0.0031 for POPRES), and N_1 and N_2 are total chromosome counts for each of our two populations. c was estimated using the strictly clumped data sets described above. For each SNP, the negative log likelihoods of observing the Northern- and Southern-European allele frequency difference was calculated using R and summed over all independent SNPs ($r^2 < 0.1$) genome wide or in groups of 500 independent SNPs sorted by GIANT height-association *P* value.

To model the effect of drift and selection on the observed difference between Northern- and Southern-European allele frequencies, we first

doi:10.1038/ng.2368

estimated the expected number of allele frequency differences that could be attributed to selection using the following equation (see **Supplementary Note** for derivation):

$$\Delta AF_{\mathrm{Sel}} \approx T \times \left(\frac{wp^2 + wp + p}{1 + 2wp} - p \right)$$

where p is the ancestral allele frequency (estimated as the average of Northern- and Southern-European allele frequencies), T is the number of generations since the two populations have split and w is the selective pressure experienced by the population under different models of ongoing selection. Note that the above equation for changes in allele frequency over time is only an approximation, as the change in allele frequency per generation is also a function of the allele frequency itself. However, we showed that the effect is negligible when the changes in allele frequency are very small between generations, as is the case here (**Supplementary Fig. 10**). See **Supplementary Note** for further details.

Ruling out potential ascertainment biases. A number of additional biases could have influenced our results, including ascertainment bias due to GIANT cohort collection, HapMap SNP ascertainment and our choice of phenotype. See **Supplementary Note** for detailed descriptions of analyses showing that these potential ascertainment biases did not influence our results.

32. Splansky, G.L. *et al*. The third generation cohort of the National Heart, Lung, and Blood Institute's Framingham Heart Study: design, recruitment, and initial examination. *Am. J. Epidemiol.* **165**, 1328–1335 (2007).