Human Molecular Genetics, 2009, Vol. 18, No. 2 ***373–380***
doi:10.1093/hmg/ddn350
Advance Access published on October 24, 2008

Common variants in the *JAZF1* gene associated with height identified by linkage and genome-wide association analysis

Åsa Johansson[1], Fabio Marroni[2], Caroline Hayward[3], Christopher S. Franklin[4], Anatoly V. Kirichenko[5], Inger Jonasson[1], Andrew A. Hicks[2], Veronique Vitart[3], Aaron Isaacs[6], Tatiana Axenovich[5], Susan Campbell[3], Malcolm G. Dunlop[8], Jamie Floyd[9], Nick Hastie[3], Albert Hofman[6], Sara Knott[10], Ivana Kolcic[11], Irene Pichler[2], Ozren Polasek[11], Fernando Rivadeneira[7], Albert Tenesa[8], André G. Uitterlinden[7], Sarah H. Wild[4], Irina V. Zorkoltseva[5], Thomas Meitinger[12], James F. Wilson[4], Igor Rudan[4,13,14], Harry Campbell[4], Cristian Pattaro[2], Peter Pramstaller[2,15], Ben A. Oostra[6], Alan F. Wright[3], Cornelia M. van Duijn[6], Yurii S. Aulchenko[6] and Ulf Gyllensten[1,*] on behalf of the EUROSPAN Consortium

[1]Rudbeck Laboratory, Department of Genetics and Pathology, Uppsala University, SE-751 85 Uppsala, Sweden, [2]Institute of Genetic Medicine, European Academy, Viale Druso 1, I-39100 Bolzano, Italy, [3]MRC Human Genetics Unit, Institute of Genetics and Molecular Medicine, Western General Hospital, Edinburgh EH4 2XU, UK, [4]Public Health Sciences, University of Edinburgh, Teviot Place, Edinburgh EH8 9AG, UK, [5]Siberian Division of Russian Academy of Sciences, Institute of Cytology and Genetics, 630090 Novosibirsk, Russia, [6]Genetic Epidemiology Unit, Department of Epidemiology and Clinical Genetics and, [7]Genetic Laboratory, Department of Internal Medicine, Erasmus University Medical Center, PO Box 2040, 3000 CA, Rotterdam, The Netherlands, [8]Colon Cancer Genetics Group, Institute of Genetics and Molecular Medicine, MRC Human Genetics Unit, University of Edinburgh, Edinburgh EH4 2XU, UK, [9]The Roslin Institute, Royal (Dick) School of Veterinary Studies, Roslin BioCentre, University of Edinburgh, Midlothian EH25 9PS, UK, [10]Institute of Evolutionary Biology, University of Edinburgh, West Mains Road, Edinburgh EH9 3JT, UK, [11]Andrija Stampar School of Public Health, University of Zagreb Medical School, Rockefellerova 4, 10000 Zagreb, Croatia, [12]Institute of Human Genetics, Helmholtz Zentrum München, German Research Center for Environmental Health (GmbH), Ingolstaedter Landstr. 1, D-85764 Neuherberg, Germany, [13]Croatian Centre for Global Health, University of Split Medical School, Soltanska 2, 21000 Split, Croatia, [14]Institute for Clinical Medical Research, University Hospital 'Sestre Milosrdnice', Zagreb, Croatia and [15]Department of Neurology, Central Hospital, Bolzano, Italy

Received August 27, 2008; Revised and Accepted October 16, 2008

Genes for height have gained interest for decades, but only recently have candidate genes started to be identified. We have performed linkage analysis and genome-wide association for height in approximately 4000 individuals from five European populations. A total of five chromosomal regions showed suggestive linkage and in one of these regions, two SNPs (rs849140 and rs1635852) were associated with height (nominal $P = 7.0 \times 10^{-8}$ and $P = 9.6 \times 10^{-7}$, respectively). In total, five SNPs across the genome showed an association with height that reached the threshold of genome-wide significance (nominal $P < 1.6 \times 10^{-7}$). The association with height was replicated for two SNPs (rs1635852 and rs849140) using three independent studies (n = 31 077, n=1268 and n = 5746) with overall meta P-values of 9.4×10^{-10} and 5.3×10^{-8}. These SNPs are located in the *JAZF1* gene, which has recently been associated with type II diabetes, prostate and endometrial cancer. *JAZF1* is a transcriptional repressor of *NR2C2*, which results in low IGF1 serum concentrations, perinatal and

*To whom correspondence should be addressed. Tel: +46 184714909; Fax: +46 184714931; Email: ulf.gyllensten@genpat.uu.se

early postnatal hypoglycemia and growth retardation when knocked out in mice. Both the linkage and association analyses independently identified the *JAZF1* region affecting human height. We have demonstrated, through replication in additional independent populations, the consistency of the effect of the *JAZF1* SNPs on height. Since this gene also has a key function in the metabolism of growth, *JAZF1* represents one of the strongest candidates influencing human height identified so far.

INTRODUCTION

The genetic architecture of anthropometric measures is largely unknown. Height displays a very high heritability, indicating a strong contribution of genetic factors to its normal variation, but progress in identifying these factors has been relatively slow. During the last year, a number of genes or chromosomal regions have been identified that are associated with height (1–5). The large sample sizes needed and the low overlap between studies in terms of the regions identified indicate a complex inheritance pattern with many contributing gene variants of small effect, which explain a relatively low fraction of the variance in body height.

One approach to increase the power of linkage and association analyses is to focus on populations with an increased level of linkage disequilibrium (LD) and reduced diversity. LD mapping has been suggested to be more efficient in populations with increased LD, and a smaller sample size will be required using this approach to achieve the same power as in general population studies (6). Recent findings have demonstrated that studies of isolated populations also benefit from a reduction in environmental heterogeneity (7), which, in combination with lower genetic variability, may further increase the power to detect genomic regions affecting a complex trait compared with that of similar studies undertaken in general populations.

We have performed linkage and genome-wide association (GWA) analyses for height in four (linkage) or five (association) geographically and environmentally distinct European populations (Fig. 1). Data for linkage studies were available from populations of northern Swedish villages (Sweden), Alpine settlements in South Tyrol (Italy), the Vis island in the Adriatic Sea (Croatia) and a Dutch village (Netherlands), from now on referred to as the Swedish, South Tyrolean, Croatian and Dutch populations. GWA data were available for these four populations and a fifth additional population from the Orkney Islands (Scotland), subsequently referred to as the Scottish population.

RESULTS

Trait distribution and heritability

The distribution of height was similar in the different populations (Table 1). The highest average height (176 and 161.4 cm for males and females, respectively) was seen in the Croatian population and the lowest average height (173.1 and 160.6 cm for males and females, respectively) in the South Tyrolean population. The heritability was high and highly significant in all populations analyzed separately or jointly, and when stratified by gender (Table 2). The heritability, estimated in the linkage analysis, varied from 0.79 for the South Tyrolean population to 1 for the Swedish population and was higher when estimated for males and females separately. In contrast, the heritability based on the joint analysis of the five populations was similar to the average heritability estimated for the populations separately.

Linkage analysis

Linkage data were available from four of the populations. A total of five chromosomal regions reached the threshold of suggestive linkage (Table 3). For three of these regions, the highest LOD score was obtained in the joint analysis either for males and females together or separately, even though one population was the major contributor to the LOD score (Fig. 2). In the remaining two linkage regions (Table 3), the overall LOD score was mainly due to strong effects in one of the populations.

Association mapping in linkage regions

We next investigated the association between SNPs and the trait in the linkage regions. The five linkage regions were large and included between 887 and 3176 SNPs each, with a total of 9852 examined. Only in the region on chromosome 7 did any of the SNPs show evidence of association with height in the meta-analysis (Table 3). In this region, two neighboring SNPs (rs849140 and rs1635852) were associated with height ($P < 0.05$ after Bonferroni correction for 9852 SNPs tested). Both these SNPs are located within the *JAZF1* gene.

Genome-wide association

The meta-analysis of the GWA data from the five populations identified five SNPs that reached the threshold of genome-wide significance ($P < 0.05$) after Bonferroni correction for the number of SNPs tested: rs12106790 (nominal $P = 3.8 \times 10^{-8}$), rs11205415 (nominal $P = 6.7 \times 10^{-8}$), rs849140 (nominal $P = 7.0 \times 10^{-8}$), rs1772810 (nominal $P = 1.0 \times 10^{-7}$) and rs17051743 (nominal $P = 1.6 \times 10^{-7}$). None of these SNPs was associated (nominal $P < 0.05$) in all five populations, and the strongest signal was in four of the cases obtained for males separately and in one case for females (Supplementary Material, Table S2). One SNP was associated in four of the populations, two SNPs were associated in three populations and two SNPs were associated in two populations. All five SNPs associated with height were located within or in proximity (<40 kb upstream) to a known gene, and one of the SNPs (rs849140) was located within the *JAZF1*, and

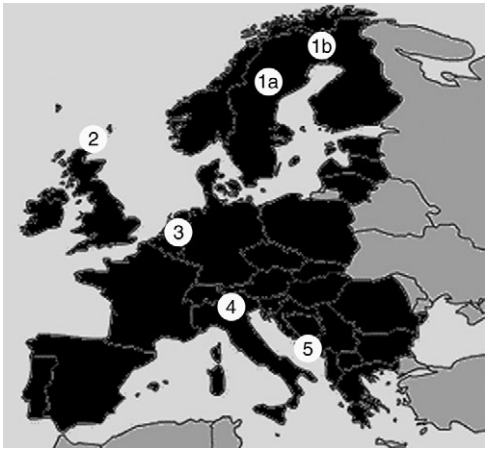

Figure 1. Sampling locations for the populations in the study. 1a and 1b is the Swedish, 2 the Scottish, 3 the Dutch, 4 the South Tyrolean and 5 the Croatian sampling locations.

corresponds to one of the SNPs under the linkage peak describe earlier.

Replication of our top SNP association findings

We used three different studies to attempt to replicate our findings on height. The recently published data from Gudbjartsson *et al*. (1) include information on all SNPs associated with height ($P < 10^{-4}$) within a large data set of 25 174 Icelanders, 2876 Dutch, 1770 European Americans, 3025 individuals of European descent from the Diabetes Genetics Initiative (DGI) and 1148 African Americans. Only one of the SNPs (rs1635852) associated with height in our data set was present among the top SNPs ($P < 10^{-4}$) identified in the data set by Gudbjartsson *et al*., where it had a *P*-value of 7.7×10^{-5} in the European populations alone ($n = 31\,077$) or 8.2×10^{-5} when the African American populations were also included ($n = 32\,225$). This SNP is located 5709 bp downstream of the SNP showing the strongest association with height in our data set (rs849140) and was detected as a candidate SNP within one of our linkage regions (Fig. 3). Both rs1635852 and rs849140 are located in the last intron of the *JAZ1F* gene and are in strong LD ($r^2 = 0.61$, 0.60, 0.75, 0.59 and 0.47 in the Croatian, Dutch, Scottish, South Tyrolean and Swedish populations, respectively).

We further investigated our six top-ranked SNP–height associations in 1268 individuals (494 cases with colorectal cancer and 774 healthy controls) of Scottish descent and recruited from the general Scottish population for the COGS (COlorectal Cancer Genetics Susceptibility) Study (8,9). One (rs849140) of the SNPs within *JAZ1F* was associated (nominal $P = 0.03$) with height in this study. Finally, we used data from the Rotterdam Study ($n = 5746$), a population-based follow-up Dutch cohort (10). In this cohort, both rs1635852 and rs849140 were associated with height ($P = 0.003$ and $P = 0.001$, respectively). In all the 11 populations of European descent [five from European Special Population Network (EUROSPAN), four from Gudbjartsson *et al*., one from the Rotterdam Study and one from the COGS], for which the information on effect sizes was available, the effect was in the same direction (Table 4), and the overall meta *P*-values including all available populations were estimated to be 9.4×10^{-10} and 5.3×10^{-8} for rs1635852 and rs849140, respectively.

Table 1. Average (standard deviation) for age (years) and height (cm) in the five populations by sex

Population	Gender	Age at study average (standard deviation)	Height average (standard deviation)
Croatian	Males ($n = 305$)	54.9 (15.9)	176 (7.2)
	Females ($n = 415$)	56.5 (16.5)	161.4 (6.6)
Dutch	Males ($n = 354$)	50.1 (14.5)	173.5 (7.1)
	Females ($n = 564$)	48.8 (14.6)	161.3 (6.3)
Scottish	Males ($n = 334$)	54.2 (15.7)	174.8 (6.7)
	Females ($n = 385$)	53.0 (15.7)	161.0 (6.1)
Swedish	Males ($n = 309$)	43.2 (10.9)	174.7 (8.1)
	Females ($n = 347$)	45.2 (10.8)	160.8 (6.8)
South Tyrolean	Males ($n = 475$)	45.6 (15.8)	173.1 (7.4)
	Females ($n = 622$)	46.1 (16.5)	160.6 (6.9)

Table 2. Heritability of height

	Height Heritability (standard error)	*P*-value
Joint both sexes	0.89 (0.027)	3.3×10^{-151}
Joint males	1.00 (NA)	1.8×10^{-45}
Joint females	0.95 (0.051)	1.6×10^{-57}
Croatian	0.93 (0.076)	7.7×10^{-21}
Dutch	0.89 (0.038)	4.6×10^{-76}
Swedish	1.00 (NA)	1.2×10^{-27}
South Tyrolean	0.79 (0.063)	5.7×10^{-33}

Only the populations from the linkage analysis are included—Croatian, Dutch, Swedish and South Tyrolean.

Replication of recent findings in our data set

We also tested the findings of three recently published meta-analyses (1,2,4) in our data set. A total of 61 SNPs were significantly associated with height in the three studies, 33 of which could be evaluated in our study (Supplementary Material, Table S3). Eighteen of these 33 SNPs (located in 27 different regions) showed an association ($P < 0.05$) with height in our sample. Two of these SNPs were strongly associated ($P < 0.05$ after correcting for 33 SNPs tested) with height in our sample; rs6854783 ($P = 4.26 \times 10^{-5}$) located in the *HHIP* region and rs8756 ($P = 4.4 \times 10^{-4}$) located in the *HMGA2* gene.

DISCUSSION

We have performed joint linkage and association under the linkage peaks, along with GWA analyses in a set of diverse European populations. These populations were selected because prior genealogical information and/or previous

Table 3. Chromosomal regions with evidence of linkage to height and the best association for each region

Trait	Chromosome	deCODE (cM)	Position[a] (Mb)	No SNPs	LOD[b]	Group	Best association group	*P*-value	P_c^c
Height	2	77–94	51–72	2582	3.7	Joint	Males	4.9×10^{-5}	NS
Height	7	22–40	8–29	3176	3.5	Sum[d]	Males[e]	7.0×10^{-8}	6.9×10^{-4}
Height	9	109–123	108–120	2014	3.2	Swedish	Males	1.5×10^{-5}	NS
Height	16	93–104	72–80	1193	2.9	Dutch	Males	7.3×10^{-6}	NS
Height	17	91–95	52–63	887	3.2	Joint	Males	6.7×10^{-5}	NS

[a]Build 36.2.
[b]Suggestive LOD of 2.69 is based on seven subpopulations: Swedish, South Tyrolean, Croatian, Dutch, Females, Males and All populations jointly. Suggestive LOD of >3.47 based on combining LOD scores for the four subpopulations (Swedish, Croatian, South Tyrolean, Dutch).
[c]*P*-value corrected for the number of SNPs in the region and the total number of regions evaluated for the trait. NS, not significant ($P > 0.05$).
[d]Sum of the LOD scores for the four subpopulations (Swedish, Croatian, South Tyrolean, Dutch).
[e]The best association is to rs849140, which is located in the *JAZF1* gene.

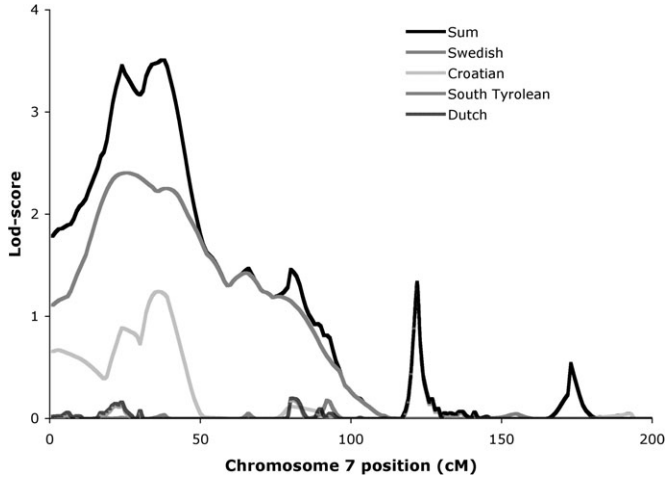

Figure 2. QTL on chromosome 7 for height. The signal in the summarized LOD score is mainly based on the signal from the Swedish population, with contribution also from the Croatian populations.

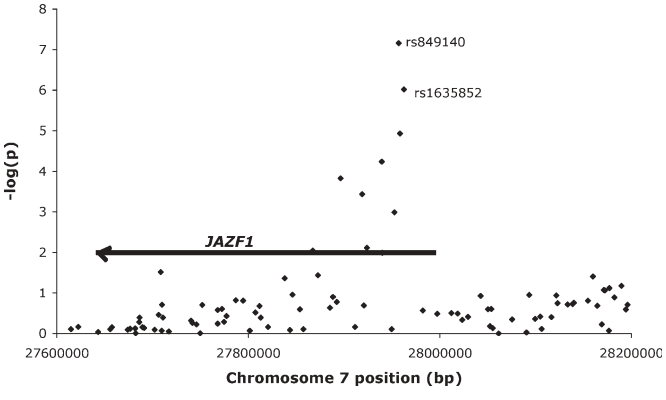

Figure 3. Manhattan plot of the *JAZF1* region on chromosome 7. Associations ($-\log_{10}$ nominal *P*-value) for males are showed for all SNPs in the *JAZF1* region. The top SNPs (rs1635852 from the linkage region) and rs849140 (from the linkage region and the genome-wide analysis) are all located within the first intron of the *JAZF1* gene.

molecular studies indicated that they have increased LD and decreased genetic diversity compared with general populations. The linkage analysis results showed little overlap between the four populations. Most of the regions with suggestive linkage were identified in the joint analysis, but in the majority of cases, the dominant contribution to the LOD score came from a single population. This is expected, since linkage is most powerful to detect rare alleles which have large effects on the trait and which are more likely to vary across populations. Of the five linkage regions, only the one on chromosome 7 also showed evidence for association. This low overlap might reflect the different power of linkage and association methods in detecting rare alleles with large effect versus common variants with small effects and implies that the specific variants underlying the linkage signal might not be those responsible for the association findings. However, we cannot exclude the possibility that some of the regions identified in the linkage analysis represent false positives. Combining linkage and association analysis allowed us to identify one additional SNP influencing height compared with performing association analysis alone, but the fact that the chromosome 7 region was identified both by linkage and GWA is interesting and might also reflect multiple polymorphisms with different frequencies. The overlap between the genome-wide findings in the five populations was also modest. The statistically most significant association in each of the populations showed very little overlap across populations. Similarly, about half of SNPs recently reported to be associated with height (Supplementary Material, Table S3) showed an association ($P < 0.05$) in our data set (1,2,4). This is higher than expected by chance, but the significant association with height remained only for two of these SNPs after correcting for the number of SNPs tested (33 SNPs).

In conclusion, we have identified novel common variants associated with normal variation in height in a study of 3925 individuals from five diverse European populations. This is one of the first studies of common variants for human height performed on family-based population samples across Europe. In total, we identified two SNPs associated to variation in height, which also replicates in up to six different populations. The SNP (rs849140) showing the strongest association with height replicated in 1268 unrelated individuals of Scottish ancestry and ($n = 5746$) of Dutch descent. A nearby SNP

Table 4. Sex- and age-adjusted effect sizes (standard error) in centimeters for the two SNPs in the *JAZF1* gene associated to height

	EUROSPAN[a]					Rotterdam Study (10)	COGS (8,9)	Gudbjartsson *et al.* (1)[b]				POOLED	
SNP (effective allele)	Croatian	Dutch	Scottish	South Tyrolean	Swedish	Dutch	Scottish	Iceland	Holland	US Caucasian	DGI	β (SE β)	*P*-value
rs849140 (A)	0.309 (0.201)	0.194 (0.150)	0.523 (0.161)	0.234 (0.161)	0.228 (0.223)	0.394 (0.118)	0.571 (0.267)	NA	NA	NA	NA	0.342 (0.063)	5.3×10^{-8}
rs1635852 (A)	0.344 (0.175)	0.251 (0.170)	0.361 (0.166)	0.179 (0.124)	0.281 (0.215)	0.344 (0.116)	0.390 (0.260)	0.203 (0.0655)	0.385 (0.226)	0.210 (0.191)	0.308 (0.193)	0.248 (0.041)	9.4×10^{-10}

NA, data not available.
[a]All effect sizes are for males and females together.
[b]Effect sizes and standard errors were converted from percentages of standard deviations and *z*-scores into cm by setting the standard deviation to 7 cm.

(rs1635852) was also associated with height in a recently published data set (four populations of European descent) by Gudbjartsson *et al*. (1) from approximately 31 000 individuals. This SNP also showed an association was also present in the 5746 individuals of Dutch descent. Interestingly, the effect sizes were quite similar across populations ranging from 0.18 cm for rs1635852 in the South Tyrolean population to 0.57 cm for rs849140 in the Scottish population from the COGS study. The two SNPs, rs1635852 and rs849140, are both located in the chromosome 7 linkage region (Fig. 3) and are in LD with each other. Both SNPs are located in the first intron of the *JAZF1* gene. *JAZF1* has recently been associated with type 2 diabetes (11), prostate cancer (12) and endometrial cancer (13). *JAZF1*, also known as *TIP27*, encodes a transcriptional repressor of *NR2C2*, also known as *TAK1* or *TR4* (14). Mice deficient in *NR2C2* show low IGF1 serum concentrations and perinatal and early postnatal hypoglycemia, as well as growth retardation (15). As a transcriptional repressor of a gene that causes growth retardation in mice, *JAZF1* is a good candidate also to affect variation in human height. *JAZF1* has a key function in the metabolism of growth, and our results show that it has a consistent pattern of association to variation in height across 11 populations. This makes *JAZF1* one of the best candidate genes, identified so far, influencing the variation in human height.

MATERIALS AND METHODS

Subjects

The five populations used for the initial analysis are all part of the EUROSPAN (European Special Population Network) project, the aims of which are to study the genetic and environmental determinants of quantitative phenotypes of clinical importance. The Swedish samples are part of the Northern Swedish Population Health Study (NSPHS). Samples for the linkage analysis were collected from the southern part of the Swedish mountain region (County of Västerbotten) and samples for the association analysis from the northern part of the Swedish mountain region (County of Norrbotten, Parish of Karesuando) (Fig. 1). There is no overlap between the individuals in the Swedish linkage and association cohorts. Samples from South Tyrol were collected as part of an extended genetic study (MICROS) from settlements in Venosta Valley. The Dutch samples were collected within the ERF study, which is a family-based study that includes 3000 inhabitants of a genetically isolated community in the south-western area of the Netherlands (16). The Croatian samples were from the villages of Vis and Komiza on the Dalmatian island of Vis. The Orkney Complex Disease Study (ORCADES) is an ongoing family-based genetic epidemiology collection in the isolated Scottish archipelago of Orkney. More information about each EUROSPAN study population can be found in Supplementary Material. Two additional populations were used to replicate the initial findings. The COGS study in Scotland is a national population-based case–control study of colorectal cancer occurring in individuals aged 55 years and under presenting to Scottish hospitals (8,9). From the COGS study, 1268 individuals were genotyped for the SNPs of interest. The Rotterdam

Study (10) is a prospective cohort study that started in 1990 in Ommoord, a suburb of Rotterdam, among 10 994 men and women aged 55 years and over. Height measurements used in this study were obtained at baseline between 1990 and 1993. For this study, we used data on 5746 participants for whom genotypic and height data were available (10,17).

Genotyping

Linkage analysis. A total of 3448 individuals (1890 females and 1518 males) from four populations were genotyped (591 Croatian, 1455 Dutch, 926 South Tyrolean and 436 Swedish). The Swedish, South Tyrolean and Croatian samples were typed for 390, 1113, 747 microsatellite markers, respectively, whereas the Dutch individuals were genotyped for 6008 SNPs. For complete genotyping information, see references (18–20). The individuals from the four populations represented 138, 156, 121 and 119 families, with a total pedigree size of 2806, 1346, 1940 and 850 individuals for the Croatian, Dutch, South Tyrolean and Swedish population, respectively.

Genome-wide association. DNA samples ($n = 4200$) were genotyped according to the manufacturer's instructions on using Illumina's HumanHap300 Genotyping BeadChip. Analysis of the raw data was done using the BeadStudio software with the recommended parameters for the Infinium assay and using the genotype cluster files provided by Illumina. Samples with a call rate $<97\%$, identical twins and genetic outliers (identified by classical multidimensional scaling) were excluded from the analysis, resulting in a total of $n = 3925$ (790 Dutch, 656 Swedish, 1079 South Tyrolean, 709 Croatian and 695 Scottish) samples for the downstream analysis. Multidimensional scaling was also used to test for population stratifications. The Dutch, South Tyrolean and Croatian individuals in the association cohorts overlap partly with those in the linkage cohorts. In the initial QC, most SNPs were included in the primary analysis, independent of allele frequencies and with a call rate $>90\%$, with the exception or SNPs deviating strongly form Hardy–Weinberg equilibrium ($P < 1.0 \times 10^{-10}$), resulting in 317465, 318049, 318235, 318236, 318237 in the Croatian, Dutch, Scottish, Swedish and South Tyrolean populations, respectively. QC summary for all SNPs discussed throughout this manuscript is included in Supplementary Material, Table S1. SNPs tested for replication in the Rotterdam Study were genotyped using Infinium II HumanHap550 V.3 BeadChips according to the manufacturer's protocol (Illumina Inc., San Diego, CA, USA).

Statistical analyses

The trait (height) was rank-transformed to be normally distributed with average (=0) and standard deviation (=1) for each population and gender separately. Heritabilities and evidence for linkage were estimated using SOLAR (21).

Linkage analysis. Multipoint-identical-by-descent (MIBD) matrixes were calculated for each population separately using LOKI or Merlin software (22,23), using the deCODE genetic map as a reference. MIBD matrixes of different populations were merged and imported into SOLAR (21), which was used to perform the linkage analysis. Analyses were performed independently for each population and on the combined data, for males and females separately and both sexes together. The threshold for suggestive linkage was set to 2.69 when analyzing the populations/subgroups (Swedish, South Tyrolean, Croatian, Dutch, Females and Males) separately. This is based on correcting the *P*-value for six independent tests. LOD variates are additive, so we adjusted the threshold for suggestive linkage to 3.47 by compensating for four degrees of freedom when combining the LOD scores from the four populations.

GWA analysis and meta-analysis

GWA analyses were performed using the GenABEL R library (24–26) for each population, both combining the genders and analyzing them separately. Genomic control (27) was used to correct standard errors of the effect estimates. The inflation factor lambdas for females/males/all were 1.11/1.07/1.24, 1.41/1.50/1.81, 1.23/1.19/1.45, 1.34/1.30/1.62 and 2.63/2.63/4.19 for the Croatian, Dutch, Scottish, South Tyrolean and Swedish populations, respectively. The meta-analysis was performed by adding the effects (β) and the standard error of the effects (SE β) for each population and gender as:

$$\mathrm{SE}\,\beta_{\mathrm{meta}} = \frac{1}{\sqrt{\sum(1/\mathrm{SE}\,\beta_i^2)}}$$

$$\beta_{\mathrm{meta}} = \mathrm{SE}\,\beta_{\mathrm{meta}}^2 \sum \frac{\beta_i}{\mathrm{SE}\,\beta_i^2}.$$

For each population, standard errors of the estimates were corrected using genomic control procedure to account for relatedness. The *P*-value for the meta-analysis was assessed using χ^2 statistics. Only populations that contributed to the association were included. The genome-wide threshold for significance ($P = 0.05$) after using the Bonferroni approach to correct for multiple testing (318237 SNPs) was calculated to be $P = 1.6 \times 10^{-7}$. To get an unbiased estimate of the effect sizes for the most significant SNPs, we used a score test for association in family-based samples (28), implemented in GenABEL (24–26).

SUPPLEMENTARY MATERIAL

Supplementary Material is available at *HMG* Online.

ACKNOWLEDGEMENTS

We are grateful to all the individuals participating in these studies and the data collection teams in the different locations. We are also grateful to Uppsala Genome Center for performing the genotyping, and the contribution of samples from the Medical Biobank in Umeå is greatly appreciated. We are grateful for the contribution of the district nurse Svea Hennix in the Karesuando study. Special thanks to Lorraine

Anderson and the ORCADES research nurses and Rosa Bisset and Kay Lindsay for administrative support. For the MICROS study, we thank the primary care practitioners Raffaela Stocker, Stefan Waldner, Toni Pizzecco, Josef Plangger, Ugo Marcadent and the personnel of the Hospital of Silandro (Department of Laboratory Medicine) for their participation and collaboration in the research project. We are also grateful to Professor Pavao Rudan and the staff of the Institute for Anthropological Research in Zagreb, Croatia (organization of the field work, anthropometric and physiological measurements and DNA extraction); Professor Ariana Vorko-Jovic and the staff and medical students of the Andrija Stampar School of Public Health of the Faculty of Medicine, University of Zagreb, Croatia (questionnaires, genealogial reconstruction and data entry); Dr Branka Salzer from the biochemistry laboratory 'Salzer', Croatia (measurements of biochemical traits); local general practitioners and nurses (recruitment and communication with the study population); the employees of several other Croatian institutions who participated in the field work, including but not limited to the University of Rijeka and Split, Croatia; Croatian Institute of Public Health; Institutes of Public Health in Split and Dubrovnik, Croatia. We thank Pascal Arp, Mila Jhamai, Dr Michael Moorhouse, Marijn Verkerk and Sander Bervoets for their help in creating the database for the Rotterdam Study.

Conflict of Interest statement. None declared.

FUNDING

All studies are part of the EUROSPAN (European Special Populations Research Network), which was supported by European Commission FP6 STRP grant number 018947 (LSHG-CT-2006-01947). The Northern Swedish Population Health Study was supported by grants from The Swedish Natural Sciences Research Council, The Foundation for Strategic Research (SSF) and The Linneaus Centre for Bioinformatics (LCB). The ORCADES study was supported by the Scottish Executive Health Department, the Royal Society, and DNA extractions were performed at the Wellcome Trust Clinical Research Facility. In South Tyrol, the Ministry of Health of the Autonomous Province of Bolzano and the South Tyrolean Sparkasse Foundation supported the study. The ERF study was supported by the grants from the Netherlands Organization for Scientific Research (NWO, 91203014), joint grant NWO–the Russian Foundation for Basic Research (NWO-RFBR, 047.017.043) and the Center for Medical Systems Biology (CMSB). The Vis study in the Croatian island of Vis was supported through the grants from the Medical Research Council UK to H.C., A.F.W. and I.R.; and Ministry of Science, Education and Sport of the Republic of Croatia to I.R. (number 108-1080315-0302). The COGS study was funded by grants from Cancer Research UK (C348/A3758 and A8896, C48/A6361); Medical Research Council (G0000657-53203); Scottish Executive Chief Scientist's Office (K/OPR/2/2/D333, CZB/4/449); Centre Grant from CORE as part of the Digestive Cancer Campaign. The Rotterdam Study was funded by the European Commission (HEALTH-F2-2008-201865, GEFOS), The Netherlands Organisation of Scientific Research NWO Investments (175.010.2005.011, 911-03-012), the Research Institute for Diseases in the Elderly (NWO 014-93-015, RIDE2), the Netherlands Genomics Initiative (NGI)/Netherlands Organisation for Scientific Research (NWO 050-060-810, NCHA). The Rotterdam Study is also funded by Erasmus Medical Center and Erasmus University, Rotterdam, Netherlands Organization for the Health Research and Development (ZonMw), the Research Institute for Diseases in the Elderly (RIDE), the Ministry of Education, Culture and Science, the Ministry for Health, Welfare and Sports, the European Commission (DG XII) and the Municipality of Rotterdam.

REFERENCES

1. Gudbjartsson, D.F., Walters, G.B., Thorleifsson, G., Stefansson, H., Halldorsson, B.V., Zusmanovich, P., Sulem, P., Thorlacius, S., Gylfason, A., Steinberg, S. *et al.* (2008) Many sequence variants affecting diversity of adult human height. *Nat. Genet.*, **40**, 609–615.
2. Lettre, G., Jackson, A.U., Gieger, C., Schumacher, F.R., Berndt, S.I., Sanna, S., Eyheramendy, S., Voight, B.F., Butler, J.L., Guiducci, C. *et al.* (2008) Identification of ten loci associated with height highlights new biological pathways in human growth. *Nat. Genet.*, **40**, 584–591.
3. Sanna, S., Jackson, A.U., Nagaraja, R., Willer, C.J., Chen, W.-M., Bonnycastle, L.L., Shen, H., Timpson, N., Lettre, G., Usala, G. *et al.* (2008) Common variants in the GDF5-UQCC region are associated with variation in human height. *Nat. Genet.*, **40**, 198–203.
4. Weedon, M.N., Lango, H., Lindgren, C.M., Wallace, C., Evans, D.M., Mangino, M., Freathy, R.M., Perry, J.R.B., Stevens, S., Hall, A.S. *et al.* (2008) Genome-wide association analysis identifies 20 loci that influence adult height. *Nat. Genet.*, **40**, 575–583.
5. Weedon, M.N., Lettre, G., Freathy, R.M., Lindgren, C.M., Voight, B.F., Perry, J.R.B., Elliott, K.S., Hackett, R., Guiducci, C., Shields, B. *et al.* (2007) A common variant of HMGA2 is associated with adult and childhood height in the general population. *Nat. Genet.*, **39**, 1245–1250.
6. Chapman, N.H. and Wijsman, E.M. (1998) Genome screens using linkage disequilibrium tests: optimal marker characteristics and feasibility. *Am. J. Hum. Genet.*, **63**, 1872–1885.
7. Marroni, F., Grazio, D., Pattaro, C., Devoto, M. and Pramstaller, P. (2008) Estimates of genetic and environmental contribution to 43 quantitative traits support sharing of a homogeneous environment in an isolated population from South Tyrol, Italy. *Hum. Hered.*, **65**, 175–182.
8. Tenesa, A., Farrington, S.M., Prendergast, J.G.D., Porteous, M.E., Walker, M., Haq, N., Barnetson, R.A., Theodoratou, E., Cetnarskyj, R., Cartwright, N. *et al.* (2008) Genome-wide association scan identifies a colorectal cancer susceptibility locus on 11q23 and replicates risk loci at 8q24 and 18q21. *Nat. Genet.*, **40**, 631–637.
9. Zanke, B.W., Greenwood, C.M.T., Rangrej, J., Kustra, R., Tenesa, A., Farrington, S.M., Prendergast, J., Olschwang, S., Chiang, T., Crowdy, E. *et al.* (2007) Genome-wide association scan identifies a colorectal cancer susceptibility locus on chromosome 8q24. *Nat. Genet.*, **39**, 989–994.
10. Hofman, A., Breteler, M., van Duijn, C., Krestin, G., Pols, H., Stricker, B., Tiemeier, H., Uitterlinden, A., Vingerling, J. and Witteman, J. (2007) The Rotterdam Study: objectives and design update. *Eur. J. Epidemiol.*, **22**, 819–829.
11. Zeggini, E., Scott, L.J., Saxena, R., Voight, B.F., Marchini, J.L., Hu, T., de Bakker, P.I.W., Abecasis, G.R., Almgren, P., Andersen, G. *et al.* (2008) Meta-analysis of genome-wide association data and large-scale replication identifies additional susceptibility loci for type 2 diabetes. *Nat. Genet.*, **40**, 638–645.
12. Thomas, G., Jacobs, K.B., Yeager, M., Kraft, P., Wacholder, S., Orr, N., Yu, K., Chatterjee, N., Welch, R., Hutchinson, A. *et al.* (2008) Multiple loci identified in a genome-wide association study of prostate cancer. *Nat. Genet.*, **40**, 310–315.
13. Koontz, J.I., Soreng, A.L., Nucci, M., Kuo, F.C., Pauwels, P., van den Berghe, H., Cin, P.D., Fletcher, J.A. and Sklar, J. (2001) Frequent fusion of the *JAZF1* and *JJAZ1* genes in endometrial stromal tumors. *Proc. Natl Acad. Sci. USA*, **98**, 6348–6353.
14. Nakajima, T., Fujino, S., Nakanishi, G., Kim, Y.-S. and Jetten, A.M. (2004) TIP27: a novel repressor of the nuclear orphan receptor TAK1/TR4. *Nucleic Acids Res.*, **32**, 4194–4204.

15. Collins, L.L., Lee, Y.F., Heinlein, C.A., Liu, N.C., Chen, Y.T., Shyr, C.R., Meshul, C.K., Uno, H., Platt, K.A. and Chang, C. (2004) Growth retardation and abnormal maternal behavior in mice lacking testicular orphan nuclear receptor 4. *Proc. Natl Acad. Sci. USA*, **101**, 15058–15063.
16. Aulchenko, Y.S., Heutink, P., Mackay, I., Bertoli-Avella, A.M., Pullen, J., Vaessen, N., Rademaker, T.A.M., Sandkuijl, L.A., Cardon, L., Oostra, B. *et al.* (2004) Linkage disequilibrium in young genetically isolated Dutch population. *Eur. J. Hum. Genet.*, **12**, 527–534.
17. Hofman, A., Grobbee, D.E., de Jong, P.T. and van den Ouweland, F.A. (1991) Determinants of disease and disability in the elderly: the Rotterdam Elderly Study. *Eur. J. Epidemiol.*, **7**, 403–422.
18. Johansson, A., Vavruch-Nilsson, V., Edin-Liljegren, A., Sjolander, P. and Gyllensten, U. (2005) Linkage disequilibrium between microsatellite markers in the Swedish Sami relative to a worldwide selection of populations. *Hum. Genet.*, **116**, 105–113.
19. Liu, F., Arias-Vásquez, A., Sleegers, K., Aulchenko, Y.S., Kayser, M., Sanchez-Juan, P., Feng, B.-J., Bertoli-Avella, A.M., van Swieten, J., Axenovich, T.I. *et al.* (2007) A genomewide screen for late-onset Alzheimer disease in a genetically isolated Dutch population. *Am. J. Hum. Genet.*, **81**, 17–31.
20. Pattaro, C., Marroni, F., Riegler, A., Mascalzoni, D., Pichler, I., Volpato, C., Dal Cero, U., De Grandi, A., Egger, C., Eisendle, A. *et al.* (2007) The genetic study of three population microisolates in South Tyrol (MICROS): study design and epidemiological perspectives. *BMC Med. Genet.*, **8**, 29.
21. Almasy, L. and Blangero, J. (1998) Multipoint quantitative-trait linkage analysis in general pedigrees. *Am. J. Hum. Genet.*, **62**, 1198–1211.
22. Abecasis, G.R., Cherny, S.S., Cookson, W.O. and Cardon, L.R. (2002) Merlin—rapid analysis of dense genetic maps using sparse gene flow trees. *Nat. Genet.*, **30**, 97–101.
23. Heath, S.C. (1997) Markov chain Monte Carlo segregation and linkage analysis for oligogenic models. *Am. J. Hum. Genet.*, **61**, 748–760.
24. Amin, N., van Duijn, C.M. and Aulchenko, Y.S. (2007) A genomic background based method for association analysis in related individuals. *PLoS ONE*, **2**, e1274.
25. Aulchenko, Y.S., de Koning, D.J. and Haley, C. (2007) Genomewide rapid association using mixed model and regression: a fast and simple method for genomewide pedigree-based quantitative trait loci association analysis. *Genetics*, **177**, 577–585.
26. Aulchenko, Y.S., Ripke, S., Isaacs, A. and van Duijn, C.M. (2007) GenABEL: an R library for genome-wide association analysis. *Bioinformatics*, **23**, 1294–1296.
27. Bacanu, S.A., Devlin, B. and Roeder, K. (2000) The power of genomic control. *Am. J. Hum. Genet.*, **66**, 1933–1944.
28. Chen, W.M. and Abecasis, G.R. (2007) Family-based association tests for genome-wide association scans. *Am. J. Hum. Genet.*, **81**, 913–926.