Rare and low-frequency coding variants alter human adult height

A full list of authors and affiliations appears at the end of the paper.

Correspondence to:

Joel N. Hirschhorn (joelh@broadinstitute.org)

Panos Deloukas (p.deloukas@qmul.ac.uk)

Guillaume Lettre (guillaume.lettre@umontreal.ca)

Summary: 150 words

Main text: 2,993 words

Three figures and three tables

SUMMARY

Height is a highly heritable, classic polygenic trait with ~700 common associated variants identified so far through genome-wide association studies. Here, we report 83 new height-associated coding variants with lower minor allele frequencies (range of 0.1-4.8%) and effects of up to 2 cm/allele (*e.g.* in *IHH*, *STC2*, *AR* and *CRISPLD2*), >10 times the average effect of common variants. In functional follow-up studies, rare height-increasing variants of *STC2* (+1-2 cm/allele) compromised proteolytic inhibition of PAPP-A and increased cleavage of IGFBP-4 *in vitro*, resulting in higher bioavailability of insulin-like growth factors. These 83 height-associated variants overlap genes mutated in monogenic growth disorders and highlight new biological candidates (*e.g. ADAMTS3, IL11RA, NOX4*) and pathways (*e.g.* proteoglycan/glycosaminoglycan synthesis) involved in growth. Our results demonstrate that sufficiently large sample sizes can uncover rare and low-frequency variants of moderate to large effect associated with polygenic human phenotypes, and that these variants implicate relevant genes and pathways.

INTRODUCTION

Human height is a highly heritable, polygenic trait[1,2]. The contribution of common DNA sequence variation to inter-individual differences in adult height has been systematically evaluated through genome-wide association studies (GWAS). This approach has thus far identified 697 independent variants located within 423 loci that together explain ~20% of the heritability of height[3]. As is typical of complex traits and diseases, most of the height alleles discovered so far are common (minor allele frequency (MAF) >5%) and are mainly located outside coding regions, complicating the identification of the relevant genes or functional variants. Identifying coding variants associated with a complex trait in new or known loci has the potential to pinpoint causal genes. Furthermore, the extent to which rare (MAF <1%) and low-frequency (1%< MAF ≤ 5%) coding variants also influence complex traits and diseases remains an open question. Many recent DNA sequencing studies have identified only few such variants[4-8], but this limited success could be due to their modest sample size[9]. Some studies have suggested that common sequence variants may explain the majority of the heritable variation in adult height[10], making it timely to assess whether and to what extent rare and low-frequency coding variation contributes to the genetic landscape of this model polygenic trait.

In this study, we used an ExomeChip[11] to test the association between 241,453 variants (83% coding with MAF ≤5%) and adult height variation in 711,418 individuals (discovery and validation sample sizes are 458,927 and 252,491, respectively). The main goals of our project were to determine whether rare and low-frequency coding variants influence the architecture of a model complex human trait, such as adult height, and to discover and characterize new genes and biological pathways implicated in human growth.

RESULTS

32 rare and 51 low-frequency coding variants associated with adult height

We conducted single-variant meta-analyses in a discovery sample of 458,927 individuals, of whom 381,625 were of European ancestry. We validated our association results in an independent set of 252,491 participants. We first performed standard single-variant association analyses; technical details of the discovery and validation steps are in the **Online Methods** (**Supplementary Figs 1-6**, **Supplementary Tables 1-11**). In total, we found 606 independent ExomeChip variants at array-wide significance ($P<2x10^{-7}$), including 252 non-synonymous or splice site variants **(Online Methods** and **Supplementary Table 11**). Focusing on non-synonymous or splice site variants with MAF <5%, our single-variant analyses identified 32 rare and 51 low-frequency height-associated variants (**Tables 1-2**). To date, these 83 height variants (MAF range 0.1-4.8%) represent the largest set of validated rare and low-frequency coding variants associated with any complex human trait or disease. Among these 83 variants, there are 81 missense, one nonsense (in *CCND3*), and one essential acceptor splice site (in *ARMC5*) variants.

We observed a strong inverse relationship between MAF and effect size that is consistent with our statistical power to discover genetic associations (**Fig. 1**). The largest effect sizes were observed for four rare missense variants, located in the androgen receptor gene *AR* (rs137852591, MAF=0.21%, $P_{combined}=2.7x10^{-14}$ under recessive model), in *CRISPLD2* (rs148934412, MAF=0.08%, $P_{combined}=2.4x10^{-20}$), in *IHH* (rs142036701, MAF=0.08%, $P_{combined}=1.9x10^{-23}$), and in *STC2* (rs148833559, MAF=0.1%, $P_{combined}=1.2x10^{-30}$). Carriers of the rare *STC2* missense variant are ~2.1 cm taller than non-carriers, whereas carriers of the remaining three variants (or hemizygous men that carry the X-linked *AR*-rs137852591 rare

allele) are ~2 cm shorter than non-carriers. In comparison, the mean effect size of common height alleles is ten times smaller in the same dataset. Across all 83 rare and low-frequency coding variants, the minor alleles were evenly distributed between height-increasing and -decreasing effects (48% vs. 52%, respectively) (**Fig 1.** and **Tables 1-2**).

Coding variants in new and known height loci, and heritability explained

As expected, many of the height-associated variants in this ExomeChip effort are located near common variants previously associated with height. Of the 83 rare and low-frequency coding variants, 49 fell within 1 Mb of a known height signal, but all were found to be independent after conditional analysis using the July 2015 release of the imputed UK Biobank dataset (which contains individual-level genotype data for both ExomeChip and previously associated common height variants); the remaining 34 define new loci. In addition, we found a further 85 common variants and one low-frequency synonymous variant (in *ACHE*) that define novel loci. Thus, our study identified a total of 120 new height loci, plus 49 additional independent signals from rare and low-frequency coding variants at known loci (**Supplementary Table 11**). Because the sample size of the UK Biobank is smaller than our discovery sample size (120,084 vs. 458,927, respectively), we sought to validate the UK Biobank conditional results using an orthogonal imputation-based methodology implemented in the full discovery set (**Online Methods**). As shown in **Supplementary Fig. 7**, both analytical frameworks produced largely compatible results.

We used the UK Biobank dataset to estimate the contribution of the new height variants to heritability, which is h^2~80% for adult height[2]. In combination, the 83 rare and low-frequency

variants explained 1.7% of the heritability of height. The newly identified novel common variants accounted for another 2.4%, and all independent variants, known and unknown, together explained 27.4% of heritability. By comparison, the 697 known height SNPs explain 23.3% of height heritability in the same dataset. We observed a modest yet significant positive trend between MAF and heritability explained per variant (P=0.012, **Supplementary Fig. 8**), with each common variant explaining slightly more heritability than rare or low-frequency variants (0.029% vs. 0.021%, **Supplementary Fig. 8**).

Gene-based association results

To increase power to find rare or low-frequency coding variants associated with height, we performed gene-based analyses (**Online Methods** and **Supplementary Tables 12-14**). In European-ancestry individuals, the SKAT[12] test (variants with MAF <5% annotated as nonsense, stop-loss, splice site, or damaging missense) identified 99 genes with >1 variant and P_{SKAT} $<2x10^{-6}$ (**Supplementary Table 13**). These 99 genes are enriched for those involved in syndromes of abnormal skeletal growth (previously identified from the Online Mendelian Inheritance in Man (OMIM) database), located near height SNPs identified by GWAS, or predicted to be causal with bioinformatic tools (**Supplementary Fig. 9**)[3,13].

After accounting for gene-based signals explained by a single variant driving the association statistics (**Supplementary Fig. 10**), we identified ten genes that harbor more than one coding variant associated with height variation, and for which the gene-based results remained significant after conditioning on genotypes at nearby common height-associated variants (**Table 3** and **Supplementary Table 15**). Using the same gene-based tests in an independent dataset of 59,804 individuals genotyped on the same exome array, we replicated three genes at P<0.05

(**Table 3**). Further evidence for replication in these genes was seen at the level of single variants (**Supplementary Table 16**). For one of the genes, *OSGIN1,* we found that conditioning on rare variants from this gene affected the results from the single variant analysis. Specifically, two independent variants in *CRISPLD2* (rs149615348, MAF=0.7%, P=9.1x10^{-12}; rs2326458, MAF=26%, P=2.7x10^{-15}) became less significant after conditioning on *OSGIN1* variants (**Supplementary Fig. 11**). Despite this result, *CRISPLD2* is a promising height candidate gene as a third variant in *CRISPLD2*, the rare missense rs148934412 (MAF=0.08%, P=7.6x10^{-14}), remains highly significant after conditioning on *OSGIN1* variants (**Supplementary Fig. 11**). From the gene-based results, three genes – *CSAD*, *NOX4*, and *UGGT2* – fell outside of the loci found by single-variant analyses and are implicated in human height for the first time.

Coding variants implicate biological pathways in human skeletal growth

Prior pathway analyses of height loci identified by GWAS have highlighted gene sets related to both general biological processes (such as chromatin modification and regulation of embryonic size) and more skeletal growth-specific pathways (chondrocyte biology, extracellular matrix (ECM), and skeletal development)[3]. We used two different methods, DEPICT[13] and PASCAL[14] (**Online Methods**), to perform pathway analyses using the ExomeChip results to test whether non-synonymous variants could either independently confirm the relevance of these previously highlighted pathways (and further implicate specific genes in these pathways), or identify new pathways. To compare the pathways emerging from coding and non-coding variation, we applied DEPICT separately on (1) exome array-wide significant coding variants independent of known GWAS signals and (2) non-coding GWAS loci, excluding all novel height-associated genes implicated by coding variants. We identified a total of 496 and 1,623 enriched gene sets,

146 respectively, at a false discovery rate (FDR) <1% (**Supplementary Tables 17-18**); similar
147 analyses with PASCAL yielded 362 and 278 enriched gene sets (**Supplementary Tables 19-20**).
148 Comparison of the results revealed largely shared biology for coding and non-coding variants,
149 especially in the DEPICT analyses, but some pathways showed stronger enrichment with either
150 coding or non-coding variation. In general, coding variants more strongly implicated pathways
151 specific to skeletal growth (such as ECM and bone growth), while GWAS signals highlighted
152 more global biological processes (such as transcription factor binding and embryonic
153 size/lethality)(**Supplementary Fig. 12)**. The two gene sets significant in both DEPICT and
154 PASCAL analyses and that were uniquely implicated by coding variants were "BCAN protein
155 protein interaction subnetwork" and "proteoglycan binding." Both of these pathways relate to the
156 biology of proteoglycans, which are proteins (such as aggrecan) that contain glycosaminoglycans
157 (such as chrondroitin sulfate) and that have well-established connections to skeletal growth[15].
158
159 We also examined which height-associated genes identified by ExomeChip analyses were
160 driving enrichment of pathways such as proteoglycan binding. Using unsupervised clustering
161 analysis to aid in visualization, we observed that a cluster of height-associated genes is strongly
162 implicated in a group of correlated pathways that include biology related to
163 proteoglycans/glycosaminoglycans (**Fig. 2** and **Supplementary Fig. 13**). Strikingly, many of
164 these genes are already annotated in OMIM as underlying disorders of skeletal growth; as such,
165 the remaining genes may be strong candidates for harboring variants that cause Mendelian
166 growth disorders. Within this group are genes that are largely uncharacterized (*SUSD5*), have
167 relevant biochemical functions (*GLT8D2*, a glycosyltransferase studied mostly in the context of
168 the liver[16]; *LOXL4*, a lysyl oxidase expressed in cartilage[17]), modulate pathways known to affect

skeletal growth (*FIBIN*, *SFRP4*)[18,19] or lead to increased body length when knocked out in mice (*SFRP4*)[20].

Functional characterization of rare STC2 variants

To begin exploring whether the identified rare coding variants affect protein function, we performed *in vitro* functional analyses of two rare coding variants in a particularly compelling and novel candidate gene, *STC2*. Over-expression of *STC2* diminishes growth in mice by covalent binding and inhibition of the proteinase PAPP-A, which specifically cleaves IGF binding protein-4 (IGFBP-4), leading to reduced levels of bioactive insulin-like growth factors (**Fig. 3A**)[21]. Although there was no prior genetic evidence implicating *STC2* variation in human growth, the *PAPPA* and *IGFBP4* genes were both implicated in height GWAS[3], and rare mutations in *PAPPA2* cause severe short stature[22], emphasizing the likely relevance of this pathway in humans. The two *STC2* height-associated variants are rs148833559 (p.Arg44Leu, MAF=0.096%, $P_{discovery}=5.7x10^{-15}$) and rs146441603 (p.Met86Ile, MAF=0.14%, $P_{discovery}=2.1x10^{-5}$). These rare alleles increase height by 1.9 and 0.9 cm, respectively, suggesting that they both partially impair STC2 function. In functional studies, STC2 with these amino acid substitutions were expressed at similar levels to wild-type, but showed clear, partial defects in binding to PAPP-A and in inhibition of PAPP-A-mediated cleavage of IGFBP-4 (**Fig. 3B-D**). Thus, the genetic analysis successfully identified rare coding alleles that have demonstrable and predicted functional consequences, strongly confirming the role of these variants and the *STC2* gene in human growth.

Pleiotropic effects and clinical significance

Previous GWAS studies have reported pleiotropic or secondary effects on other phenotypes for many common variants associated with adult height[3,23]. Therefore, we explored to which extent the identified coding variants are associated with 17 human complex phenotypes for which well-powered meta-analysis results were available. Of the 606 height variants identified by single-variant analyses in this study, we found that 96 were associated with at least one of the other investigated traits at array-wide significance ($P<2x10^{-7}$), including one rare and five low-frequency missense variants (see below, **Supplementary Fig. 14**, and **Supplementary Table 21**). Overall, the 606 height signals were enriched for variants nominally associated with body mass index (BMI; $P_{binomial}=1.2x10^{-10}$), LDL-cholesterol (LDL-C; $P_{binomial}=3.5x10^{-6}$), total cholesterol (TC; $P_{binomial}=4.4\ x10^{-8}$), triglycerides (TG; $P_{binomial}=8.9x10^{-7}$) and coronary artery disease (CAD; $P_{binomial}=6.0x10^{-10}$) (**Supplementary Table 22**).

Of the rare and low-frequency missense variants associated with other traits at array-wide significance, the minor alleles at rs77542162 (*ABCA6*, MAF=1.7%) and rs28929474 (*SERPINA1*, MAF=1.8%) were associated with increased height and increased levels of LDL-C and TC, whereas the minor allele at rs3208856 in *CBLC* (MAF=3.4%) was associated with increased height, HDL-cholesterol (HDL-C) and TG, but lower LDL-C and TC levels. The minor allele at rs141845046 (*ZBTB7B,* MAF=2.8%) was associated with both increased height and BMI. The minor alleles at the other two missense variants associated with shorter stature, rs201226914 in *PIEZO1* (MAF=0.2%) and rs35658696 in *PAM* (MAF=4.8%), were associated with decreased glycated haemoglobin (HbA1c) and increased type 2 diabetes risk, respectively. Consistent with a recent report[24], the most pleiotropic variant that we found was the missense rs13107325 (MAF=6.2%) in *SLC39A8*: the minor allele is associated with decreased height,

increased BMI, decreased HDL-C, LDL-C, and TC, but increased TG, and decreased systolic and diastolic blood pressure (**Supplementary Fig. 14** and **Supplementary Table 21**). Rare mutations in *SLC39A8* cause variable short stature phenotypes[25,26], whereas common variants in this gene were previously associated with metabolic syndrome, inflammation, and hypertension[27-30].

Our set of variants associated with height includes several missense variants in genes underlying monogenic syndromes affecting skeletal growth such as *ACAN* (MIM 165800, 608361), *PTH1R* (MIM 60002, 215045), *IHH* (MIM 607778, 112500), *FBN2* (MIM 121050), *ADAMTS10* (MIM 277600) and *ADAMTS17* (MIM 613195) (**Supplementary Table 11**). To further explore the clinical significance of height variants, we queried the ClinVar database and retrieved information on 8,736 variants, including 1,446 markers that are, or predicted to be, pathogenic (**Supplementary Fig. 15** and **Supplementary Table 23**). Of this group, The NIH Genetic Testing Registry recommends testing for four height-associated variants. Two coding variants (rs80356487, MAF=0.03%, β=-1.7 cm; rs1801175, MAF=0.04%, β=-1.2 cm) are located in *G6PC*. Mutations in *G6PC* cause glycogen storage disorder type Ia (von Gierke Disease), which is characterized by growth retardation, delayed puberty, and metabolic abnormalities (MIM 232200). The other two variants are rs1800562 (MAF=6.0%, β=+0.2 cm) in *HFE*, which causes type-1 hemochromatosis (MIM 235200), and rs28929474 (MAF=1.8%, β=+0.8 cm) in the α-1-antitrypsin gene *SERPINA1* (**Supplementary Table 23**). When homozygous, the *SERPINA1*-rs28929474 variant is a cause of emphysema and liver disease in European-descent individuals, and an important risk factor for severe liver disease in cystic fibrosis patients[31]. This is intriguing given that the low-frequency *SERPINA1* allele at rs28929474 is associated with increased height and milder complication in patients with cystic fibrosis due to improve lung functions[32].

DISCUSSION

We undertook an association study of nearly 200,000 non-synonymous variants in 711,418 individuals, and identified 32 rare and 51 low-frequency coding variants associated with adult height. Furthermore, gene-based testing discovered 10 genes that harbor several additional rare/low-frequency variants associated with height, including three genes (*CSAD*, *NOX4*, *UGGT2*) in loci not previously implicated in height. In total, our results highlight 89 genes (10 from gene-based testing and 79 from single-variant analyses (four genes have 2 independent coding variants)) that are likely to modulate human growth, and 24 alleles segregating in the general population that affect height by more than 1 cm (**Tables 1-3**). The rare and low-frequency coding variants explain 1.7% of the heritable variation in adult height. When considering all rare, low-frequency, and common height-associated variants validated in this study, we can now explain 27.4% of the heritability. On a per variant basis, we found that common height SNPs explain more heritability than rare or low-frequency variants (0.029% vs. 0.021%). This suggests that the effect size of rare/low-frequency variants, despite being larger than for common SNPs (**Fig. 1**), is not as large as initially anticipated. Overall, our findings provide strong evidence that rare and low-frequency coding variants contribute to the genetic architecture of height, a model complex human trait.

Our analyses revealed many coding variants in genes mutated in monogenic skeletal growth disorders (**Supplementary Fig. 9**), confirming the presence of allelic series (from familial penetrant mutations to mild effect common variants) in the same genes for related growth phenotypes in humans. We used gene set enrichment-type analyses to demonstrate the functional connectivity between the genes that harbor coding height variants, highlighting known as well as novel biological pathways that regulate height in humans (**Fig. 2**, **Supplementary Fig. 13** and

Supplementary Tables 17-20), and newly implicating genes such as *SUSD5, GLT8D2*, *LOXL4*, *FIBIN,* and *SFRP4* that have not been previously connected with skeletal growth. Additional interesting height candidate genes include *NOX4*, *ADAMTS3* and *ADAMTS6, PTH1R,* and *IL11RA* (**Tables 1-2**, **Supplementary Tables 15** and **24**). *NOX4*, identified through gene-based testing, encodes NADPH oxidase 4, an enzyme that produces reactive oxygen species, a biological pathway not previously implicated in human growth. *Nox4*$^{-/-}$ mice display higher bone density and reduced numbers of osteoclasts, a cell type essential for bone repair, maintenance, and remodelling[12]. We also found rare coding variants in *ADAMTS3* and *ADAMTS6*, genes that encode metalloproteinases that belong to the same family than several other human growth syndromic genes (e.g. *ADAMTS2*, *ADAMTS10*, *ADAMTSL2*). Moreover, we discovered a rare missense variant in *PTH1R* that encodes a receptor of the parathyroid hormone (PTH): PTH-PTH1R signaling is important for bone resorption and mutations in *PTH1R* cause chondrodysplasia in humans[33]. Finally, we replicated the association between a low-frequency missense variant in the cytokine gene *IL11*, but also found a new low-frequency missense variant in its receptor gene *IL11RA*. The IL11-IL11RA axis has been shown to play an important role in bone formation in the mouse[34,35]. Thus, our data confirm the relevance of this signaling cascade in human growth as well. Taken together, the identification of specific genes implicated in human height variation has the potential: (1) to elucidate biological mechanisms that control growth, (2) to provide candidate genes for orphan syndromes characterized by abnormal height phenotypes, and (3) to guide the development of new therapeutic strategies for growth defects. In that regard, the identification of rare missense height-increasing variants of large effect size in *STC2*, and the functional characterization of their effect on IGF signaling, is particularly promising.

ACKNOWLEDGMENTS

A full list of acknowledgments appears in the **Supplementary Note**. Part of this work was conducted using the UK Biobank resource.

AUTHOR CONTRIBUTIONS

Writing Group (wrote and edited manuscript)

Panos Deloukas, Timothy M. Frayling, Mariaelisa Graff, Joel N. Hirschhorn, Guillaume Lettre, Ken Sin Lo, Yingchang Lu, Eirini Marouli, M. Carolina Medina-Gomez, Fernando Rivadeneira.

Data preparation group (checked and prepared data from contributing cohorts for meta-analyses and replication)

Tonu Esko, Mariaelisa Graff, Heather Highland, Anne Justice, Tugce Karaderi, Ken Sin Lo, Adam E. Locke, Yingchang Lu, Eirini Marouli, Nicholas G.D. Masca, M. Carolina Medina-Gomez, Poorva Mudgal, Maggie C.Y. Ng, Manuel A. Rivas, Claudia Schurmann, Kathy Stirrups, Valérie Turcot, Sailaja Vedantam, Thomas W. Winkler, Kristin L. Young.

Height meta-analyses (discovery and replication, single-variant and gene-based)

Panos Deloukas, Timothy M. Frayling, Mariaelisa Graff, Joel N. Hirschhorn, Guillaume Lettre, Daijiang J. Liu, Ken Sin Lo, Yingchang Lu, Eirini Marouli, M. Carolina Medina-Gomez, Fernando Rivadeneira, Andrew R. Wood.

UK Biobank-based integration of height association signals group and heritability analyses

Panos Deloukas, Timothy M. Frayling, Guillaume Lettre, Zoltán Kutalik, Ken Sin Lo, Eirini Marouli, Sina Rüeger, Andrew R. Wood.

Pleiotropy working group

Gonçalo Abecasis, Michael Boehnke, James P. Cook, Panos Deloukas, Fotios Drenos, Jose C.

Florez, Heather Highland, Sekar Kathiresan, Cecilia M. Lindgren, Dajiang J. Liu, Ruth J.F. Loos, Anubha Mahajan, Eirini Marouli, Mark I. McCarthy, Patricia B. Munroe, Gina M. Peloso, John R. B. Perry, Katherine S. Ruth, Cristen J. Willer.

Biological and clinical enrichment, and pathway analyses

Rebecca S. Fine, Joel N. Hirschhorn, Zoltán Kutalik, David Lamparter, Guillaume Lettre, Ken Sin Lo, Tune H. Pers.

Functional characterization of STC2

Troels R. Kjaer, Claus Oxvig.

ONLINE METHODS

Study design & participants

The discovery cohort consisted of 147 studies comprising 458,927 adult individuals of the following ancestries: 1) European descent (N=381,625), 2) African (N=27,494), 3) South Asian (N=29,591), 4) East Asian (N=8,767); 5) Hispanic (N=10,776) and 6) Saudi (N=695). Discovery meta-analysis was carried out in each ancestry group (except the Saudi) separately as well as in the All group. Replication was undertaken in individuals of European ancestry only (**Supplementary Tables 1-3**). Conditional analyses were undertaken only in the European descent group (106 studies, N=381,625).

Phenotype

Height (in centimeters) was corrected for age and the genomic principal components (derived from GWAS data, the variants with MAF >1% on ExomeChip, or ancestry informative markers available on the ExomeChip), as well as any additional study-specific covariates (e.g. recruiting center), in a linear regression model. For studies with non-related individuals, residuals were calculated separately by sex, whereas for family-based studies sex was included as a covariate in the model. Additionally, residuals for case/control studies were calculated separately. Finally, residuals were subject to inverse normal transformation.

Genotype calling

The majority of studies followed a standardized protocol and performed genotype calling using the designated manufacturer software, which was then followed by zCall[36]. For 10 studies participating in the Cohorts for Heart and Aging Research in Genomic Epidemiology

(CHARGE) Consortium, the raw intensity data for the samples from seven genotyping centers were assembled into a single project for joint calling[11]. Study-specific quality control (QC) measures of the genotyped variants was implemented before association analysis (**Supplementary Tables 1-2**).

Study-level statistical analyses

Individual cohorts were analyzed separately for each ancestry population, with either RAREMETALWORKER (http://genome.sph.umich.edu/wiki/RAREMETALWORKER) or RVTEST (http://zhanxw.github.io/rvtests/), to associate inverse normal transformed height data with genotype data taking potential cryptic relatedness (kinship matrix) into account in a linear mixed model. These software are designed to perform score-statistics based rare-variant association analysis, can accommodate both unrelated and related individuals, and provide single-variant results and variance-covariance matrix. The covariance matrix captures linkage disequilibrium (LD) relationships between markers within 1 Mb, which is used for gene-level meta-analyses and conditional analyses[37]. Single-variant analyses were performed for both additive and recessive models.

Centralized quality-control

The individual study data were investigated for potential existence of ancestry population outliers based on 1000 Genome Project phase 1 ancestry reference populations. A centralized QC procedure implemented in EasyQC[38] was applied to individual study association summary statistics to identify outlying studies: (1) assessment of possible problems in height transformation, (2) comparison of allele frequency alignment against 1000 Genomes Project

phase 1 reference data to pinpoint any potential strand issues, and (3) examination of quantile-quantile (QQ) plots per study to identify any problems arising from population stratification, cryptic relatedness and genotype biases.

Meta-analyses

Meta-analyses were carried out in parallel by two different analysts at two sites. We excluded variants if they had call rate <95%, Hardy-Weinberg equilibrium $P<1x10^{-7}$, or large allele frequency deviations from reference populations (>0.6 for all ancestry analyses and >0.3 for ancestry-specific population analyses). We also excluded from downstream analyses markers not present on the Illumina ExomeChip array 1.0, variants on the Y-chromosome or the mitochondrial genome, indels, multiallelic variants, and problematic variants based on the Blat-based sequence alignment analyses. Significance for single-variant analyses was defined at array-wide level ($P<2x10^{-7}$, Bonferroni correction for 250,000 variants). For the gene-based analyses, we applied two different sets of criteria to select variants, based on coding variant annotation from five prediction algorithms (PolyPhen2 HumDiv and HumVar, LRT, MutationTaster and SIFT)[39]. The mask labeled *"broad"* included variants with a MAF <0.05 that are nonsense, stop-loss, splice site, as well as missense variants that are annotated as damaging by at least one program mentioned above. The mask labeled *"strict"* included only variants with MAF <0.05 that are nonsense, stop-loss, splice site, as well as missense variants annotated as damaging by all five algorithms. We used two tests for gene-based testing, namely the SKAT[12] and VT[40] tests. Statistical significance for gene-based tests was set at a Bonferroni-corrected threshold of $P<2x10^{-6}$ (threshold for 25,000 genes; we did not correct for the four tests given that they are correlated and that we sought validation in independent studies).

394

395 ***Genomic inflation***

396 We observed a marked genomic inflation of the test statistics even after adequate control for
397 population stratification (linear mixed model) arising mainly from common markers; λ_{GC} in
398 European-ancestry was 1.2 and 2.7 for all and common markers, respectively (**Supplementary**
399 **Fig. 2** and **Supplementary Table 8**). Such inflation is expected for a highly polygenic trait like
400 height, and is consistent with our very large sample size[3,41]. To confirm this, we applied the
401 recently developed linkage disequilibrium (LD) score regression method to our height
402 ExomeChip results[42], with the caveats that the method was developed (and tested) with >200,000
403 common markers available. We restricted our analyses to 15,848 common variants (MAF ≥5%)
404 from the European-ancestry meta-analysis, and matched them to pre-computed LD scores for the
405 European reference dataset[42]. The intercept of the regression of the χ^2 statistics from the height
406 meta-analysis on the LD score estimate the inflation in the mean χ^2 due to confounding bias,
407 such as cryptic relatedness or population stratification. The intercept is 1.4 (standard error =
408 0.07), which is small when compared to the λ_{GC} of 2.7. The ratio statistic of (intercept -1) /
409 (mean χ^2 -1) is 0.067 (standard error = 0.012), well within the normal range[42], suggesting that
410 most of the inflation (~93%) observed in the height association statistics is due to polygenic
411 effects (**Supplementary Fig. 3**).

412

413 Furthermore, to exclude the possibility that some of the observed associations between height
414 and rare/low-frequency variants could be due to allele calling problems in the smaller studies, we
415 performed a sensitivity meta-analysis with primarily Europe-ancestry studies totaling >5,000

participants. We found very concordant effect sizes, suggesting that smaller studies do not bias our results (**Supplementary Fig. 4**).

Conditional analyses

The RAREMETAL R-package[43] and the GCTA v1.24[44] software were used to identify independent height association signals across the European descent meta-analysis results. RAREMETAL performs conditional analyses by using covariance matrices in order to distinguish true signals from those driven by LD with adjacent known variants. First, we identified the lead variants ($P<2x10^{-7}$) based on a 1 Mb window centered on the most significantly associated variant and performed LD pruning ($r^2<0.3$) to avoid downstream problems in the conditional analyses due to co-linearity. We then conditioned on the LD-pruned set of lead variants in RAREMETAL and kept new lead signals at $P<2x10^{-7}$. The process was repeated until no additional signal emerged below the pre-specified P-value threshold. The use of a 1Mb window in RAREMETAL can obscure dependence between conditional signals in adjacent intervals in regions of extended LD. To detect such instances, we performed joint analyses using GCTA in the ARIC and UK ExomeChip reference panels, both of which comprise >10,000 individuals of European descent. Gene-based conditional analyses were also performed in RAREMETAL.

The newly discovered 120 height variants were conditioned on the previously published GWAS height variants[3] in the first release of the UK Biobank imputed dataset using regression methodology implemented in BOLT-LMM[45]. We also explored an alternative approach based on approximate conditional analysis[44]. This latter method relies on summary statistics available

from the same cohort, thus we first imputed summary statistics[46] for exome variants, using summary statistics from the Wood et *al*. 2014 study[3]. Conversely, we imputed the top variants from the Wood et *al*. 2014 study using the summary statistics from the ExomeChip. Subsequently, we calculated effect sizes for each exome variant conditioned on the Wood et *al*. 2014 top variants in two ways. First, we conditioned the imputed summary statistics of the exome variant on the summary statistics of the Wood et *al*. 2014 top variants that fell within 5 Mb of the target ExomeChip variant. Second, we conditioned the summary statistics of the ExomeChip variant on the imputed summary statistics of the Wood et *al*. 2014 hits. We then selected the option that yielded a higher imputation quality. For poorly tagged variants ($\hat{r}^2 < 0.8$), we simply used up-sampled HapMap summary statistics for the approximate conditional analysis. Pairwise SNP-by-SNP correlations were estimated from the UK10K data (TwinsUK[47] and ALSPAC[48] studies , N=3,781).

Description of the single-variant analyses

We conducted single-variant meta-analyses in a discovery sample of 458,927 individuals of different ancestries using both additive and recessive genetic models (**Supplementary Fig. 1** and **Supplementary Tables 1-4**). The combined additive analyses identified 1,455 unique variants that reached array-wide significance ($P<2x10^{-7}$), including 578 non-synonymous and splice site variants (**Supplementary Tables 5-7**). Under the additive model, we observed a high genomic inflation of the test statistics (*e.g.* λ_{GC} of 2.7 in European-ancestry studies for common markers, **Supplementary Fig. 2** and **Supplementary Table 8**), although replication results (see below) and additional sensitivity analyses (see above) suggest that it is consistent with polygenic inheritance as opposed to population stratification, cryptic relatedness, or technical artifacts

462 (**Supplementary Figs 3-4**). The majority of these 1,455 association signals (1,241; 85.3%) were
463 found in the European-ancestry meta-analysis (85.5% of the discovery sample size)
464 **(Supplementary Fig. 5)**. Nevertheless, we discovered eight associations within five loci in our
465 all-ancestry analyses that are driven by African studies (including one missense variant in the
466 growth hormone gene *GH1* (rs151263636), **Supplementary Fig. 6**), three height variants found
467 only in African studies, and one rare missense marker associated with height in South Asians
468 only (**Supplementary Table 7**).
469
470 Several studies, totaling 252,491 independent individuals of European ancestry, became
471 available after the completion of the discovery analyses, and were thus used for validation of our
472 experiment. First, we considered the 81 variants with suggestive association in the discovery
473 analyses ($2x10^{-7}<P_{discovery}\leq 2x10^{-6}$). Of those 81 variants, 55 reached significance after combining
474 discovery and replication results based on $P_{combined}<2x10^{-7}$ (**Supplementary Table 9**).
475 Furthermore, recessive modeling confirmed seven new independent markers with $P_{combined}<2x10^{-7}$
476 , including one rare missense variant (rs137852591, MAF 0.21%) in the *AR* gene
477 (**Supplementary Table 10**). To test the independence and integrate all height markers from the
478 discovery and replication phase, we used conditional analyses and GCTA "joint" modeling[44] in
479 the combined discovery and replication set. This resulted in the identification of 606 independent
480 height variants, including 252 non-synonymous or splice site variants (**Supplementary Table**
481 **11**). Of the 606 variants, 605 had concordant direction of effect between the discovery and
482 validation studies, and 595 variants had a $P_{validation}<0.05$ (482 variants with $P_{validation}<8x10^{-5}$,
483 Bonferroni correction for 606 tests), suggesting a very low false discovery rate (**Supplementary**
484 **Table 11**).

485

486 ***Pathway analyses***

487 DEPICT is a computational framework that uses probabilistically-defined reconstituted gene sets
488 to perform gene set enrichment and gene prioritization[13]. For a description about gene set
489 reconstitution please refer to references [13] and [49]. In brief, reconstitution was performed by
490 extending pre-defined gene sets (such as Gene Ontology terms, canonical pathways, protein-
491 protein interaction subnetworks and rodent phenotypes) with genes co-regulated with genes in
492 these pre-defined gene set using large-scale microarray-based transcriptomics data. In order to
493 adapt the gene set enrichment part of DEPICT for ExomeChip data, we made two principal
494 changes. First and foremost, because DEPICT for GWAS incorporates all genes within a given
495 LD block around each index SNP, we modified DEPICT to take as input only the gene directly
496 impacted by the coding SNP. Second, we adapted the way DEPICT adjust for confounders (such
497 as gene length) by generating null ExomeChip association results using Swedish ExomeChip
498 data (Malmö Diet and Cancer (MDC), All New Diabetics in Scania (ANDIS), and Scania
499 Diabetes Registry (SDR) cohorts, N=11,899) and randomly assigning phenotypes from a normal
500 distribution before conducting association analysis (see **Supplementary Note**). For the gene set
501 enrichment analysis of the ExomeChip data, we used significant non-synonymous variants
502 statistically independent of known GWAS hits (and that were present in the null ExomeChip
503 data; See Supplementary Note for details). For gene set enrichment analysis of the GWAS data,
504 we used all loci (1) with a non-coding index SNP and (2) that did not contain any of the novel
505 ExomeChip genes. In visualizing the analysis, we used affinity propagation clustering[50] to group
506 the most similar reconstituted gene sets based on their gene memberships (See Supplementary
507 Note). Within a "meta-gene set", the best P-value of any member gene set was used as

representative for comparison. DEPICT for ExomeChip was written using the Python programming language and the code can be found at https://github.com/perslab/ec-depict.

We also applied the PASCAL pathway analysis tool[14] to genome-wide association summary statistics for all coding variants. In brief, the method derives gene-based scores (both SUM and MAX statistics) and subsequently tests for the over-representation of high gene scores in predefined biological pathways. We used standard pathway libraries from KEGG, REACTOME and BIOCARTA, and also added dichotomized (Z-score>3) reconstituted gene sets from DEPICT[13]. To accurately estimate SNP-by-SNP correlations even for rare variants, we used the UK10K data (TwinsUK[47] and ALSPAC[48] studies , N=3781). In order to separate the contribution of regulatory variants from the coding variants, we also applied PASCAL to association summary statistics of only regulatory variants (20 kb upstream, gene body excluded) from the Wood et *al*. study[3]. In this way, we could classify pathways driven principally by coding, regulatory or mixed signals.

Validation

We performed single-variant and gene-based association analyses for eight validation studies, totaling 59,804 participants, genotyped on the Exomechip using RAREMETAL[37]. We sought additional evidence for association for the top signals in two independent studies in the UK (UK Biobank) and Iceland (deCODE), comprising 120,084 and 72,613 individuals, respectively. We used the same QC and analytical methodology as described above. Genotyping and study descriptives are provided in **Supplementary Tables 1-3**. For the combined analysis, we used the inverse-variance weighted fixed effects meta-analysis method using METAL[51]. Significant

associations were defined as those with a combined meta-analysis (discovery and validation) $P_{combined}<2x10^{-7}$.

STC2 functional experiments

Mutagenesis, cell culture and transfection. For the generation of STC2 mutants (R44L and M86I), wild-type STC2 cDNA contained in pcDNA3.1/Myc-His(-) (Invitrogen)[21] was used as a template. Mutagenesis was carried out using Quickchange (Stratagene), and all constructs were verified by sequence analysis. Recombinant wild-type STC2 and variants were expressed in human embryonic kidney (HEK) 293T cells (293tsA1609neo) maintained in high-glucose DMEM supplemented 10% fetal bovine serum, 2 mM glutamine, nonessential amino acids, and gentamicin. Cells ($6x10^6$) were plated onto 10 cm-dishes and transfected 18 h later by calcium phosphate coprecipitation using 10 μg plasmid DNA. Media were harvested 48 h post transfection, cleared by centrifugation, and stored at -20°C until use. Protein concentrations (58-66 nM) were determined by TRIFMA using antibodies described previously[21]. PAPP-A was expressed stably in HEK293T cells as previously reported[52]. Expressed levels of PAPP-A (27.5 nM) were determined by a commercial ELISA (AL-101, Ansh Labs, TX).

STC2 and PAPP-A complex formation. Culture supernatants containing wild-type STC2 or variants were adjusted to 58 nM, added an equal volume of culture supernatant containing PAPP-A corresponding to a 2.1-fold molar excess, and incubated at 37°C. Samples were taken at 1, 2, 4, 6, 8, 16, and 24 h and stored at -20^{o}C.

553 *Analysis of proteolytic activity.* Specific proteolytic cleavage of ^{125}I-labeled IGFBP-4 is
554 described in detail elsewhere[53]. Briefly, the PAPP-A:STC2 complex mixtures were diluted
555 (1:190) to a concentration of 145 pM PAPP-A and mixed with preincubated ^{125}I-IGFBP4 (10
556 nM) and IGF-1 (100 nM) in 50 mM Tris-HCl, 100 mM NaCl, 1 mM $CaCl_2$. Following 1 h
557 incubation at 37^{o}C, reactions were terminated by the addition of SDS-PAGE sample buffer
558 supplemented with 25 mM EDTA. Substrate and co-migrating cleavage products were separated
559 by 12% nonreducing SDS-PAGE and visualized by autoradiography using a storage phosphor
560 screen (GE Healthcare) and a Typhoon imaging system (GE Healthcare). Band intensities were
561 quantified using ImageQuant TL 8.1 software (GE Healthcare).
562
563 *Western blotting.* STC2 and covalent complexes between STC2 and PAPP-A were blotted onto
564 PVDF membranes (Millipore) following separation by 3-8% SDS-PAGE. The membranes were
565 blocked with 2% Tween-20, and equilibrated in 50 mM Tris-HCl, 500 mM NaCl, 0.1% Tween-
566 20, pH 9 (TST). For STC2, the membranes were incubated with goat polyclonal anti-STC2
567 (R&D systems, AF2830) at 0.5 μg/ml in TST supplemented with 2% skim milk for 1 h at 20°C.
568 For PAPP-A:STC2 complexes, the membranes were incubated with rabbit polyclonal anti-
569 PAPP-A[54] at 0.63 μg/ml in TST supplemented with 2% skim milk for 16 h at 20°C. Membranes
570 were washed with TST and subsequently incubated with polyclonal swine anti-rabbit IgG-HRP
571 (DAKO, P0217) or polyclonal rabbit anti-goat IgG-HRP (DAKO, P0449), respectively, diluted
572 1:2000 in TST supplemented with 2% skim milk for 1 h at 20°C. Following washing with TST,
573 membranes were developed using enhanced chemiluminescence (ECL Prime, GE Healthcare).
574 Images were captured using an ImageQuant LAS 4000 instrument (GE Healthcare).
575

576 ***Pleiotropy analyses***

577 We accessed ExomeChip data from GIANT (BMI, waist-hip ratio), GLGC (total cholesterol
578 (TC), triglycerides (TG), HDL-cholesterol (HDL-C), LDL-cholesterol (LDL-C)), IBPC (systolic
579 and diastolic blood pressure), MAGIC (glycaemic traits), REPROGEN (age at menarche and
580 menopause), and DIAGRAM (type 2 diabetes). For coronary artery disease, we accessed 1000
581 Genomes Project-imputed GWAS data released by CARDIoGRAMplusC4D[55].
582

583 **URLs**

584 ClinVar, http://www.ncbi.nlm.nih.gov/clinvar/

585 DEPICT, http://www.broadinstitute.org/mpg/depict/

586 ExomeChip, http://genome.sph.umich.edu/wiki/Exome_Chip_Design

587 ExomeDEPICT, https://github.com/perslab/ec-depict

588 OMIM, http://omim.org/

589 PASCAL, http://www2.unil.ch/cbg/index.php?title=Pascal

590 RAREMETALWORKER, http://genome.sph.umich.edu/wiki/RAREMETALWORKER

591 RVTEST, http://zhanxw.github.io/rvtests/

Table 1. Rare variants associated with adult height. 32 coding or splice site variants with minor allele frequency <1% in European-ancestry participants that have $P_{combined}$ <2x10^{-7}. All markers are significant under an additive genetic model, except *AR*-rs137852591, which was discovered using the recessive model in the all-ancestry analysis. The direction of the effect (Beta) and effect allele frequency (AF) is given for the alternate (Alt) allele. Genomic coordinates are on build 37 of the human genome. For each variant, we provide the most severe annotation using the ENSEMBL Variant Effect Predictor (VEP) tool. N, sample size; Ref, reference allele; SE, standard error.

					Discovery (N up to 381,625)				Validation (N up to 252,491)				Combined (N up to 634,116)			
Variant	**Chr:Pos**	**Ref/Alt**	**Gene**	**Annotation**	**AF**	**Beta**	**SE**	**P-value**	**AF**	**Beta**	**SE**	**P-value**	**AF**	**Beta**	**SE**	**P-value**
rs150341307	1:32673514	G/C	*IQCC*	missense	0.002	-0.141	0.026	7.92E-08	0.004	-0.116	0.025	3.83E-06	0.003	-0.128	0.018	1.34E-12
rs143365597	1:41540902	G/A	*SCMH1*	missense	0.004	0.188	0.018	1.58E-25	0.006	0.169	0.024	9.42E-13	0.005	0.181	0.014	1.35E-36
rs114233776	1:41618297	G/A	*SCMH1*	missense	0.006	-0.119	0.015	1.92E-15	0.006	-0.11	0.019	1.32E-08	0.006	-0.116	0.012	1.80E-22
rs145659444	1:149902342	C/T	*MTMR11*	missense	0.007	0.067	0.015	4.16E-06	0.006	0.083	0.019	7.11E-06	0.007	0.073	0.012	3.03E-10
rs144712473	1:183495812	A/G	*SMG7*	missense	0.006	-0.094	0.014	4.97E-11	0.008	-0.067	0.017	8.94E-05	0.007	-0.083	0.011	1.61E-14
rs144673025	1:223178026	T/C	*DISP1*	missense	0.008	-0.078	0.013	1.11E-09	0.007	-0.086	0.018	1.22E-06	0.008	-0.081	0.011	1.27E-14
rs142036701	2:219924961	G/T	*IHH*	missense	0.001	-0.32	0.04	1.09E-15	0.003	-0.263	0.043	1.48E-09	0.002	-0.294	0.029	1.85E-23
rs147445258	2:220078652	C/T	*ABCB6*	missense	0.01	-0.086	0.012	3.43E-13	0.009	-0.064	0.018	4.40E-04	0.01	-0.079	0.01	2.47E-15
rs121434601	3:46939587	C/T	*PTH1R*	missense	0.003	0.154	0.023	1.30E-11	0.003	0.192	0.031	5.48E-10	0.003	0.168	0.019	1.14E-19
rs141374503	4:73179445	C/T	*ADAMTS3*	missense	0.003	-0.119	0.021	1.82E-08	0.004	-0.089	0.023	1.32E-04	0.004	-0.106	0.016	1.30E-11
rs149385790	4:120422407	T/G	*PDE5A*	missense	0.001	0.257	0.031	7.50E-17	0.005	0.19	0.033	1.28E-08	0.003	0.226	0.023	2.65E-23
rs146301345	5:32784907	G/A	*NPR3*	missense	0.003	0.128	0.022	1.05E-08	0.002	0.166	0.035	1.78E-06	0.003	0.139	0.019	7.91E-14
rs61736454	5:64766798	G/A	*ADAMTS6*	missense	0.002	-0.152	0.026	7.82E-09	0.002	-0.182	0.032	1.37E-08	0.002	-0.164	0.02	4.80E-16
rs78727187	5:127668685	G/T	*FBN2*	missense	0.006	0.183	0.015	2.47E-33	0.006	0.181	0.02	5.06E-20	0.006	0.182	0.012	1.47E-52
rs148833559	5:172755066	C/A	*STC2*	missense	0.001	0.29	0.037	5.69E-15	0.001	0.368	0.043	1.32E-17	0.001	0.323	0.028	1.15E-30
rs148543891	6:155450779	A/G	*TIAM2*	missense	0.003	-0.124	0.022	1.45E-08	0.001	-0.016	0.082	8.50E-01	0.003	-0.117	0.021	3.96E-08
rs41511151	7:73482987	G/A	*ELN*	missense	0.004	-0.086	0.018	2.63E-06	0.007	-0.061	0.019	1.51E-03	0.006	-0.074	0.013	2.31E-08
rs112892337	8:135614553	G/C	*ZFAT*	missense	0.004	0.196	0.019	4.42E-26	0.004	0.184	0.024	1.20E-14	0.004	0.191	0.015	6.12E-38
rs75596750	8:135622851	G/A	*ZFAT*	missense	0.001	0.255	0.036	1.54E-12	0.002	0.339	0.039	5.94E-18	0.002	0.293	0.027	2.05E-28
rs138273386	11:27016360	G/A	*FIBIN*	missense	0.004	-0.12	0.017	5.79E-12	0.005	-0.076	0.024	1.56E-03	0.004	-0.105	0.014	3.26E-14
rs138059525	11:94533444	G/A	*AMOTL1*	missense	0.009	-0.096	0.012	9.01E-16	0.007	-0.089	0.017	3.84E-07	0.008	-0.094	0.01	2.84E-21
rs147996581	12:58138971	G/A	*TSPAN31*	missense	0.003	-0.116	0.022	8.26E-08	0.001	-0.268	0.09	2.85E-03	0.003	-0.125	0.021	5.50E-09
rs13141	12:121756084	G/A	*ANAPC5*	missense	0.009	-0.082	0.012	1.09E-11	0.011	-0.105	0.016	1.44E-11	0.01	-0.091	0.01	1.45E-21
rs150494621	15:44153571	C/T	*WDR76*	missense	0.008	0.063	0.013	1.56E-06	0.014	0.054	0.015	3.42E-04	0.011	0.059	0.01	2.32E-09
rs141308595	15:89424870	G/T	*HAPLN3*	missense	0.001	-0.267	0.037	2.84E-13	0.002	-0.234	0.035	2.43E-11	0.002	-0.25	0.025	1.02E-22
rs141923065	16:31474091	A/G	*ARMC5*	splice_acceptor	0.006	0.104	0.015	5.88E-12	0.013	0.057	0.018	1.16E-03	0.009	0.084	0.011	1.62E-13
rs34667348	16:47684830	C/A	*PHKB*	missense	0.005	0.121	0.016	3.96E-14	0.005	0.033	0.020	1.04E-01	0.005	0.088	0.013	3.43E-12
rs140385822	16:67470505	G/A	*HSD11B2*	missense	0.002	-0.148	0.028	1.27E-07	0.002	-0.124	0.035	3.38E-04	0.002	-0.139	0.022	1.97E-10
rs149615348	16:84900645	G/A	*CRISPLD2*	missense	0.007	-0.095	0.014	9.13E-12	0.008	-0.098	0.017	4.34E-09	0.008	-0.096	0.011	2.92E-19
rs148934412	16:84902472	G/A	*CRISPLD2*	missense	0.001	-0.297	0.04	7.75E-14	0.001	-0.317	0.058	3.49E-08	0.001	-0.304	0.033	2.36E-20
rs201226914	16:88798919	G/T	*PIEZO1*	missense	0.002	-0.187	0.027	5.27E-12	0.002	-0.241	0.043	1.99E-08	0.002	-0.202	0.023	8.68E-19
rs137852591	23:66941751	C/G	*AR*	missense	0.002	-0.304	0.061	7.05E-07	0.008	-0.333	0.058	7.12E-09	0.005	-0.319	0.042	2.67E-14

598 **Table 2. Low-frequency variants associated with adult height.** 59 variants (51 coding) with minor allele frequency between 1 and 5% in
599 European-ancestry participants that have $P_{combined}$ <2x10^{-7} under an additive genetic model. For *TTN*-rs16866412 and *NOL8*-rs921122, the
600 association is significant (P<2x10^{-7}) upon conditional analysis. The direction of the effect (Beta) and effect allele frequency (AF) is given for the
601 alternate (Alt) allele. For each variant, we provide the most severe annotation using the ENSEMBL Variant Effect Predictor (VEP) tool. N, sample
602 size; Ref, reference allele; SE, standard error

					Discovery (N up to 381,625)				Validation (N up to 252,491)				Combined (N up to 634,116)			
Variant	**Chr:Pos**	**Ref/Alt**	**Gene**	**Annotation**	**AF**	**Beta**	**SE**	**P-value**	**AF**	**Beta**	**SE**	**P-value**	**AF**	**Beta**	**SE**	**P-value**
rs41292521	1:51873967	G/A	*EPS15*	missense	0.020	0.045	0.008	5.07E-08	0.023	0.065	0.010	7.60E-11	0.021	0.053	0.006	2.56E-17
rs61730011	1:119427467	A/C	*TBX15*	missense	0.042	-0.059	0.006	1.61E-24	0.046	-0.056	0.007	4.19E-15	0.044	-0.058	0.005	2.79E-36
rs11580946	1:150551327	G/A	*MCL1*	missense	0.014	0.061	0.010	2.16E-09	0.015	0.085	0.012	7.86E-12	0.015	0.070	0.008	1.55E-19
rs141845046	1:154987704	C/T	*ZBTB7B*	missense	0.028	0.058	0.007	7.30E-17	0.025	0.061	0.010	4.46E-10	0.027	0.059	0.006	3.46E-25
rs79485039	1:180886140	C/T	*KIAA1614*	missense	0.026	0.034	0.007	1.41E-06	0.031	0.030	0.009	4.51E-04	0.028	0.033	0.006	2.63E-09
rs52826764	2:20205541	C/T	*MATN3*	missense	0.026	-0.071	0.007	2.67E-23	0.028	-0.084	0.010	6.60E-19	0.027	-0.076	0.006	3.74E-41
rs16859517	2:219949184	C/T	*NHEJ1*	intron	0.036	0.059	0.006	5.96E-21	0.036	0.064	0.008	1.12E-15	0.036	0.061	0.005	8.20E-37
rs16866412	2:179474668	G/A	*TTN*	missense	0.013	-0.053	0.010	1.35E-07	0.010	-0.019	0.015	2.15E-01	0.012	-0.042	0.008	3.44E-07
rs7571816	2:233077064	A/G	*DIS3L2*	intron	0.025	-0.060	0.007	2.35E-16	0.023	-0.079	0.010	2.58E-15	0.024	-0.066	0.006	6.46E-31
rs2229089	3:14214524	G/A	*XPC*	missense	0.031	-0.038	0.007	1.22E-08	0.035	-0.020	0.008	1.68E-02	0.033	-0.030	0.005	1.29E-08
rs76208147	3:47162886	C/T	*SETD2*	missense	0.019	0.048	0.009	2.24E-08	0.016	0.062	0.012	2.22E-07	0.018	0.053	0.007	1.65E-13
rs35713889	3:49162583	C/T	*LAMB2*	missense	0.039	0.043	0.006	3.28E-12	0.045	0.060	0.007	1.33E-16	0.041	0.050	0.005	3.49E-27
rs9838238	3:98600385	T/C	*DCBLD2*	missense	0.047	0.029	0.005	1.23E-07	0.051	0.027	0.007	5.62E-05	0.048	0.028	0.004	1.68E-12
rs11722554	4:5016883	G/A	*CYTL1*	missense	0.040	-0.049	0.006	2.01E-17	0.034	-0.057	0.009	6.68E-11	0.038	-0.052	0.005	1.86E-25
rs61730641	4:87730980	C/T	*PTPN13*	missense	0.015	-0.086	0.010	1.94E-19	0.016	-0.094	0.012	1.38E-15	0.016	-0.089	0.008	9.43E-32
rs116807401	4:135121721	T/C	*PABPC4L*	missense	0.017	0.065	0.009	1.39E-13	0.016	0.045	0.012	1.33E-04	0.017	0.058	0.007	7.54E-16
rs28925904	4:144359490	C/T	*GAB1*	missense	0.019	-0.048	0.008	1.04E-08	0.023	-0.036	0.010	3.24E-04	0.021	-0.043	0.006	4.29E-12
rs34343821	4:154557616	C/T	*KIAA0922*	missense	0.011	0.059	0.011	7.75E-08	0.015	0.056	0.012	5.75E-06	0.013	0.058	0.008	2.18E-12
rs35658696	5:102338811	A/G	*PAM*	missense	0.048	-0.025	0.005	3.76E-06	0.053	-0.031	0.007	8.47E-06	0.050	-0.027	0.004	1.63E-10
rs34821177	5:126250812	C/T	*MARCH3*	missense	0.036	0.034	0.006	4.25E-08	0.029	0.027	0.009	2.45E-03	0.034	0.032	0.005	1.67E-10
rs62623707	5:135288632	A/G	*LECT2*	missense	0.044	-0.030	0.006	1.02E-07	0.049	-0.024	0.007	4.77E-04	0.046	-0.027	0.005	1.36E-09
rs34471628	5:172196752	A/G	*DUSP1*	missense	0.036	0.048	0.006	4.00E-14	0.042	0.036	0.007	1.26E-06	0.039	0.043	0.005	1.93E-20
rs28932177	5:176637471	G/A	*NSD1*	missense	0.028	0.063	0.007	2.38E-17	0.027	0.065	0.009	2.62E-12	0.028	0.064	0.006	4.27E-30
rs78247455	5:176722005	G/A	*NSD1*	missense	0.023	-0.083	0.008	1.86E-26	0.025	-0.085	0.010	8.42E-18	0.024	-0.084	0.006	2.32E-41
rs7757648	6:30851933	G/A	*DDR1*	intron	0.013	-0.075	0.013	1.11E-08	0.011	-0.079	0.018	1.24E-05	0.012	-0.076	0.011	4.64E-13
rs34427075	6:34730395	C/T	*SNRPC*	synonymous	0.014	-0.117	0.010	9.21E-33	0.016	-0.139	0.012	9.59E-31	0.015	-0.126	0.008	3.45E-60
rs33966734	6:41903798	C/A	*CCND3*	stop_gained	0.013	-0.140	0.017	5.51E-17	0.011	-0.101	0.018	3.41E-08	0.012	-0.122	0.012	1.28E-22
rs17277546	7:99489571	G/A	*TRIM4*	3'UTR	0.049	0.034	0.005	3.28E-10	0.052	0.038	0.007	2.26E-07	0.050	0.035	0.004	1.40E-17
rs7636	7:100490077	G/A	*ACHE*	synonymous	0.043	-0.037	0.006	8.59E-10	0.035	-0.019	0.009	2.92E-02	0.040	-0.031	0.005	2.98E-10
rs17480616	7:135123060	G/C	*CNOT4*	missense	0.028	0.060	0.007	2.31E-17	0.030	0.054	0.009	5.04E-10	0.029	0.058	0.005	3.90E-26
rs3136797	8:42226805	C/G	*POLB*	missense	0.018	0.044	0.009	1.95E-06	0.021	0.026	0.010	1.30E-02	0.019	0.036	0.007	1.88E-07
rs11575580	9:34660864	C/T	*IL11RA*	missense	0.016	-0.064	0.009	5.20E-13	0.020	-0.030	0.011	4.42E-03	0.018	-0.050	0.007	4.01E-13
rs921122	9:95063947	C/T	*NOL8*	missense	0.039	0.041	0.009	2.56E-06	0.040	0.018	0.008	3.45E-02	0.040	0.029	0.006	3.33E-06

rs41274586	10:79580976	G/A	*DLG5*	missense	0.017	-0.058	0.009	2.72E-11	0.017	-0.076	0.012	5.15E-11	0.017	-0.065	0.007	7.66E-20
rs41291604	10:97919011	A/G	*ZNF518A*	missense	0.040	0.031	0.006	9.94E-08	0.040	0.022	0.008	3.05E-03	0.040	0.028	0.005	3.91E-09
rs71455793	11:65715204	G/A	*TSGA10IP*	missense	0.039	-0.058	0.006	1.82E-21	0.046	-0.072	0.007	1.41E-23	0.042	-0.064	0.005	1.52E-43
rs4072796	12:7548996	C/G	*CD163L1*	missense	0.035	0.034	0.006	4.11E-08	0.037	0.015	0.008	6.68E-02	0.036	0.027	0.005	1.87E-08
rs61743810	12:69140339	G/C	*SLC35E3*	missense	0.022	-0.047	0.008	1.13E-09	0.023	-0.036	0.010	5.11E-04	0.022	-0.043	0.006	1.29E-11
rs117801489	12:104408832	T/C	*GLT8D2*	missense	0.017	0.053	0.009	8.72E-10	0.028	0.062	0.010	5.82E-10	0.022	0.057	0.007	1.60E-17
rs2066674	13:50842259	G/A	*DLEU1*	intron	0.044	0.073	0.006	2.33E-37	0.041	0.084	0.008	7.02E-25	0.043	0.077	0.005	5.66E-57
rs17880989	14:23313633	G/A	*MMP14*	missense	0.027	0.041	0.007	1.72E-08	0.029	0.052	0.009	7.81E-09	0.028	0.045	0.006	3.27E-16
rs34354104	14:24707479	G/A	*GMPR2*	missense	0.048	0.045	0.005	3.67E-16	0.050	0.047	0.007	1.34E-11	0.049	0.046	0.004	2.13E-29
rs117295933	14:45403699	C/A	*KLHL28*	missense	0.016	-0.045	0.009	1.55E-06	0.025	-0.036	0.010	4.13E-04	0.020	-0.041	0.007	3.05E-09
rs41286548	14:70633411	C/T	*SLC8A3*	missense	0.021	-0.054	0.008	2.49E-11	0.026	-0.045	0.009	2.02E-06	0.023	-0.050	0.006	2.03E-16
rs28929474	14:94844947	C/T	*SERPINA1*	missense	0.018	0.124	0.009	1.39E-45	0.019	0.139	0.011	2.50E-34	0.019	0.130	0.007	1.72E-75
rs41286560	14:101349454	G/T	*RTL1*	missense	0.024	-0.050	0.007	1.17E-11	0.028	-0.033	0.009	2.12E-04	0.026	-0.044	0.006	2.50E-15
rs116858574	15:34520687	T/C	*EMC4*	missense	0.014	0.047	0.010	1.16E-06	0.014	0.028	0.012	2.19E-02	0.014	0.040	0.008	1.60E-07
rs34815962	15:72462255	C/T	*GRAMD2*	missense	0.019	0.073	0.009	8.72E-17	0.023	0.074	0.010	3.66E-13	0.021	0.073	0.007	1.28E-27
rs16942341	15:89388905	C/T	*ACAN*	synonymous	0.026	-0.129	0.007	4.30E-72	0.028	-0.146	0.009	1.08E-56	0.027	-0.135	0.006	3.79E-130
rs61733564	16:4812705	A/G	*ZNF500*	missense	0.032	0.056	0.007	8.61E-17	0.032	0.044	0.009	2.34E-07	0.032	0.051	0.005	2.89E-21
rs113388806	16:24804954	A/T	*TNRC6A*	missense	0.040	0.036	0.006	1.08E-09	0.047	0.041	0.008	1.65E-07	0.043	0.038	0.005	1.90E-15
rs8052655	16:67409180	G/A	*LRRC36*	missense	0.043	-0.054	0.006	1.08E-18	0.043	-0.055	0.008	3.91E-13	0.043	-0.054	0.005	6.40E-31
rs77542162	17:67081278	A/G	*ABCA6*	missense	0.017	0.049	0.010	2.17E-06	0.023	0.051	0.010	5.58E-07	0.020	0.050	0.007	5.57E-12
rs77169818	18:74980601	A/T	*GALR1*	missense	0.047	-0.048	0.006	3.60E-18	0.038	-0.035	0.008	3.64E-05	0.044	-0.044	0.005	5.11E-19
rs3208856	19:45296806	C/T	*CBLC*	missense	0.034	0.036	0.007	1.48E-07	0.034	0.021	0.008	1.19E-02	0.034	0.030	0.005	2.96E-08
rs4252548	19:55879672	C/T	*IL11*	missense	0.026	-0.114	0.007	1.02E-57	0.022	-0.101	0.010	2.28E-23	0.025	-0.110	0.006	5.32E-81
rs147110934	19:55993436	G/T	*ZNF628*	missense	0.021	-0.084	0.010	2.28E-18	0.022	-0.098	0.011	1.17E-18	0.022	-0.090	0.007	6.33E-34
rs77885044	22:28501414	C/T	*TTC28*	missense	0.012	-0.067	0.010	9.47E-11	0.017	-0.069	0.012	3.24E-09	0.014	-0.068	0.008	3.93E-19
rs147348682	22:42095658	T/G	*MEI1*	missense	0.025	0.041	0.007	2.25E-08	0.034	0.024	0.009	6.59E-03	0.029	0.034	0.006	3.70E-10

603

Table 3. Ten height genes implicated by gene-based testing. These genes meet our three criteria for statistical significance: (1) gene-based $P<2x10^{-6}$, (2) the gene does not include variants with $P<2x10^{-7}$, and (3) the gene-based P-value is at least two orders of magnitude smaller than the P-value for the most significant variant within the gene. For each gene, we provide P-values for the four different gene-based tests applied. P-values in bold are the most significant results for a given gene. [1]Replication results using the same test and (when possible) variants in 59,804 European-ancestry individuals. [2]When the gene is located in a locus identified by our single-variant analysis (1 Mb window), we conditioned the gene-based association result by genotypes at the single variant. [3]If the gene falls within a known GWAS height locus, we mention if it was predicted to be causal using bioinformatic tools (ref. [3]). NA, not applicable.

Gene	Discovery gene-based P-value				Replication P-value[1]	Conditional P-value[2]	Note[3]
	SKAT-broad	VT-broad	SKAT-strict	VT-strict			
OSGIN1	**$4.3x10^{-11}$**	$4.5x10^{-5}$	0.19	0.18	**0.048**	$7.7x10^{-11}$	Known locus. No predicted causal genes.
CRISPLD1	$2.2x10^{-7}$	**$1.5x10^{-10}$**	$8.5x10^{-6}$	$8.9x10^{-7}$	0.50	NA	Known locus, sentinel GWAS SNP not tested on ExomeChip. *CRISPLD1* was predicted to be causal.
CSAD	$2.3x10^{-8}$	**$6.0x10^{-10}$**	0.83	0.59	0.54	NA	New locus.
SNED1	$1.9x10^{-5}$	**$2.3x10^{-9}$**	NA	NA	0.083	$1.4x10^{-9}$	Known locus. *SNED1* was not predicted to be causal.
G6PC	$1.3x10^{-5}$	**$3.6x10^{-8}$**	$5.5x10^{-6}$	$1.3x10^{-6}$	0.24	$3.9x10^{-8}$	Known locus, *G6PC* was not predicted to be causal. *G6PC* is mutated in glycogen storage disease Ia.
NOX4	$5.1x10^{-6}$	**$1.8x10^{-7}$**	NA	NA	**0.013**	NA	New locus.
UGGT2	$3.0x10^{-5}$	**$2.0x10^{-7}$**	$2.3x10^{-5}$	$4.8x10^{-7}$	0.64	NA	New locus.
FLNB	$2.2x10^{-6}$	$5.1x10^{-4}$	**$2.4x10^{-9}$**	$3.2x10^{-6}$	**0.016**	$3.6x10^{-9}$	Known locus. *FLNB* was predicted to be causal. *FLNB* is mutated in atelosteogenesis type I.
B4GALNT3	$2.4x10^{-5}$	$1.9x10^{-5}$	$1.8x10^{-5}$	**$3.4x10^{-7}$**	0.79	$7.7x10^{-7}$	Known locus. *B4GALNT3* was predicted to be causal.
CCDC3	$6.3x10^{-4}$	$6.3x10^{-6}$	$3.0x10^{-7}$	**$5.5x10^{-9}$**	0.080	$1.6x10^{-9}$	Known locus. *CCDC3* was predicted to be causal.

Figure legends

Figure 1. Variants with a larger effect size on height variation tend to be rarer. We observe an inverse relationship between the effect size (from the combined "discovery+ validation" analysis, in cm on the *y*-axis) and the minor allele frequency (MAF) for the height variants (*x*-axis, from 0 to 50%). We included in this figure the 606 height variants with $P<2x10^{-7}$.

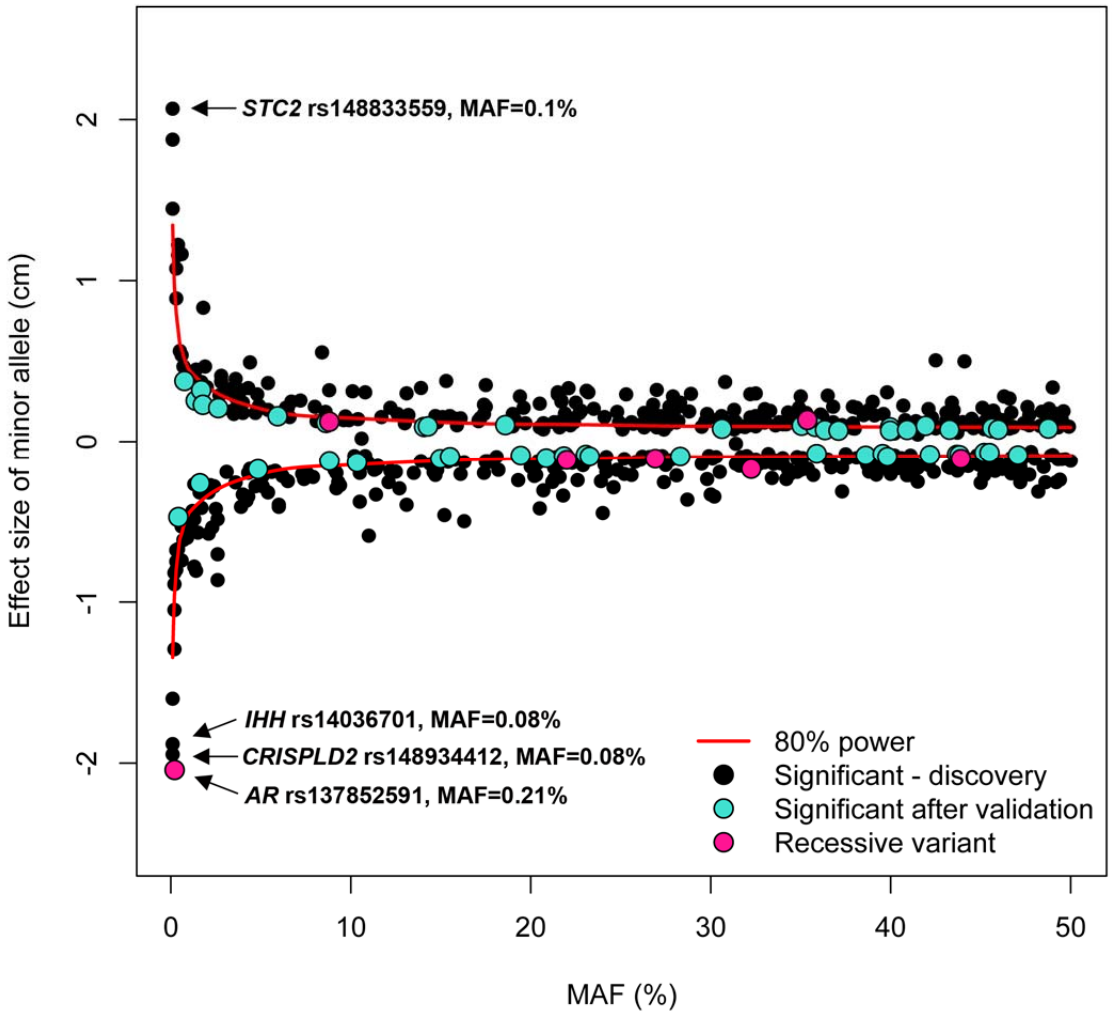


Figure 2. Heat map showing subset of DEPICT gene set enrichment results. The full heat map is available as **Supplementary Fig. 13**. For any given square, the color indicates how strongly the corresponding gene (shown on the x-axis) is predicted to belong to the reconstituted gene set (y-axis). This value is based on the gene's Z-score for gene set inclusion in DEPICT's reconstituted gene sets, where red indicates a higher Z-score and blue indicates a lower one. The proteoglycan binding pathway (bold) was uniquely implicated by coding variants (as opposed to common variants) by both DEPICT and the PASCAL method. To visually reduce redundancy and increase clarity, we chose one representative "meta-gene set" for each group of highly correlated gene sets based on affinity propagation clustering (**Supplementary Note**). Heat map intensity and DEPICT P-values correspond to the most significantly enriched gene set within the meta-gene set; meta-gene sets are listed with their database source. Annotations for the genes indicate whether the gene has OMIM annotation as underlying a disorder of skeletal growth (black and grey) and the minor allele frequency of the significant ExomeChip (EC) variant (shades of blue; if multiple variants, the lowest-frequency variant was kept). Annotations for the gene sets indicate if the gene set was also found significant for EC by the PASCAL method (yellow and grey) and if the gene set was found significant by DEPICT for EC only or for both EC and GWAS (purple and green). Abbreviations: GO: Gene Ontology; MP: mouse phenotype in the Mouse Genetics Initiative; PPI: protein-protein interaction in the InWeb database.

(Figure 2)

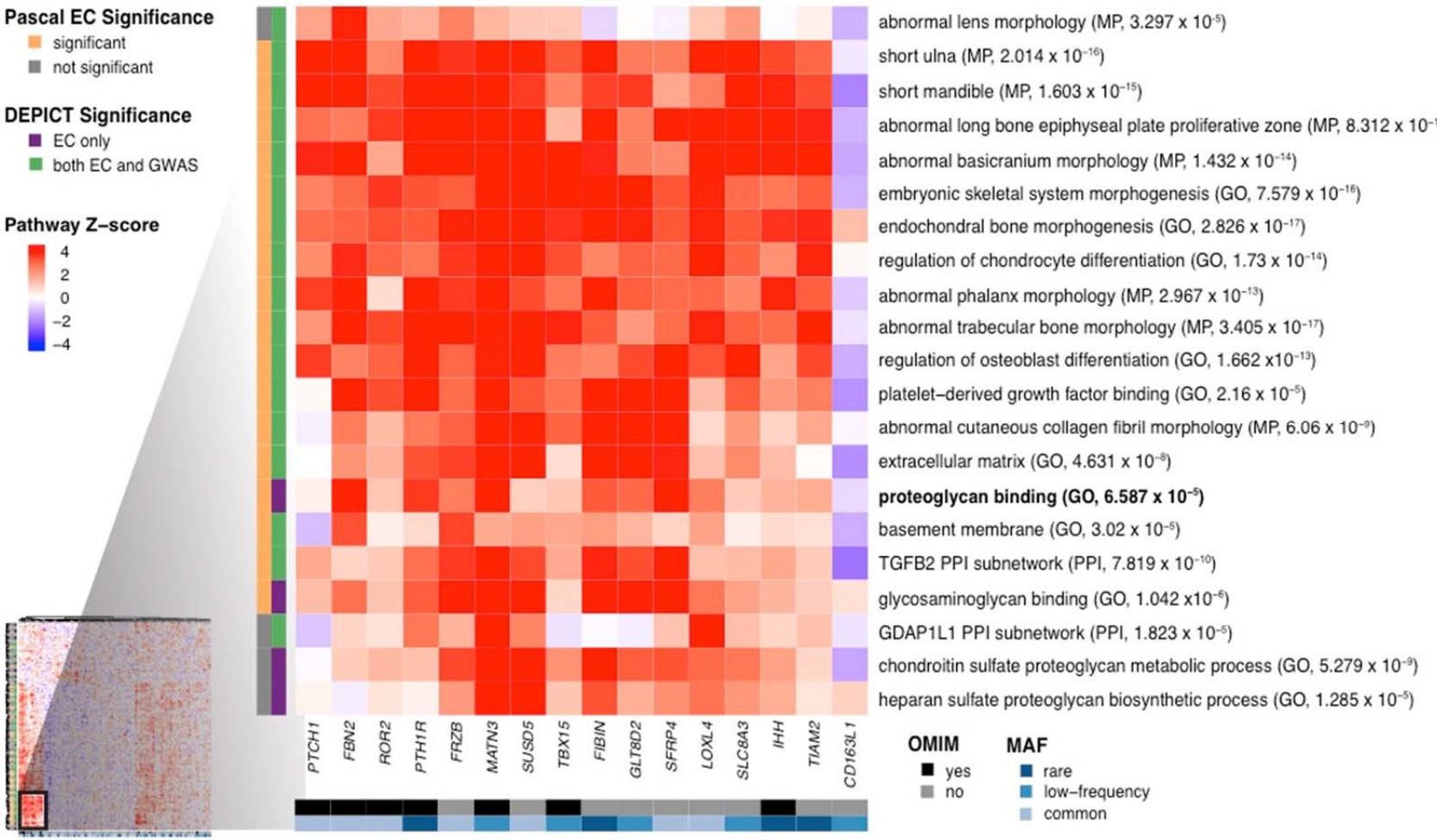
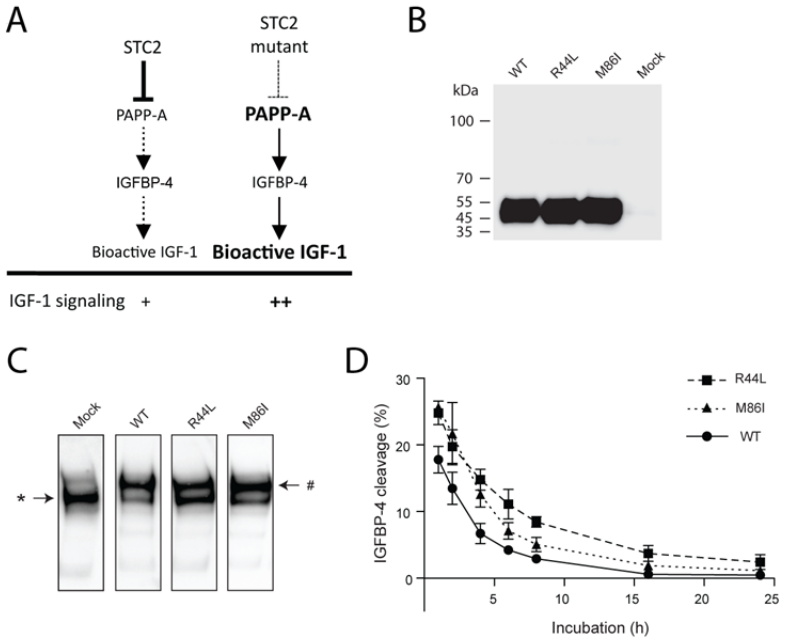

Pascal EC Significance
significant
not significant
DEPICT Significance
EC only
both EC and GWAS
Pathway Z–score
4
2
0
−2
−4
abnormal lens morphology (MP, 3.297 x 10-5)
short ulna (MP, 2.014 x 10-16)
short mandible (MP, 1.603 x 10-15)
abnormal long bone epiphyseal plate proliferative zone (MP, 8.312 x 10-18)
abnormal basicranium morphology (MP, 1.432 x 10-14)
embryonic skeletal system morphogenesis (GO, 7.579 x 10-16)
endochondral bone morphogenesis (GO, 2.826 x 10-17)
regulation of chondrocyte differentiation (GO, 1.73 x 10-14)
abnormal phalanx morphology (MP, 2.967 x 10-13)
abnormal trabecular bone morphology (MP, 3.405 x 10-17)
regulation of osteoblast differentiation (GO, 1.662 x10-13)
platelet–derived growth factor binding (GO, 2.16 x 10-5)
abnormal cutaneous collagen fibril morphology (MP, 6.06 x 10-9)
extracellular matrix (GO, 4.631 x 10-8)
proteoglycan binding (GO, 6.587 x 10-5)
basement membrane (GO, 3.02 x 10-5)
TGFB2 PPI subnetwork (PPI, 7.819 x 10-10)
glycosaminoglycan binding (GO, 1.042 x10-6)
GDAP1L1 PPI subnetwork (PPI, 1.823 x 10-5)
chondroitin sulfate proteoglycan metabolic process (GO, 5.279 x 10-9)
heparan sulfate proteoglycan biosynthetic process (GO, 1.285 x 10-5)
PTCH1
FBN2
ROR2
PTH1R
FRZB
MATN3
SUSD5
TBX15
FIBIN
GLT8D2
SFRP4
LOXL4
SLC8A3
IHH
TIAM2
CD163L1
OMIM
yes
no
MAF
rare
low–frequency
common

Figure 3. STC2 mutants p.Arg44Leu (R44L) and p.Met86Ile (M86I) show compromised proteolytic inhibition of PAPP-A. (**A**) Schematic representation of the role of STC2 in IGF-1 signaling. Partial inactivation of STC2 by height-associated DNA sequence variation could increase bioactive IGF-1 through reduced inhibition of PAPP-A. (**B**) Western blot analysis of recombinant STC2 wild-type and variants R44L and M86I. (**C**) Covalent complex formation between PAPP-A and STC2 wild-type or variants R44L and M86I. Separately synthesized proteins were analyzed by PAPP-A Western blotting following incubation for 8 h. In the absence of STC2 (Mock lane), PAPP-A appears as a single 400 kDa band (*). Following incubation with wild-type STC2, the majority of PAPP-A is present as the approximately 500 kDa covalent PAPP-A:STC2 complex (#), in which PAPP-A is devoid of proteolytic activity towards IGFBP-4. Under similar conditions, incubation with variants R44L or M86I appeared to cause less covalent complex formation with PAPP-A. The gels are representative of at least three independent experiments. (**D**) PAPP-A proteolytic cleavage of IGFBP-4 following incubation with wild-type STC2 or variants for 1-24 h. Wild-type STC2 causes reduction in PAPP-A activity, with complete inhibition of activity following 24 h incubation. Both STC2 variants show increased IGFBP-4 cleavage (*i.e.* less inhibition) for all time points analyzed. Mean and standard deviations of three independent experiments are shown. One-way repeated measures analysis of variance followed by Dunnett's post-test showed significant differences between STC2 wild-type and variants R44L ($P<0.001$) and M86I ($P<0.01$).

665 **(Figure 3)**



666

667 **References**

668 1 Fisher, R. A. The Correlation Between Relatives on the Supposition of Mendelian
669 Inheritance. *Transactions of the Royal Society of Edinburgh* **52**, 399-433 (1918).
670 2 Silventoinen, K. *et al.* Heritability of adult body height: a comparative study of twin
671 cohorts in eight countries. *Twin research : the official journal of the International*
672 *Society for Twin Studies* **6**, 399-408, doi:10.1375/136905203770326402 (2003).
673 3 Wood, A. R. *et al.* Defining the role of common variation in the genomic and
674 biological architecture of adult human height. *Nat Genet* **46**, 1173-1186,
675 doi:10.1038/ng.3097 (2014).
676 4 Flannick, J. *et al.* Loss-of-function mutations in SLC30A8 protect against type 2
677 diabetes. *Nat Genet* **46**, 357-363, doi:10.1038/ng.2915 (2014).
678 5 Steinthorsdottir, V. *et al.* Identification of low-frequency and rare sequence variants
679 associated with elevated or reduced risk of type 2 diabetes. *Nat Genet* **46**, 294-298,
680 doi:10.1038/ng.2882 (2014).
681 6 Gudmundsson, J. *et al.* A study based on whole-genome sequencing yields a rare
682 variant at 8q24 associated with prostate cancer. *Nat Genet* **44**, 1326-1329,
683 doi:10.1038/ng.2437 (2012).
684 7 Sidore, C. *et al.* Genome sequencing elucidates Sardinian genetic architecture and
685 augments association analyses for lipid and blood inflammatory markers. *Nat Genet*
686 **47**, 1272-1281, doi:10.1038/ng.3368 (2015).
687 8 Danjou, F. *et al.* Genome-wide association analyses based on whole-genome
688 sequencing in Sardinia provide insights into regulation of hemoglobin levels. *Nat*
689 *Genet* **47**, 1264-1271, doi:10.1038/ng.3307 (2015).

9 Zuk, O. *et al.* Searching for missing heritability: designing rare variant association studies. *Proc Natl Acad Sci U S A* **111**, E455-464, doi:10.1073/pnas.1322563111 (2014).

10 Yang, J. *et al.* Genetic variance estimation with imputed variants finds negligible missing heritability for human height and body mass index. *Nat Genet* **47**, 1114-1120, doi:10.1038/ng.3390 (2015).

11 Grove, M. L. *et al.* Best practices and joint calling of the HumanExome BeadChip: the CHARGE Consortium. *PloS one* **8**, e68095, doi:10.1371/journal.pone.0068095 (2013).

12 Wu, M. C. *et al.* Rare-variant association testing for sequencing data with the sequence kernel association test. *Am J Hum Genet* **89**, 82-93, doi:10.1016/j.ajhg.2011.05.029 (2011).

13 Pers, T. H. *et al.* Biological interpretation of genome-wide association studies using predicted gene functions. *Nat Commun* **6**, 5890, doi:10.1038/ncomms6890 (2015).

14 Lamparter, D., Marbach, D., Rueedi, R., Kutalik, Z. & Bergmann, S. Fast and Rigorous Computation of Gene and Pathway Scores from SNP-Based Summary Statistics. *PLoS Comput Biol* **12**, e1004714, doi:10.1371/journal.pcbi.1004714 (2016).

15 Schwartz, N. B. & Domowicz, M. Chondrodysplasias due to proteoglycan defects. *Glycobiology* **12**, 57R-68R (2002).

16 Wei, H. S., Wei, H. L., Zhao, F., Zhong, L. P. & Zhan, Y. T. Glycosyltransferase GLT8D2 positively regulates ApoB100 protein expression in hepatocytes. *Int J Mol Sci* **14**, 21435-21446, doi:10.3390/ijms141121435 (2013).

17 Ito, H. *et al.* Molecular cloning and biological activity of a novel lysyl oxidase-related gene expressed in cartilage. *J Biol Chem* **276**, 24023-24029, doi:10.1074/jbc.M100861200 (2001).

18 Wakahara, T. *et al.* Fibin, a novel secreted lateral plate mesoderm signal, is essential for pectoral fin bud initiation in zebrafish. *Dev Biol* **303**, 527-535, doi:10.1016/j.ydbio.2006.11.041 (2007).

19 Kawano, Y. & Kypta, R. Secreted antagonists of the Wnt signalling pathway. *J Cell Sci* **116**, 2627-2634, doi:10.1242/jcs.00623 (2003).

20 Mastaitis, J. *et al.* Loss of SFRP4 Alters Body Size, Food Intake, and Energy Expenditure in Diet-Induced Obese Male Mice. *Endocrinology* **156**, 4502-4510, doi:10.1210/en.2015-1257 (2015).

21 Jepsen, M. R. *et al.* Stanniocalcin-2 inhibits mammalian growth by proteolytic inhibition of the insulin-like growth factor axis. *J Biol Chem* **290**, 3430-3439, doi:10.1074/jbc.M114.611665 (2015).

22 Dauber, A. *et al.* Mutations in pregnancy-associated plasma protein A2 cause short stature due to low IGF-I availability. *EMBO Mol Med*, doi:10.15252/emmm.201506106 (2016).

23 Lango Allen, H. *et al.* Hundreds of variants clustered in genomic loci and biological pathways affect human height. *Nature* **467**, 832-838, doi:nature09410 [pii] 10.1038/nature09410 (2010).

24 Pickrell, J. K. *et al.* Detection and interpretation of shared genetic influences on 42 human traits. *Nat Genet* **48**, 709-717, doi:10.1038/ng.3570 (2016).

25 Park, J. H. *et al.* SLC39A8 Deficiency: A Disorder of Manganese Transport and Glycosylation. *Am J Hum Genet* **97**, 894-903, doi:10.1016/j.ajhg.2015.11.003 (2015).

26 Boycott, K. M. *et al.* Autosomal-Recessive Intellectual Disability with Cerebellar Atrophy Syndrome Caused by Mutation of the Manganese and Zinc Transporter Gene SLC39A8. *Am J Hum Genet* **97**, 886-893, doi:10.1016/j.ajhg.2015.11.002 (2015).

27 Ehret, G. B. *et al.* Genetic variants in novel pathways influence blood pressure and cardiovascular disease risk. *Nature* **478**, 103-109, doi:10.1038/nature10405 (2011).

28 Willer, C. J. *et al.* Discovery and refinement of loci associated with lipid levels. *Nature genetics* **45**, 1274-1283, doi:10.1038/ng.2797 (2013).

29 Speliotes, E. K. *et al.* Association analyses of 249,796 individuals reveal 18 new loci associated with body mass index. *Nature genetics* **42**, 937-948, doi:10.1038/ng.686 (2010).

30 Wain, L. V. *et al.* Genome-wide association study identifies six new loci influencing pulse pressure and mean arterial pressure. *Nature genetics* **43**, 1005-1011, doi:10.1038/ng.922 (2011).

31 Bartlett, J. R. *et al.* Genetic modifiers of liver disease in cystic fibrosis. *Jama* **302**, 1076-1083, doi:10.1001/jama.2009.1295 (2009).

32 North, T. L. *et al.* A study of common Mendelian disease carriers across ageing British cohorts: meta-analyses reveal heterozygosity for alpha 1-antitrypsin deficiency increases respiratory capacity and height. *J Med Genet* **53**, 280-288, doi:10.1136/jmedgenet-2015-103342 (2016).

33 Karaplis, A. C. *et al.* Inactivating mutation in the human parathyroid hormone receptor type 1 gene in Blomstrand chondrodysplasia. *Endocrinology* **139**, 5255-5258, doi:10.1210/endo.139.12.6522 (1998).

34 Sims, N. A. *et al.* Interleukin-11 receptor signaling is required for normal bone remodeling. *J Bone Miner Res* **20**, 1093-1102, doi:10.1359/JBMR.050209 (2005).

35 Takeuchi, Y. *et al.* Interleukin-11 as a stimulatory factor for bone formation prevents bone loss with advancing age in mice. *J Biol Chem* **277**, 49011-49018, doi:10.1074/jbc.M207804200 (2002).

36 Goldstein, J. I. *et al.* zCall: a rare variant caller for array-based genotyping: genetics and population analysis. *Bioinformatics* **28**, 2543-2545, doi:10.1093/bioinformatics/bts479 (2012).

37 Liu, D. J. *et al.* Meta-analysis of gene-level tests for rare variant association. *Nat Genet* **46**, 200-204, doi:10.1038/ng.2852 (2014).

38 Winkler, T. W. & Day, F. R. Quality control and conduct of genome-wide association meta-analyses. **9**, 1192-1212, doi:10.1038/nprot.2014.071 (2014).

39 Purcell, S. M. *et al.* A polygenic burden of rare disruptive mutations in schizophrenia. *Nature* **506**, 185-190, doi:10.1038/nature12975 (2014).

40 Price, A. L. *et al.* Pooled association tests for rare variants in exon-resequencing studies. *Am J Hum Genet* **86**, 832-838, doi:10.1016/j.ajhg.2010.04.005 (2010).

41 Yang, J. *et al.* Genomic inflation factors under polygenic inheritance. *European Journal of Human Genetics* **19**, 807-812, doi:10.1038/ejhg.2011.39 (2011).

42 Bulik-Sullivan, B. K. *et al.* LD Score regression distinguishes confounding from polygenicity in genome-wide association studies. *Nat Genet* **47**, 291-295, doi:10.1038/ng.3211 (2015).

43 Feng, S., Liu, D., Zhan, X., Wing, M. K. & Abecasis, G. R. RAREMETAL: fast and powerful meta-analysis for rare variants. *Bioinformatics (Oxford, England)* **30**, 2828-2829, doi:10.1093/bioinformatics/btu367 (2014).

44 Yang, J. *et al.* Conditional and joint multiple-SNP analysis of GWAS summary statistics identifies additional variants influencing complex traits. *Nature genetics* **44**, 369-S363, doi:10.1038/ng.2213 (2012).

45 Loh, P. R. *et al.* Efficient Bayesian mixed-model analysis increases association power in large cohorts. *Nat Genet* **47**, 284-290, doi:10.1038/ng.3190 (2015).

46 Pasaniuc, B. *et al.* Fast and accurate imputation of summary statistics enhances evidence of functional enrichment. *Bioinformatics* **30**, 2906-2914, doi:10.1093/bioinformatics/btu416 (2014).

47 Moayyeri, A., Hammond, C. J., Valdes, A. M. & Spector, T. D. Cohort Profile: TwinsUK and healthy ageing twin study. *Int J Epidemiol* **42**, 76-85, doi:10.1093/ije/dyr207 (2013).

48 Boyd, A. *et al.* Cohort Profile: the 'children of the 90s'--the index offspring of the Avon Longitudinal Study of Parents and Children. *Int J Epidemiol* **42**, 111-127, doi:10.1093/ije/dys064 (2013).

49 Fehrmann, R. S. *et al.* Gene expression analysis identifies global gene dosage sensitivity in cancer. *Nat Genet* **47**, 115-125, doi:10.1038/ng.3173 (2015).

50 Frey, B. J. & Dueck, D. Clustering by passing messages between data points. *Science* **315**, 972-976, doi:10.1126/science.1136800 (2007).

51 Willer, C. J., Li, Y. & Abecasis, G. R. METAL: Fast and efficient meta-analysis of genomewide association scans. *Bioinformatics* **26**, 2190-2191, doi:btq340 [pii] 10.1093/bioinformatics/btq340 (2010).

52 Overgaard, M. T. *et al.* Expression of recombinant human pregnancy-associated plasma protein-A and identification of the proform of eosinophil major basic protein as its physiological inhibitor. *The Journal of biological chemistry* **275**, 31128-31133, doi:10.1074/jbc.M001384200 (2000).

53 Gyrup, C. & Oxvig, C. Quantitative analysis of insulin-like growth factor-modulated proteolysis of insulin-like growth factor binding protein-4 and -5 by pregnancy-associated plasma protein-A. *Biochemistry* **46**, 1972-1980, doi:10.1021/bi062229i (2007).

54 Oxvig, C., Sand, O., Kristensen, T., Kristensen, L. & Sottrup-Jensen, L. Isolation and characterization of circulating complex between human pregnancy-associated plasma protein-A and proform of eosinophil major basic protein. *Biochimica et biophysica acta* **1201**, 415-423 (1994).

55 Nikpay, M. *et al.* A comprehensive 1,000 Genomes-based genome-wide association meta-analysis of coronary artery disease. *Nat Genet* **47**, 1121-1130, doi:10.1038/ng.3396 (2015).

Authors

Eirini Marouli[1]*, Mariaelisa Graff[2]*, Carolina Medina-Gomez[3, 4]*, Ken Sin Lo[5]*, Andrew R Wood[6]*, Troels R Kjaer[7]*, Rebecca S Fine[8-10]*, Yingchang Lu[11-13]*, Claudia Schurmann[12, 13], Heather M Highland[2, 14], Sina Rüeger[15, 16], Gudmar Thorleifsson[17], Anne E Justice[2], David Lamparter[16, 18], Kathleen E Stirrups[1, 19], Valérie Turcot[5], Kristin L Young[2], Thomas W Winkler[20], Tõnu Esko[8, 10, 21], Tugce Karaderi[22], Adam E Locke[23, 24], Nicholas GD Masca[25, 26], Maggie CY Ng[27, 28], Poorva Mudgal[27], Manuel A Rivas[8, 29], Sailaja Vedantam[8-10], Anubha Mahajan[22], Katja K Aben[30, 31], Linda S Adair[32], Dewan S Alam[33], Eva Albrecht[34], Kristine H Allin[35], Matthew Allison[36], Philippe Amouyel[37-39], Emil V Appel[35], Dominique Arveiler[40, 41], Folkert W Asselbergs[42-44], Paul L Auer[45], Beverley Balkau[46], Bernhard Banas[47], Lia E Bang[48], Marianne Benn[49, 50], Sven Bergmann[16, 18], Lawrence F Bielak[51], Matthias Blüher[52, 53], Heiner Boeing[54], Eric Boerwinkle[55, 56], Carsten A Böger[20, 47], Lori L Bonnycastle[57], Jette Bork-Jensen[35], Michiel L Bots[58], Erwin P Bottinger[12], Donald W Bowden[27, 28, 59], Ivan Brandslund[60, 61], Gerome Breen[62], Murray H Brilliant[63], Linda Broer[4], Amber A Burt[64], Adam S Butterworth[65, 66], David J Carey[67], Mark J Caulfield[1, 68], John C Chambers[69-71], Daniel I Chasman[72-74], Yii-Der Ida Chen[75], Rajiv Chowdhury[65], Cramer Christensen[76], Audrey Y Chu[73, 77], Massimiliano Cocca[78], Francis S Collins[57], James P Cook[79], Janie Corley[80, 81], Jordi Corominas Galbany[82], Amanda J Cox[27, 28, 83], Gabriel Cuellar-Partida[84, 85], John Danesh[65, 66, 86], Gail Davies[80, 81], Paul IW de Bakker[87, 88], Gert Jan de Borst[89], Simon de Denus[5, 90], Mark CH de Groot[91, 92], Renée de Mutsert[93], Ian J Deary[80, 81], George Dedoussis[94], Ellen W Demerath[95], Anneke I den Hollander[96], Joe G Dennis[97], Emanuele Di Angelantonio[65, 66], Fotios Drenos[98, 99], Mengmeng Du[100, 101], Alison M Dunning[102], Douglas F Easton[97, 102], Tapani Ebeling[103, 104], Todd L Edwards[105], Patrick T Ellinor[106, 107], Paul Elliott[108], Evangelos Evangelou[70, 109], Aliki-Eleni Farmaki[94], Jessica D Faul[110], Mary F Feitosa[111], Shuang Feng[112], Ele Ferrannini[113, 114], Marco M Ferrario[115], Jean Ferrieres[116], Ian Ford[117], Myriam Fornage[118], Paul W Franks[119-121], Ruth Frikke-Schmidt[50, 122], Tessel E Galesloot[31], Wei Gan[22], Ilaria Gandin[123], Paolo Gasparini[123, 124], Vilmantas Giedraitis[125], Ayush Giri[105], Giorgia Girotto[123, 124], Scott D Gordon[85], Penny Gordon-Larsen[126, 127], Mathias Gorski[20, 47], Niels Grarup[35], Megan L. Grove[55], Vilmundur Gudnason[128, 129], Xiuqing Guo[75], Stefan Gustafsson[130], Torben Hansen[35], Kathleen M Harris[126, 131], Tamara B Harris[132], Andrew T Hattersley[133], Caroline Hayward[134], Liang He[135, 136], Iris M Heid[20, 34], Kauko Heikkilä[136, 137], Øyvind Helgeland[138, 139], Jussi Hernesniemi[140-142], Alex W Hewitt[143-145], Lynne J Hocking[146, 147], Mette Hollensted[35], Oddgeir L Holmen[148], G. Kees Hovingh[149], Joanna MM Howson[65], Carel B Hoyng[96], Paul L Huang[106], Kristian Hveem[150], M. Arfan Ikram[3, 151, 152], Erik Ingelsson[130, 153], Anne U Jackson[112], Jan-Håkan Jansson[154, 155], Gail P Jarvik[64, 156], Gorm B Jensen[157], Min A Jhun[51], Yucheng Jia [75], Xuejuan Jiang[158, 159], Stefan Johansson[139, 160], Marit E Jørgensen[161, 162], Torben Jørgensen[50, 163, 164], Pekka Jousilahti[165], J Wouter Jukema[166, 167], Bratati Kahali[168-170], René S Kahn[171], Mika Kähönen[172], Pia R Kamstrup[49, 50], Stavroula Kanoni[1], Jaakko Kaprio[136, 137, 173], Maria Karaleftheri[174], Sharon LR Kardia[51], Fredrik Karpe[175, 176], Frank Kee[177], Renske Keeman[178], Lambertus A Kiemeney[31], Hidetoshi Kitajima[22], Kirsten B Kluivers[31], Thomas Kocher[179], Pirjo Komulainen[180], Jukka Kontto[173], Jaspal S Kooner[69, 71, 181], Charles Kooperberg[182], Peter Kovacs[52], Jennifer Kriebel[183-185], Helena Kuivaniemi[67, 186], Sébastien Küry [187], Johanna Kuusisto[188], Martina La Bianca[189], Markku Laakso[188], Timo A Lakka[180, 190], Ethan M Lange[191], Leslie A Lange[191], Carl D Langefeld [192], Claudia Langenberg[193], Eric B Larson[64, 194, 195], I-Te Lee[196,

[197], Terho Lehtimäki[141, 142], Cora E Lewis[198], Huaixing Li[199], Jin Li[200], Ruifang Li-Gao[93], Honghuang Lin[201], Li-An Lin[118], Xu Lin[199], Lars Lind[202], Jaana Lindström[203], Allan Linneberg[203], Yeheng Liu[75], Yongmei Liu[204], Artitaya Lophatananon[205], Jian'an Luan[193], Steven A Lubitz[106, 107], Leo-Pekka Lyytikäinen[141, 142], David A Mackey[144], Pamela AF Madden[206], Alisa K Manning[106, 107, 207], Satu Männistö[173], Gaëlle Marenne[86], Jonathan Marten[134], Nicholas G Martin[85], Angela L Mazul[2], Karina Meidtner[183, 208], Andres Metspalu[21], Paul Mitchell[209], Karen L Mohlke[191], Dennis O Mook-Kanamori[93, 210], Anna Morgan[123], Andrew D Morris[211], Andrew P Morris[22, 79], Martina Müller-Nurasyid[34, 212, 213], Patricia B Munroe[1, 68], Mike A Nalls[214], Matthias Nauck[215, 216], Christopher P Nelson[25, 26], Matt Neville[175, 176], Sune F Nielsen[49, 50], Kjell Nikus[217], Pål R Njølstad[138, 139], Børge G Nordestgaard[49, 50], Ioanna Ntalla[1], Jeffrey R O'Connel[218], Heikki Oksa[219], Loes M Olde Loohuis[220], Roel A Ophoff[171, 220], Katharine R Owen[175, 176], Chris J Packard[117], Sandosh Padmanabhan[117], Colin NA Palmer[221], Gerard Pasterkamp[222, 223], Aniruddh P Patel[8, 74, 106], Alison Pattie[81], Oluf Pedersen[35], Peggy L Peissig[63], Craig E Pennell[224], Markus Perola[173, 225, 226], James A Perry[218], Thomas N Person[63], Ailith Pirie[102], Ozren Polasek[211, 227], Danielle Posthuma[228, 229], Olli T Raitakari[230, 231], Asif Rasheed[232], Rainer Rauramaa[180, 233], Dermot F Reilly[234], Alex P Reiner[182, 235], Frida Renström[119, 236], Paul M Ridker[73, 74, 237], John D Rioux[5, 238], Neil Robertson[22, 175], Antonietta Robino[189], Olov Rolandsson[154, 239], Igor Rudan[211], Danish Saleheen[232, 240], Veikko Salomaa[173], Nilesh J Samani[25, 26], Kevin Sandow[75], Yadav Sapkota[85], Naveed Sattar[117], Marjanka K Schmidt[178], Pamela J Schreiner[241], Matthias B Schulze[183, 208], Robert A Scott[193], Marcelo P Segura-Lepe[70], Svati Shah[242], Xueling Sim[112, 243], Suthesh Sivapalaratnam[106, 244, 245], Kerrin S Small[246], Albert Vernon Smith[128, 129], Jennifer A Smith[51], Lorraine Southam[22, 86], Timothy D Spector[246], Elizabeth K Speliotes[168-170], John M Starr[80, 247], Valgerdur Steinthorsdottir[17], Heather Stringham[112], Michael Stumvoll[52, 53], Praveen Surendran[65], Leen M t Hart[248-250], Katherine E Tansey[251, 252], Jean-Claude Tardif[5, 238], Kent D Taylor[75, 253], Alexander Teumer[254], Deborah J Thompson[97], Unnur Thorsteinsdottir[17, 128], Betina H Thuesen[164], Anke Tönjes[255], Gerard Tromp[67, 256], Stella Trompet[166, 257], Emmanouil Tsafantakis[258], Jaakko Tuomilehto[203, 259-261], Anne Tybjaerg-Hansen[50, 122], Jonathan P Tyrer[102], Rudolf Uher[262], André G Uitterlinden, Sheila Ulivi[3, 4], Sander W van der Laan[223], Andries R Van Der Leij[263], Cornelia M van Duijn[3], Natasja M van Schoor[264], Jessica van Setten[42], Anette Varbo[49, 50], Tibor V Varga[119], Rohit Varma[159], Digna R Velez Edwards[265], Sita H Vermeulen[31], Henrik Vestergaard[35], Veronique Vitart[134], Thomas F Vogt[266], Diego Vozzi[124], Mark Walker[267], Feijie Wang[199], Carol A Wang[224], Shuai Wang[268], Yiqin Wang[199], Nicholas J Wareham[193], Helen R Warren[1, 68], Jennifer Wessel[269], Sara M Willems[193], James G Wilson[270], Daniel R Witte[271, 272], Michael O Woods[273], Ying Wu[191], Hanieh Yaghootkar[274], Jie Yao[75], Pang Yao[199], Laura M Yerges-Armstrong[218, 275], Robin Young[65, 117], Eleftheria Zeggini[86], Xiaowei Zhan[276], Weihua Zhang[69, 70], Jing Hua Zhao[193], Wei Zhao[240], Wei Zhao[51], He Zheng[199], Wei Zhou[168, 169], EPIC-CVD Consortium¶, The EPIC-InterAct consortium¶, CHD Exome+ Consortium¶, MAGIC Consortium¶, ExomeBP Consortium¶, T2D-Genes Consortium¶, GoT2D Genes Consortium¶, Jerome I Rotter[75], Michael Boehnke[112], Sekar Kathiresan[8, 74, 106], Mark I McCarthy[22, 175, 176], Cristen J Willer[168, 169, 277], Kari Stefansson[17, 128], Ingrid B Borecki[111], Dajiang J Liu[278], Kari E North[279], Nancy L Heard-Costa[77, 280], Tune H Pers[35, 281], Cecilia M Lindgren[22, 282], Claus Oxvig[7]*, Zoltán Kutalik[15, 16]*, Fernando Rivadeneira[3, 4]*, Ruth JF Loos[12, 13, 283]*, Timothy M Frayling[274]*, Joel N Hirschhorn[8, 10, 284]*, Panos Deloukas[1, 285]*, Guillaume Lettre[5, 238]*

*These authors contributed equally to this work.
¶A full list of members and affiliations appears in the **Supplementary Note**

Correspondence should be addressed to JNH (joelh@broadinstitute.org), PD (p.deloukas@qmul.ac.uk) or GL (guillaume.lettre@umontreal.ca).

Affiliations

1 William Harvey Research Institute, Barts and The London School of Medicine and Dentistry, Queen Mary University of London, London, EC1M 6BQ, UK
2 Department of Epidemiology, University of North Carolina, Chapel Hill, NC, 27514, USA
3 Department of Epidemiology, Erasmus Medical Center, Rotterdam, 3015 GE, The Netherlands
4 Department of Internal Medicine, Erasmus Medical Center, Rotterdam, 3015 GE, The Netherlands
5 Montreal Heart Institute, Université de Montréal, Quebec, H1T 1C8, Canada
6 Genetics of Complex Traits, Institute of Biomedical and Clinical Science, University of Exeter Medical School, Exeter, EX2 5DW, UK
7 Department of Molecular Biology and Genetics, Aarhus University, Aarhus, 8000, Denmark
8 Broad Institute of MIT and Harvard, Cambridge, MA, 02142, USA
9 Department of Genetics, Harvard Medical School, Boston, MA, 02115, USA
10 Division of Endocrinology and Center for Basic and Translational Obesity Research, Boston Childrenís Hospital, Boston, MA, 02115, USA
11 Division of Epidemiology, Department of Medicine, Vanderbilt-Ingram Cancer Center, Vanderbilt Epidemiology Center, Vanderbilt University School of Medicine, Nashville, TN, 37203, USA
12 The Charles Bronfman Institute for Personalized Medicine, Icahn School of Medicine at Mount Sinai, New York, NY, 10029, USA
13 The Genetics of Obesity and Related Metabolic Traits Program, Ichan School of Medicine at Mount Sinai, New York, NY, 10069, USA
14 Human Genetics Center, The University of Texas School of Public Health, The University of Texas Graduate School of Biomedical Sciences at Houston, The University of Texas Health Science Center at Houston, Houston, TX, 77030, USA
15 Institute of Social and Preventive Medicine, Lausanne University Hospital, Lausanne, 1010, Switzerland
16 Swiss Institute of Bioinformatics, Lausanne, 1015, Switzerland
17 deCODE Genetics/Amgen inc., Reykjavik, 101, Iceland
18 Department of Computational Biology, University of Lausanne, Lausanne, 1011, Switzerland
19 Department of Haematology, University of Cambridge, Cambridge, CB2 0PT, UK
20 Department of Genetic Epidemiology, University of Regensburg, Regensburg, D-93051, Germany
21 Estonian Genome Center, University of Tartu, Tartu, 51010, Estonia
22 Wellcome Trust Centre for Human Genetics, University of Oxford, Oxford, OX3 7BN, UK
23 Center for Statistical Genetics and Department of Biostatistics, University of Michigan, Ann Arbor, MI, 43109, USA
24 McDonnell Genome Institute, Washington University School of Medicine, Saint Louis, MO, 63108, USA
25 Department of Cardiovascular Sciences, Univeristy of Leicester, Glenfield Hospital, Leicester, LE3 9QP, UK
26 NIHR Leicester Cardiovascular Biomedical Research Unit, Glenfield Hospital, Leicester, LE3 9QP, UK
27 Center for Diabetes Research, Wake Forest School of Medicine, Winston-Salem, NC, 27157, USA
28 Center for Genomics and Personalized Medicine Research, Wake Forest School of Medicine, Winston-Salem, NC, 27157, USA
29 Nuffield Department of Clinical Medicine, Oxford, OX37BN, UK
30 Netherlands Comprehensive Cancer Organisation, Utrecht, 3501 DB, The Netherlands
31 Radboud Institute for Health Sciences, Radboud university medical center, Nijmegen, 6500 HB, The Netherlands
32 Department of Nutrition, University of North Carolina, Chapel Hill, NC, 27599, USA
33 Centre for Control of Chronic Diseases (CCCD), Dhaka, 1212, Bangladesh
34 Institute of Genetic Epidemiology, Helmholtz Zentrum München - German Research Center for Environmental Health, Neuherberg, D-85764, Germany
35 The Novo Nordisk Foundation Center for Basic Metabolic Research, Faculty of Health and Medical Sciences, University of Copenhagen, Copenhagen, 2100, Denmark
36 Department of Family Medicine & Public Health, University of California, San Diego, La Jolla, CA, 92093, USA
37 INSERM U1167, Lille, F-59019, France

38 Institut Pasteur de Lille, U1167, Lille, F-59019, France
39 Universite de Lille, U1167 - RID-AGE - Risk factors and molecular determinants of aging-related diseases, Lille, F-59019, France
40 Department of Epidemiology and Public Health, University of Strasbourg, Strasbourg, F-67085, France
41 Department of Public Health, University Hospital of Strasbourg, Strasbourg, 67081, France
42 Department of Cardiology, Division Heart & Lungs, University Medical Center Utrecht, Utrecht, The Netherlands
43 Durrer Center for Cardiogenetic Research, ICIN-Netherlands Heart Institute, Utrecht, The Netherlands
44 Institute of Cardiovascular Science, Faculty of Population Health Sciences, University College London, London, UK
45 Zilber School of Public Health, University of Wisconsin-Milwaukee, Milwaukee, WI, 53201, USA
46 INSERM U1018, Centre de recherche en Épidemiologie et Sante des Populations (CESP), Villejuif, France
47 Department of Nephrology, University Hospital Regensburg, Regensburg, 93042, Germany
48 Department of Cardiology, Rigshospitalet, Copenhagen University Hospital, Copenhagen, 2100, Denmark
49 Department of Clinical Biochemistry, Herlev and Gentofte Hospital, Copenhagen University Hospital, Herlev, 2730, Denmark
50 Faculty of Health and Medical Sciences, University of Copenhagen, Copenhagen, 2200, Denmark
51 Department of Epidemiology, School of Public Health, University of Michigan, Ann Arbor, MI, 48109, USA
52 IFB Adiposity Diseases, University of Leipzig, Leipzig, 04103, Germany
53 University of Leipzig, Department of Medicine, Leipzig, 04103, Germany
54 Department of Epidemiology, German Institute of Human Nutrition Potsdam-Rehbruecke (DIfE), Nuthetal, 14558, Germany
55 School of Public Health, Human Genetics Center, The University of Texas Health Science Center at Houston, Houston, TX, 77030, USA
56 Human Genome Sequencing Center, Baylor College of Medicine, Houston, TX, 77030 USA
57 Medical Genomics and Metabolic Genetics Branch, National Human Genome Research Institute, National Institutes of Health, Bethesda, MD, 20892, USA
58 Julius Center for Health Sciences and Primary Care, University Medical Center Utrecht, Utrecht, The Netherlands
59 Department of Biochemistry, Wake Forest School of Medicine, Winston-Salem, NC 27157, USA
60 Department of Clinical Biochemistry, Lillebaelt Hospital, Vejle, 7100, Denmark
61 Institute of Regional Health Research, University of Southern Denmark, Odense, 5000, Denmark
62 MRC Social Genetic and Developmental Psychiatry Centre, Institute of Psychiatry, Psychology and Neuroscience, King's College London & NIHR Biomedical Research Centre for Mental Health at the Maudsley, London, SE5 8AF, UK
63 Marshfield Clinic Research Foundation, Marshfield, WI, 54449, USA
64 Department of Medicine, University of Washington, Seattle, WA, 98195, USA
65 Cardiovascular Epidemiology Unit, Department of Public Health and Primary Care, University of Cambridge, Cambridge, CB1 8RN, UK
66 The National Institute for Health Research Blood and Transplant Unit (NIHR BTRU) in Donor Health and Genomics at the University of Cambridge, University of Cambridge, Cambridge, CB1 8RN, UK
67 The Sigfried and Janet Weis Center for Research, Danville, PA, 17822, USA
68 NIHR Barts Cardiovascular Research Unit, Barts and The London School of Medicine & Dentistry, Queen Mary University, London, EC1M 6BQ, UK
69 Department of Cardiology, London North West Healthcare NHS Trust, Ealing Hospital, Middlesex, UB1 3HW, UK
70 Department of Epidemiology and Biostatistics, School of Public Health, Imperial College London, London, W2 1PG, UK
71 Imperial College Healthcare NHS Trust, London, W12 0HS, UK
72 Division of Genetics, Brigham and Women's Hospital and Harvard Medical School, Boston, MA, 02115, USA
73 Division of Preventive Medicine, Brigham and Women's and Harvard Medical School, Boston, MA, 02215, USA
74 Harvard Medical School, Boston, MA, 02215, USA
75 Institute for Translational Genomics and Population Sciences, LABioMed at Harbor-UCLA Medical Center, Torrance, CA, 90502, USA
76 Medical department, Lillebaelt Hospital, Vejle, 7100, Denmark
77 NHLBI Framingham Heart Study, Framingham, MA, 01702, USA
78 Department of Medical, Surgical and Health Sciences, University of Trieste, Trieste, 34100, Italy
79 Department of Biostatistics, University of Liverpool, Liverpool, L69 3GL, UK
80 Centre for Cognitive Ageing and Cognitive Epidemiology, University of Edinburgh, Edinburgh, EH8 9JZ, UK
81 Department of Psychology, University of Edinburgh, Edinburgh, EH8 9JZ, UK
82 Department of Human Genetics, Radboud University Medical Center, Nijmegen, 6500 HB, The Netherlands

83 Molecular Basis of Disease, Griffith University, Southport, QLD, Australia
84 Diamantina Institute, University of Qeensland, Brisbane, Queensland, 4072, Australia
85 QIMR Berghofer Medical Research Institute, Brisbane, Queensland, 4006, Australia
86 Wellcome Trust Sanger Institute, Hinxton, CB10 1SA, UK
87 Department of Epidemiology, Julius Center for Health Sciences and Primary Care, University Medical Center Utrecht, Utrecht, 3584 CG, The Netherlands
88 Department of Genetics, Center for Molecular Medicine, University Medical Center Utrecht, Utrecht, 3584 CX, The Netherlands
89 Department of Surgery, Division of Surgical Specialties, University Medical Center Utrecht, Utrecht, 3584 CX, The Netherlands
90 Faculty of Pharmacy, Université de Montréal, Montreal, Quebec, H3T 1J4, Canada
91 Department of Clinical Chemistry and Haematology, Division of Laboratory and Pharmacy, University Medical Center Utrecht, Utrecht, 3508 GA, The Netherlands
92 Utrecht Institute for Pharmaceutical Sciences, Dvision Pharmacoepidemiology & Clinical Pharmacology, Utrecht University, Utrecht, 3508 TB, The Netherlands
93 Department of Clinical Epidemiology, Leiden University Medical Center, Leiden, 2300RC, The Netherlands
94 Department of Nutrition and Dietetics, School of Health Science and Education, Harokopio University, Athens, 17671, Greece
95 Division of Epidemiology & Community Health, School of Public Health, University of Minnesota, Minneapolis, MN, 55454, USA
96 Department of Ophthalmology, Radboud University Medical Center, Nijmegen, 6500 HB, The Netherlands
97 Centre for Cancer Genetic Epidemiology, Department of Public Health and Primary Care, University of Cambridge, Cambridge, CB1 8RN, UK
98 Institute of Cardiovascular Science, University College London, London, WC1E 6JF, UK
99 MRC Integrative Epidemiology Unit, School of Social & Community Medicine, University of Bristo, Bristol, BS8 2BN, UK
100 Fred Hutchinson Cancer Research Center, Public Health Sciences Division, Seattle, WA, 98109, USA
101 Memorial Sloan Kettering Cancer Center, Department of Epidemiology and Biostatistics, New York, NY, 10017, USA
102 Centre for Cancer Genetic Epidemiology, Department of Oncology, University of Cambridge, Cambridge, CB1 8RN, UK
103 Department of Medicine, Oulu University Hospital, Oulu, 90029, Finland
104 Research Unit of Internal Medicine, University of Oulu, Oulu, FI-90014, Finland
105 Division of Epidemiology, Department of Medicine, Institute for Medicine and Public Health, Vanderbilt Genetics Institute, Vanderbilt University, Nashville, TN, 37203, USA
106 Massachusetts General Hospital, Boston, MA, 02114, USA
107 Medical and Population Genetics Program, Broad Institute, Cambridge, MA, 02141, USA
108 Department of Epidemiology and Biostatistics, MRC-PHE Centre for Environment and Health, Schoold of Public Health, Imperial College London, London, W2 1PG, Uk
109 Department of Hygiene and Epidemiology, University of Ioannina Medical School, Ioannina, 45110, Greece
110 Survey Research Center, Institute for Social Research, University of Michigan, Ann Arbor, MI, 48104, USA
111 Division of Statistical Genomics, Department of Genetics, Washington University School of Medicine, St. Louis, MO, 63108, USA
112 Department of Biostatistics and Center for Statistical Genetics, University of Michigan, Ann Arbor, MI, 48109, USA
113 CNR Institute of Clinical Physiology, Pisa, Italy
114 Department of Clinical & Experimental Medicine, University of Pisa, Italy
115 Research center on Epidemiology and Preventive Medicine, dept. Of clinical and experimental medicine, University of Insubria, Varese, 21100, Italy
116 Toulouse University School of Medicine, Toulouse, TSA 50032 31059, France
117 University of Glasgow, Glasgow, G12 8QQ, UK
118 Institute of Molecular Medicine University of Texas Health Science Center at Houston, Houston, TX, 77030, USA
119 Department of Clinical Sciences, Genetic and Molecular Epidemiology Unit, Lund University, Malmö, SE-20502, Sweden
120 Department of Nutrition, Harvard School of Public Health, Boston, MA, 02115, USA
121 Department of Public Health and Clinical Medicine, Unit of Medicine, Umeå University, Umeå, 901 87, Sweden
122 Department of Clinical Biochemistry, Rigshospitalet, Copenhagen University Hospital, Copenhagen, 2100, Denmark
123 Department of Medical Sciences, University of Trieste, Trieste, 34137, Italy
124 Division of Experimental Genetics, Sidra Medical and Research Center, Doha, 26999, Qatar

Geriatrics, Department of Public Health, Uppsala University, Uppsala, 751 85, Sweden
Carolina Population Center, University of North Carolina, Chapel Hill, NC, 27514, USA
Department of Nutrition, Gillings School of Global Public Health, University of North Carolina, Chapel Hill, NC, 27514, USA
Faculty of Medicine, University of Iceland, Reykjavik, 101, Iceland
Icelandic Heart Association, Kopavogur, 201, Iceland
Department of Medical Sciences, Molecular Epidemiology and Science for Life Laboratory, Uppsala University, Uppsala, 751 41, Sweden
Department of Sociology, University of North Carolina, Chapel Hill, NC, 27514, USA
Laboratory of Epidemiology and Population Sciences, National Institute on Aging, Intramural Research Program, National Institutes of Health, Bethesda, MD, 20892, USA
University of Exeter Medical School, University of Exeter, Exeter, EX2 5DW, UK
MRCHGU, Institute of Genetics and Molecular Medicine, University of Edinburgh, Edinburgh, EH4 2XU, UK
Biodemography of Aging Research Unit, Social Science Research Institute, Duke University, Durham, NC, 27708, USA
Department of Public Health, University of Helsinki, Helsinki, FI-00014, Finland
Institute for Molecular Medicine Finland (FIMM), University of Helsinki, Helsinki, FI-00014, Finland
Department of Pediatrics, Haukeland University Hospital, Bergen, 5021, Norway
KG Jebsen Center for Diabetes Research, Department of Clinical Science, University of Bergen, Bergen, 5020, Norway
Department of Cardiology, Heart Hospital, Tampere University Hospital, Tampere, 33521, Finland
Department of Clinical Chemistry, Fimlab Laboratories, Tampere, 33520, Finland
Department of Clinical Chemistry, University of Tampere School of Medicine, Tampere, 33014, Finland
Centre for Eye Research Australia, Royal Victorian Eye and Ear Hospital, University of Melbourne, Melbourne, Victoria, 3002, Australia
Centre for Ophthalmology and Vision Science, Lions Eye Institute, University of Western Australia, Perth, Western Australia, 6009, Australia
Menzies Research Institute Tasmania, University of Tasmania, Hobart, Tasmania, 7000, Australia
Generation Scotland, Centre for Genomic and Experimental Medicine, University of Edinburgh, Edinburgh, EH4 2XU, UK
Musculoskeletal Research Programme, Division of Applied Medicine, University of Aberdeen, Aberdeen, AB25, UK
K.G. Jebsen Center for Genetic Epidemiology, Department of Public Health, NTNU, Norwegian University of Science and Technology, Trondheim, 7600, Norway
AMC, Department of Vascular Medicine, Amsterdam, 1105 AZ, The Netherlands
HUNT Research Centre, Department of Public Health and General Practice, Norwegian University of Science and Technology, Levanger, 7600, Norway
Department of Neurology, Erasmus Medical Center, Rotterdam, 3015 GE, The Netherlands
Department of Radiology, Erasmus Medical Center, Rotterdam, 3015 GE, The Netherlands
Department of Medicine, Division of Cardiovascular Medicine, Stanford University School of Medicine, Stanford, CA, 943 05, USA
Department of Public Health & Clinical Medicine, Umeå University, Umeå, SE-90185, Sweden
Research Unit Skellefteå, Skellefteå, SE-93141, Sweden
Department of Genome Sciences, University of Washington, Seattle, WA, 98195, USA
The Copenhagen City Heart Study, Frederiksberg Hospital, Frederiksberg, 2000, Denmark
Department of Preventive Medicine, Keck School of Medicine of the University of California, Los Angeles, California, USA, 90089, USA
USC Roski Eye Institute, Department of Ophthalmology, Keck School of Medicine of the University of Southern California, Los Angeles, CA, 90089, USA
Center for Medical Genetics and Molecular Medicine, Haukeland University Hospital, Bergen, 5021, Norway
National Institute of Public Health, University of Southern Denmark, Copenhagen, 1353, Denmark
Steno Diabetes Center, Gentofte, 2820, Denmark
Aalborg University, Aalborg, DK-9000, Denmark
Research Center for Prevention and Health, Capital Region of Denmark, Glostrup, DK-2600, Denmark
Department of Health, National Institute for Health and Welfare, Helsinki, 00271, Finland
Department of Cardiology, Leiden University Medical Center, Leiden, 2333, The Netherlands
The Interuniversity Cardiology Institute of the Netherlands, Utrecht, 2333, The Netherlands
Department of Computational Medicine and Bioinformatics, University of Michigan, Ann Arbor, MI, 48109, USA
Department of Internal Medicine, University of Michigan, Ann Arbor, MI, 48109, USA
Division of Gastroenterology, University of Michigan, Ann Arbor, MI, 48109, USA
Department of Psychiatry, Brain Center Rudolf Magnus, University Medical Center Utrecht, Utrecht, 3584 CG,

The Netherlands
172 Department of Clinical Physiology, University of Tampere School of Medicine, Tampere, 33014, Finland
173 National Institute for Health and Welfare, Helsinki, FI-00271, Finland
174 Echinos Medical Centre, Echinos, Greece
175 Oxford Centre for Diabetes, Endocrinology and Metabolism, Radcliffe Department of Medicine, University of Oxford, Oxford, OX3 7LE, UK
176 Oxford NIHR Biomedical Research Centre, Oxford University Hospitals Trust, Oxford, OX3 7LE, UK
177 UKCRC Centre of Excellence for Public Health Research, Queens University Belfast, Belfast, UK, BT12 6BJ, UK
178 Netherlands Cancer Institute - Antoni van Leeuwenhoek hospital, Amsterdam, 1066 CX, The Netherlands
179 Department of Restorative Dentistry, Periodontology and Endodontology, University Medicine Greifswald, Greifswald, 17475, Germany
180 Foundation for Research in Health Exercise and Nutrition, Kuopio Research Institute of Exercise Medicine, Kuopio, 70100, Finland
181 National Heart and Lung Institute, Imperial College London, Hammersmith Hospital Campus, London, W12 0NN, UK
182 Division of Public Health Sciences, Fred Hutchinson Cancer Research Center, Seattle WA, 98109, USA
183 German Center for Diabetes Research, Neuherberg, D-85764, Germany
184 Institute of Epidemiology II, Helmholtz Zentrum München - German Research Center for Environmental Health, Neuherberg, D-85764, Germany
185 Research Unit of Molecular Epidemiology, Helmholtz Zentrum München - German Research Center for Environmental Health, Neuherberg, D-85764, Germany
186 Department of Psychiatry, and Division of Molecular Biology and Human Genetics, Department of Biomedical Sciences, Faculty of Medicine and Health Sciences, Stellenbosch University, Tygerberg, Western Cape, 7505, South Africa
187 CHU Nantes, Service de Génétique Médicale, Nantes, 44093, France
188 Institute of Clinical Medicine, Internal Medicine, University of Eastern Finland and Kuopio University Hospital, Kuopio, 70210, Finland
189 Institute for Maternal and Child Health - IRCCS "Burlo Garofolo", Trieste, 34137, Italy
190 Institute of Biomedicine & Physiology, University of Eastern Finland, Kuopio, 70210, Finland
191 Department of Genetics, University of North Carolina, Chapel Hill, NC, 27514, USA
192 Department of Biostatistical Sciences and Center for Public Health Genomics, Wake Forest School of Medicine, Winston-Salem, NC, 27157, USA
193 MRC Epidemiology Unit, University of Cambridge School of Clinical Medicine, Institute of Metabolic Science, Cambridge, CB2 0QQ, UK
194 Group Health Research Institute, Seattle, WA, 98101, USA
195 Department of Health Services, University of Washington, Seattle WA 98101
196 Division of Endocrinology and Metabolism, Department of Internal Medicine, Taichung Veterans General Hospital, Taichung, Taiwan 407, Taiwan
197 School of Medicine, National Yang-Ming School of Medicine, Chung Shan Medical University, Taipei, Taiwan 112, Taiwan
198 Division of Preventive Medicine University of Alabama at Birmingham, Birmingham, AL, AL 35205, USA
199 Key Laboratory of Nutrition and Metabolism, Institute for Nutritional Sciences, Shanghai Institutes for Biological Sciences, Chinese Academy of Sciences, University of the Chinese Academy of Sciences, Shanghai, People's Republic of China, Shanghai, 200031, China
200 Department of Medicine, Division of Cardiovascular Medicine, Stanford University School of Medicine, Palo Alto, CA, 94304, USA
201 Department of Medicine, Boston University School of Medicine, Boston, MA, 02118, USA
202 Uppsala University, Uppsala, 75185, Sweden
203 Chronic Disease Prevention Unit, National Institute for Health and Welfare, Helsinki, 00271, Finland
204 Division of Public Health Sciences, Wake Forest School of Medicine, Winston-Salem, NC, 27157, USA
205 Division of Health Sciences, Warwick Medical School, Warwick University, Coventry, CV4 7AL, UK
206 Department of Psychiatry, Washington University, Saint Louis, MO, 63110, USA
207 Department of Medicine, Harvard University Medical School, Boston, MA, 02115, USA
208 Department of Molecular Epidemiology, German Institute of Human Nutrition Potsdam-Rehbruecke (DIfE), Nuthetal, 14558, Germany
209 Westmead Millennium Institute of Medical Research, Centre for Vision Research and Department of Ophthalmology, University of Sydney, Sydney, New South Wales, 2022, Australia
210 Department of Public Health and Primary Care, Leiden University Medical Center, Leiden, 2300RC, The Netherlands
211 Usher Institute of Population Health Sciences and Informatics, University of Edinburgh, Edinburgh, EH8 9AG, UK

212 Department of Medicine I, Ludwig-Maximilians-Universität, Munich, 81377, Germany
213 DZHK (German Centre for Cardiovascular Research), partner site Munich Heart Alliance, Munich, 80802, Germany
214 Laboratory of Neurogenetics, National Institute on Aging, NIH, Bethesda, MD, 20892, USA
215 DZHK (German Centre for Cardiovascular Research), partner site Greifswald, Greifswald, 17475, Germany
216 Institute of Clinical Chemistry and Laboratory Medicine, University Medicine Greifswald, Greifswald, 17475, Germany
217 Department of Cardiology, Heart Center, Tampere University Hospital and School of Medicine, University of Tampere, Tampere, 33521, Finland
218 Program in Personalized Medicine, Department of Medicine, University of Maryland School of Medicine, Baltimore, MD, 21201, US
219 Department of Medicine, Tampere University Hospital, Tampere, 33521, Finland
220 Center for Neurobehavioral Genetics, UCLA, Los Angeles, CA, 90095, USA
221 Pat Macpherson Centre for Pharmacogenetics and Pharmacogenomics, Medical Research Institute, Ninewells Hospital and Medical School, Dundee, DD1 9SY, UK
222 Laboratory of Clinical Chemistry and Hematology, Division Laboratories and Pharmacy, University Medical Center Utrecht, Utrecht, 3584 CX, The Netherlands
223 Laboratory of Experimental Cardiology, Division Heart & Lungs, University Medical Center Utrecht, Utrecht, 3584 CX, The Netherlands
224 School of Womenís and Infantsí Health, The University of Westerna Australia, Perth, Western Australia, 6009, Australia
225 University of Helsinki, Institute for Molecular Medicine (FIMM) and Diabetes and Obesity Research Program, Helsinki, FI00014, Finland
226 University of Tartu, Estonian Genome Center, Tartu, Estonia, Tartu, 51010, Estonia
227 School of Medicine, University of Split, Split, 21000, Croatia
228 Center for Neurogenomics and Cognitive Research, Department Complex Trait Genetics, VU University, Amsterdam, 1081 HV, The Netherlands
229 Neuroscience Campus Amsterdam, Department Clinical Genetics, VU Medical Center, Amsterdam, 1081 HV, The Netherlands
230 Department of Clinical Physiology and Nuclear Medicine, Turku University Hospital, Turku, 20521, Finland
231 Research Centre of Applied and Preventive Cardiovascular Medicine, University of Turku, Turku, 20520, Finland
232 Centre for Non-Communicable Diseases, Karachi, Pakistan
233 Kuopio University Hospital, Department of Clinical Physiology and Nuclear Medicine, Kuopio, 70210, Finland
234 Merck Research Laboratories, Genetics and Pharmacogenomics, Boston, MA, 02115, USA
235 Department of Epidemiology, University of Washington, Seattle, WA, 98195, USA
236 Department of Biobank Research, Umeå University, Umeå, SE-90187, Sweden
237 Division of Cardiovascular Medicine, Brigham and Womenís Hospital and Harvard Medical School, Boston, MA, 02115, USA
238 Department of Medicine, Faculty of Medicine, Université de Montréal, Montreal, Quebec, H3T 1J4, Canada
239 Department of Public Health and Clinical Medicine, Unit of Family Medicine, Umeå University, Umeå, 90185, Sweden
240 Department of Biostatistics and Epidemiology, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA, 19104, USA
241 Division of Epidemiology & Community Health University of Minnesota, Minneapolis, MN, 55454, USA
242 Duke University, Durham, NC, 27703, USA
243 Saw Swee Hock School of Public Health, National University of Singapore, National University Health System, Singapore, Singapore
244 Departement of Haematology, University of Cambridge, Cambridge, CB2 OPT, UK
245 Department of Vascular Medicine, AMC, Amsterdam, 1105 AZ, The Netherlands
246 Department of Twin Research and Genetic Epidemiology, Kingís College London, London, SE1 7EH, UK
247 Alzheimer Scotland Dementia Research Centre, University of Edinburgh, Edinburgh, EH8 9JZ, UK
248 Department of Epidemiology and Biostatistics, VU University Medical Center, Amsterdam, 1081BT, The Netherlands
249 Department of Molecular Cell Biology, Leiden University Medical Center, Leiden, 2333ZC, The Netherlands
250 Department of Molecular Epidemiology, Leiden University Medical Center, Leiden, 2333ZC, The Netherlands
251 College of Biomedical and Life Sciences, Cardiff University, Cardiff, CF14 4EP, UK
252 MRC Integrative Epidemiology Unit, School of Social and Community Medicine, University of Bristol, Bristol, BS8 2BN, UK
253 Los Angeles Biomedical Research Institute at Harbor, UCLA, Los Angeles, CA, 90095 , USA
254 Institute for Community Medicine, University Medicine Greifswald, Greifswald, 17475, Germany
255 Center for Pediatric Research, Department for Womenís and Child Health, University of Leipzig, Leipzig,

04103, Germany
256 Division of Molecular Biology and Human Genetics, Department of Biomedical Sciences, Faculty of Medicine and Health Sciences, Stellenbosch University, Tygerberg, Western Cape, 7505, South Africa
257 Department of Gerontology and Geriatrics, Leiden University Medical Center, Leiden, 2333, The Netherlands
258 Anogia Medical Centre, Anogia, Greece
259 Centre for Vascular Prevention, Danube-University Krems, Krems, 3500, Austria
260 Dasman Diabetes Institute, Dasman, 15462, Kuwait
261 Diabetes Research Group, King Abdulaziz University, Jeddah, 21589, Saudi Arabia
262 Department of Psychiatry, Dalhousie University, Halifax, B3H 4R2, Canada
263 University of Amsterdam, Department of Brain & Cognition, Amsterdam, 1018 WS, The Netherlands
264 VU University Medical Center, Department of Epidemiology and Biostatistics, Amsterdam, 1007 MB, The Netherlands
265 Department of Obstetrics and Gynecology, Institute for Medicine and Public Health, Vanderbilt Genetics Institute, Vanderbilt University, Nashville, TN, 37203, USA
266 Merck Research Laboratories, Cardiometabolic Disease, Kenilworth, NJ, 07033, USA
267 Institute of Cellular Medicine, The Medical School, Newcastle University, Newcastle, NE2 4HH, UK
268 Department of Biostatistics, Boston University School of Public Health, Boston, MA, 02118, USA
269 Departments of Epidemiology & Medicine, Diabetes Translational Research Center, Fairbanks School of Public Health & School of Medicine, Indiana University, Indiana, IN, 46202, USA
270 Department of Physiology and Biophysics, University of Mississippi Medical Center, Jackson, MS, 39216, USA
271 Danish Diabetes Academy, Odense, 5000, Denmark
272 Department of Public Health, Aarhus University, Aarhus, 8000, Denmark
273 Memorial University, Faculty of Medicine, Discipline of Genetics, St. Johnís, NL, A1B 3V6, Canada
274 Genetics of Complex Traits, University of Exeter Medical School, University of Exeter, Exeter, EX2 5DW, UK
275 GlaxoSmithKlein, King of Prussia, PA, 19406, USA
276 Department of Clinical Sciences, Quantitative Biomedical Research Center, Center for the Genetics of Host Defense, University of Texas Southwestern Medical Center, Dallas, TX, 75390, USA
277 Department of Human Genetics, University of Michigan, Ann Arbor, MI, 48109, USA
278 Department of Public Health Sciences, Institute for Personalized Medicine, the Pennsylvania State University College of Medicine, Hershey, PA, 17033, USA
279 Department of Epidemiology and Carolina Center of Genome Sciences, Chapel Hill, NC, 27514, USA
280 Department of Neurology, Boston University School of Medicine, Boston, MA, 02118, USA
281 Department of Epidemiology Research, Statens Serum Institut, Copenhagen, 2200, Denmark
282 Li Ka Shing Centre for Health Information and Discovery, The Big Data Institute, University of Oxford, Oxford, OX3 7BN, UK
283 The Mindich Child Health and Development Institute, Ichan School of Medicine at Mount Sinai, New York, NY, 10069, USA
284 Departments of Pediatrics and Genetics, Harvard Medical School, Boston, MA, 02115, USA
285 Princess Al-Jawhara Al-Brahim Centre of Excellence in Research of Hereditary Disorders (PACER-HD), King Abdulaziz University, Jeddah, 21589, Saudi Arabia