Human Molecular Genetics, 2015, Vol. 24, No. 6 1791–1800

doi: 10.1093/hmg/ddu583
Advance Access Publication Date: 26 November 2014
Association Studies Article

ASSOCIATION STUDIES ARTICLE

Meta-analysis of genome-wide association studies of adult height in East Asians identifies 17 novel loci

Meian He[1,†,*], Min Xu[2,3,†], Ben Zhang[5,†], Jun Liang[6,†], Peng Chen[7,†], Jong-Young Lee[14,†], Todd A. Johnson[15,†], Huaixing Li[19,†], Xiaobo Yang[20,22,†], Juncheng Dai[23,†], Liming Liang[2,†], Lixuan Gui[1], Qibin Qi[2], Jinyan Huang[2], Yanping Li[2], Linda S. Adair[24], Tin Aung[26], Qiuyin Cai[5], Ching-Yu Cheng[7,8,26,27], Myeong-Chan Cho[14], Yoon Shin Cho[14], Minjie Chu[23], Bin Cui[3], Yu-Tang Gao[28], Min Jin Go[14], Dongfeng Gu[29], Weiqiong Gu[3], Huan Guo[1], Yongchen Hao[29], Jie Hong[3], Zhibin Hu[23], Yanling Hu[21], Jianfeng Huang[29], Joo-Yeon Hwang[14], Mohammad Kamran Ikram[7,8,26,31,32], Guangfu Jin[23], Dae-Hee Kang[33], Chiea Chuen Khor[7,8,9,34], Bong-Jo Kim[14], Hung Tae Kim[14], Michiaki Kubo[16], Jeannette Lee[7], Juyoung Lee[14], Nanette R. Lee[35], Ruoying Li[10], Jun Li[1], JianJun Liu[7,34], Jirong Longe[5], Wei Lu[36], Xiangfeng Lu[29], Xiaoping Miao[1], Yukinori Okada[17], Rick Twee-Hee Ong[7], Gaokun Qiu[1], Mark Seielstad[9], Xueling Sim[7,37], Huaidong Song[4], Fumihiko Takeuchi[38], Toshihiro Tanaka[18], Phil R. Taylor[39], Laiyuan Wang[29], Weiqing Wang[3], Yiqin Wang[19], Chen Wu[30], Ying Wu[25], Yong-Bing Xiang[28], Ken Yamamoto[40], Handong Yang[41], Ming Liao[20], Mitsuhiro Yokota[42], Terri Young[27,43], Xiaomin Zhang[1], Norihiro Kato[38], Qing K. Wang[44], Wei Zheng[5], Frank B. Hu[2], Dongxin Lin[30], Hongbing Shen[23,‡], Yik Ying Teo[7,11,12,13,34,‡], Zengnan Mo[20,‡], Tien Yin Wong[8,26,‡], Xu Lin[19,‡], Karen L. Mohlke[25,‡], Guang Ning[3,‡], Tatsuhiko Tsunoda[15,‡], Bok-Ghee Han[14,‡], Xiao-Ou Shu[5,‡], E. Shyong Tai[7,10,27,‡], Tangchun Wu[1,‡], and Lu Qi[2,‡]

[1]MOE Key Lab of Environment and Health, School of Public Health, Tongji Medical College, Huazhong University of Science & Technology, Hubei, Wuhan 430030, China, [2]Department of Nutrition, Harvard School of Public Health, Boston, MA 02115, USA, [3]Key Laboratory for Endocrine and Metabolic Diseases of Ministry of Health, Shanghai Clinical Center for Endocrine and Metabolic Diseases, Shanghai Institute of Endocrine and Metabolic

[†] Co-first authors.
[‡] Co-last authors.

Received: June 9, 2014. **Revised and Accepted:** November 17, 2014.

Diseases, Department of Endocrinology and Metabolism and [4]State Key Laboratory of Medical Genomics, Molecular Medical Center, Shanghai Institute of Endocrine and Metabolic Diseases, Ruijin Hospital, Shanghai Jiao Tong University School of Medicine, Shanghai 200025, China, [5]Vanderbilt Epidemiology Center, Division of Epidemiology, Department of Medicine, Vanderbilt University School of Medicine, Nashville, TN 37232, USA, [6]Department of Endocrinology, The Central Hospital of Xuzhou, Affiliated Hospital of Southeast University, Xuzhou, Jiangsu 221009, China, [7]Saw Swee Hock School of Public Health, [8]Department of Ophthalmology, [9]Department of Paediatrics National University Health Systems, [10]Department of Medicine, Yong Loo Lin School of Medicine, [11]NUS Graduate School for Integrative Science and Engineering, [12]Life Sciences Institute, [13]Department of Statistics and Applied Probability, National University of Singapore, Singapore, Singapore 119077, [14]Center for Genome Science, National Institute of Health, Osong Health Technology Administration Complex, Chungcheongbuk-do 363-700, Republic of Korea, [15]Laboratory for Medical Informatics, [16]Laboratory for Genotyping Development, [17]Laboratory for Statistical Analysis, [18]Laboratory for Cardiovascular Diseases, RIKEN Center for Genomic Medicine, Yokohama 230-0045, Japan, [19]Key Laboratory of Nutrition and Metabolism, Institute for Nutritional Sciences, Shanghai Institutes for Biological Sciences, Chinese Academy of Sciences, Shanghai 200031, China, [20]Institute of Urology and Nephrology, First Affiliated Hospital & Center for Genomic and Personalized Medicine, [21]Medical Scientific Research Center and [22]Department of Occupational Health and Environmental Health, School of Public Health, Guangxi Medical University, Nanning, Guangxi 530021, China, [23]Department of Epidemiology & Biostatistics, School of Public Health, Nanjing Medical University, Nanjing 210029, China, [24]Department of Nutrition and [25]Department of Genetics, University of North Carolina, Chapel Hill, NC 27599, USA, [26]Singapore Eye Research Institute, Singapore National Eye Centre, Singapore, Singapore 168751, [27]Duke-National University of Singapore Graduate Medical School, Singapore, Singapore 169857, [28]Department of Epidemiology, Shanghai Cancer Institute, Shanghai Jiao Tong University School of Medicine, Shanghai 200025, China, [29]State Key Laboratory of Cardiovascular Disease, Department of Evidence Based Medicine, Fuwai Hospital, National Center of Cardiovascular Diseases, [30]State Key Laboratory of Molecular Oncology, Cancer Institute & Hospital, Chinese Academy of Medical Sciences and Peking Union Medical College, Beijing 100730, China, [31]Department of Ophthalmology, Erasmus Medical Center, Rotterdam 3015, The Netherlands, [32]Memory Aging & Cognition Centre, National University Health System, Singapore, Singapore 119228, [33]Department of Preventive Medicine, Seoul National University, College of Medicine, Seoul 110-799, Republic of Korea, [34]Agency for Science, Technology and Research, Genome Institute of Singapore, Singapore, Singapore 138672, [35]Office of Population Studies Foundation, University of San Carlos, Cebu City 6000, Philippines, [36]Shanghai Municipal Center for Disease Control & Prevention, 1380 Zhong Shan Road (W), Shanghai 200336, China, [37]Department of Biostatistics and Center for Statistical Genetics, University of Michigan, Ann Arbor, MI 48109-2029, USA, [38]Department of Gene Diagnostics and Therapeutics, Research Institute, National Center for Global Health and Medicine, Tokyo 162-8655, Japan, [39]Division of Cancer Epidemiology & Genetics, Genetic Epidemiology Branch, National Cancer Institute, Bethesda, MD 20892, USA, [40]Department of Molecular Genetics, Medical Institute of Bioregulation, Kyushu University, Fukuoka 812-8582, Japan, [41]Department of Cardiology, Dongfeng Central Hospital, Dongfeng Motor Corporation and Hubei University of Medicine, Shiyan, Hubei 442008, China, [42]Department of Genome Science, Aichi-Gakuin University, School of Dentistry, Nagoya 464-8650, Japan, [43]Duke Eye Center, Duke University Medical Center, Durham, NC 27710, USA, and [44]Key Laboratory of Molecular Biophysics of the Ministry of Education, Cardio-X Institute, College of Life Science and Technology and Center for Human Genome Research, Huazhong University of Science and Technology, Wuhan, Hubei 430074, China

*To whom correspondence should be addressed. Email: hemeian@hotmail.com

Abstract

Human height is associated with risk of multiple diseases and is profoundly determined by an individual's genetic makeup and shows a high degree of ethnic heterogeneity. Large-scale genome-wide association (GWA) analyses of adult height in Europeans have identified nearly 180 genetic loci. A recent study showed high replicability of results from Europeans-based GWA studies in Asians; however, population-specific loci may exist due to distinct linkage disequilibrium patterns. We carried out a GWA meta-analysis in 93 926 individuals from East Asia. We identified 98 loci, including 17 novel and 81 previously reported loci, associated with height at $P < 5 \times 10^{-8}$, together explaining 8.89% of phenotypic variance. Among the newly identified variants, 10 are commonly distributed (minor allele frequency, MAF > 5%) in Europeans, with comparable frequencies with in Asians, and 7

single-nucleotide polymorphisms are with low frequency (MAF < 5%) in Europeans. In addition, our data suggest that novel biological pathway such as the protein tyrosine phosphatase family is involved in regulation of height. The findings from this study considerably expand our knowledge of the genetic architecture of human height in Asians.

Introduction

Human height is a classical and highly heritable complex trait, with an estimated 80–90% of variation explained by genetic components (1,2). Height has been associated with risk of multiple diseases including cancer (3,4), stroke (5) and atrial fibrillation (6). Compelling evidence has also indicated that height-associated genetic variants may also predisposing to increased risk of cancer (7). Recently, a large-scale genome-wide association (GWA) meta-analysis by the genetic investigation of anthropocentric traits (GIANT) consortium identified 180 loci associated with adult height in individuals of European ancestry (8). However, the loci identified so far only explain a relatively small proportion of the variance (~10%) in the trait, leaving additional genetic variations to be identified. In addition, a high degree of ethnic heterogeneity in adult height clearly exists. A recent study indicates population-specific loci may exist due to distinct linkage disequilibrium patterns between Europeans and East Asians, despite high trans-ethnic replicability of GWA associations (9). Several previous GWA studies (GWAS) on adult height in Korean, Japanese and Chinese (10–12) were severely limited by their relatively small size, and therefore offer an incomplete picture of the genetic architecture of adult height in Asians.

To explore the population-specific genetic architecture of adult height, we performed thus far the largest GWAS meta-analysis in Asians. The current two-stage study included a total of 93 926 adults from seven countries/regions of East Asia. We reported here the identification of in total of 98 loci, including 17 novel and 81 previously reported loci, which associated with adult height in Asians at genome-wide significant level.

Results

In Stage 1, we performed a meta-analysis of GWA data from 11 studies, comprising 36 227 individuals. The genotyping data included ~2.7 million genotyped or imputed single-nucleotide polymorphisms (SNPs) present in the HapMap Phase 2 http://hapmap.ncbi.nlm.nih.gov/downloads/genotypes/2010-05_phaseIII/hapmap_format/consensus/ Chinese and Japanese reference panel. The Stage 1 analysis validated 32 previously reported loci (10–12) at a genome-wide significance level (Supplementary Material, Table S1; $P < 5 \times 10^{-8}$). To identify additional novel loci robustly associated with adult height, we carried forward at least one SNP from each of the loci (a locus is defined as a chromosomal region at which adjacent pairs of SNPs were <1 Mb) reaching $P < 1.0 \times 10^{-5}$ in Stage 1, with a total number of SNPs of 1783, into additional samples of 57 699 East Asians with GWA data available (Stage 2). We also carried forward an additional 379 previously reported SNPs (10–12) associated with height but did not reach $P < 1.0 \times 10^{-5}$ in our Stage 1 data. Following the same analysis strategy as in Stage 1, each individual study in Stage 2 tested the association of adult height with each SNP, and meta-analyses using the inverse variance weight method were performed to combine the individual study statistics of Stage 2, and then of both Stage 1 and Stage 2 using a fixed-effect model (Supplementary Material, Fig. S1).

In the combined analysis of Stages 1 and 2, we identified 17 novel loci located on 11 autosomal chromosomes (Fig. 1) at or near genes including *PTPN14*, *CRIM1*, *ZNF638*, *SLC17A2*, *UBD*, *HCG22*, *EXTL3*, *MYC*, *FADS1*, *CCND1*, *PAN2*, *PTPN9*, *CDK10*, *KDM6B*, *ILF3*, *DNM2* and *RBL1* associated with height at a genome-wide significance level (Table 1; Supplementary Material, Fig. S2; $P = 2.24 \times 10^{-8}$–6.80×10^{-35}). These loci each explained 0.02–0.19% of the variance of height. We also found five novel loci showing nearly genome-wide significant associations ($P = 1.76 \times 10^{-7}$–6.00×10^{-8}; Supplementary Material, Table S2). Supplementary Material, Table S9 presents detailed information for the nearest genes near the 17 newly identified height loci.

In addition, 81 previously reported height-associated SNPs (10–12) showed genome-wide significant associations ($P < 5 \times 10^{-8}$; Supplementary Material, Table S1) in the combined analysis of Stages 1 and 2. Furthermore, 13 previously reported loci reached nearly genome-wide significance ($P < 5 \times 10^{-7}$; Supplementary Material, Table S1) and 62 loci showed a nominally significant association ($5 \times 10^{-7} < P < 0.05$; Supplementary Material, Table S3) in the combined analysis. There were 17, 105 and 5 previously reported height-related variants among Asian, European and African, respectively, failed to be replicated in the present study.

Using the largest of our study samples (SGWAS; N = 9297; Supplementary Material, Methods), we found that the 17 newly identified loci explained 2.33% of the phenotypic variance in height; in total, the 98 genome-wide significant loci explained 8.89% of the phenotypic variance, these might be an overestimate due to winner's curse. When we removed these 98 height variants and nearby correlated SNPs from the meta-analysis, the QQ plot indicated that the number of low *P*-values still deviated from the expectation (Supplementary Material, Fig. S3).

We further compared the frequencies of the index SNPs representing 17 newly identified loci between East Asians and Europeans (HapMap CEU frequencies, http://www.broadinstitute.org/collaboration/giant/index.php/GIANT_consortium_data_files). Seven SNPs are with minor allele frequency (MAF) <5% in Europeans. Ten of the 17 newly identified SNPs are commonly distributed (MAF > 5%) in Europeans, showing similar frequencies as in Asians (Fig. 2). We noted that 6 of these 10 SNPs showed nominally significant associations with height in GIANT, although none reached genome-wide significance level.

For the newly identified loci, we also analyzed the top SNPs of each locus (discovery stage $P < 1.0 \times 10^{-5}$), conditioning on the index SNP, in the combined samples using GCTA software (13). At a genome-wide significance level, we identified a second independent association for SNP rs10172196 near *CRIM1* ($P = 3.97 \times 10^{-9}$; $r^2 = 0.04$ with the index SNP rs3755206) and rs12624 near *HCG22* ($P = 1.01 \times 10^{-8}$; $r^2 = 0.01$ with the index SNP rs2251830). In accordance with previous studies (13), we defined a locus as a chromosomal region at which adjacent pairs of SNPs were <1 Mb. Our observations are in line with a previous analysis in GIANT which found multiple independent-associated SNPs of a single locus for height (13). For each locus, the nearest gene to the strongest associated variant was prioritized as biological candidates.

We further evaluated the heterogeneity of the 98 identified loci across studies, countries/regions, and between genders within our East Asian samples. The *P*-value of <0.05 was considered nominally significant. Among these loci, only three showed significant heterogeneity across studies ($P = 0.006$–0.03; $I^2 = 79.4$–86.8%) and four showed nominally significant heterogeneity across countries/regions ($P = 0.003$–0.015; $I^2 = 67.4$–77.1%), with the *RBL1* showing

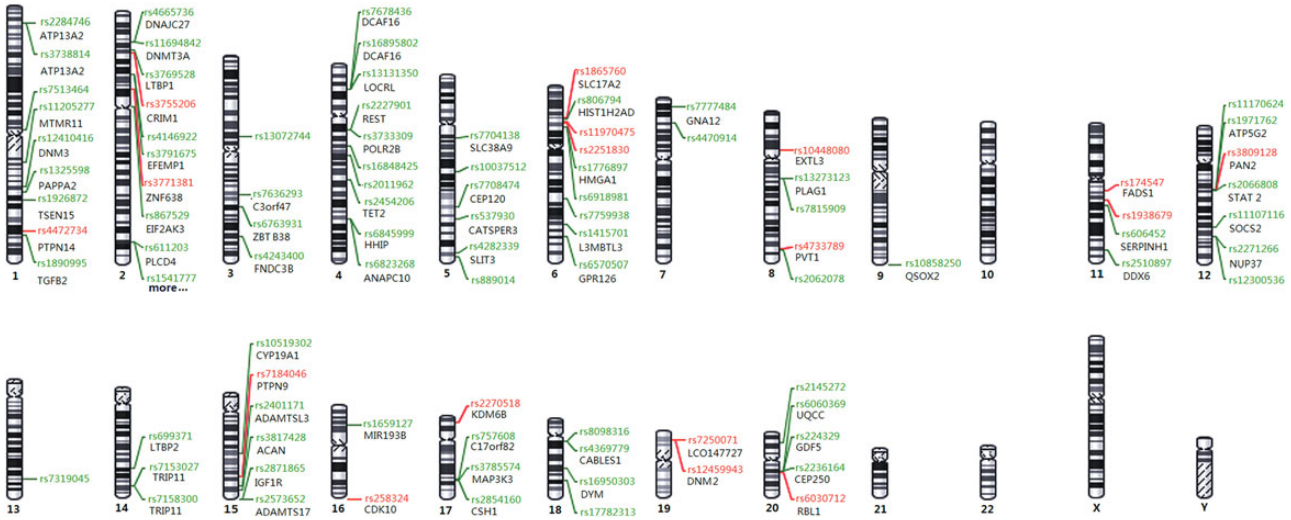

Figure 1. The 98 loci associated with adult height in East Asians. The genome location of the 17 novel loci (in red) and 81 previously reported loci (in green) was shown by each chromosome. The chromosome numbers were shown in blue. The figure was created by using Affymation (http://genepipe.ngc.sinica.edu.tw/mkr/menu.do).

Table 1. Identified novel loci associated with height in an East Asian population

SNP	Chr	Position	The nearest gene	Effect/ other allele	Stage 1				Stage 2				Combined *P*-value	Explained variance (%)
					EAF	Beta	SE	*P*-value	EAF	Beta	SE	*P*-value		
rs4472734	1	212684808	*PTPN14*	T/C	0.55	−0.034	0.007	3.98E−06	0.48	−0.024	0.006	6.42E−05	1.01E−09	0.03
rs3755206	2	36536932	*CRIM1*	T/G	0.82	0.043	0.010	7.56E−06	0.87	0.048	0.009	2.87E−08	1.01E−12	0.07
rs3771381	2	71414173	*ZNF638*	A/T	0.43	−0.035	0.008	4.06E−06	0.48	−0.045	0.006	1.77E−14	7.66E−19	0.10
rs1865760	6	26024958	*SLC17A2*	T/C	0.70	−0.042	0.008	7.34E−08	0.69	−0.033	0.007	4.23E−07	1.86E−13	0.05
rs11970475	6	29634356	*UBD*	A/G	0.18	0.051	0.010	8.91E−07	0.19	0.070	0.008	1.34E−19	2.46E−24	0.14
rs2251830	6	31124957	*HCG22*	A/C	0.37	0.036	0.008	2.15E−06	0.36	0.053	0.006	1.05E−17	5.25E−22	0.13
rs10448080	8	28660710	*EXTL3*	T/C	0.70	0.037	0.008	3.44E−06	0.68	0.023	0.007	4.97E−04	2.24E−08	0.02
rs4733789	8	128903585	*MYC*	T/C	0.46	0.037	0.007	6.04E−07	0.46	0.025	0.006	4.12E−05	2.53E−10	0.03
rs174547	11	61327359	*FADS1*	T/C	0.53	0.050	0.008	2.60E−10	0.62	0.037	0.006	7.52E−10	5.88E−18	0.07
rs1938679	11	68981277	*CCND1*	T/C	0.35	−0.036	0.008	6.90E−06	0.33	−0.032	0.006	4.52E−07	1.47E−11	0.05
rs3809128	12	54996186	*PAN2*	T/C	0.17	−0.100	0.012	9.31E−17	0.21	−0.083	0.009	4.03E−20	6.80E−35	0.19
rs7184046	15	73653205	*PTPN9*	C/G	0.35	0.036	0.008	7.09E−06	0.42	0.030	0.006	1.61E−06	1.77E−10	0.04
rs258324	16	88281756	*CDK10*	T/G	0.31	0.058	0.009	1.77E−10	0.27	0.054	0.007	8.11E−14	4.38E−23	0.12
rs2270518	17	7699247	*KDM6B*	T/C	0.23	0.049	0.010	2.40E−06	0.19	0.033	0.008	2.62E−05	8.12E−10	0.04
rs7250071	19	10626819	*ILF3*	T/C	0.62	−0.036	0.008	3.08E−06	0.63	−0.042	0.006	9.82E−12	1.94E−16	0.08
rs12459943	19	10720508	*DNM2*	A/G	0.33	−0.035	0.008	8.93E−06	0.36	−0.030	0.006	1.50E−06	5.43E−11	0.04
rs6030712	20	35070812	*RBL1*	A/G	0.51	−0.048	0.007	4.31E−11	0.39	−0.021	0.006	6.56E−04	8.66E−12	0.02

Chr., chromosome. Positions are based on National Center for Biotechnology Information build 36. EAF, effect allele frequency in Asians, estimated from Stages 1 and 2. The effect sizes (β) obtained from the combined data were used to estimate the explained variance.

nominally significant heterogeneity both between the studies and between countries/regions (Supplementary Material, Table S5 and S6). Our data indicate that the vast majority of the genetic variants identified in our study have consistent effects on height across various populations in Asia. In addition, 15 loci displayed nominally significant sexual dimorphism in association with height ($P <$ 0.001–0.048; $I^2 = 74.3$–97.5%; Supplementary Material, Table S7). However, none of the above heterogeneity tests remained significant after controlling for multiple comparisons.

To explore biological connections between the 98 height loci which reached GWAS significance in East Asians, we performed a pathway analysis based on the Stage 1 meta-analysis results using the meta-analysis gene-set enrichment of variant associations (MAGENTA) method (14,15). Our analysis revealed seven biological pathways and six molecular functions nominally enriched ($P < 0.05$), many of which lie within the validated height loci (Supplementary Material, Table S10). Among them, the histone gene set remained significant after controlling for multiple testing. We also applied the GRAIL text-mining algorithm to explore non-random and literature-based connectivity among the genes (Supplementary Material, Fig. S5). Among the genes implicated by proximity to the height-associated loci identified in the present study, 98 genes were identified as significant ($P < 0.01$), all obtaining false discovery rate (FDR) < 0.05. Forty SNPs reached GRAIL *P*-value < 0.01 and FDR < 0.05, indicating that the 40 regions tagged by these SNPs contained genes which were related to other implicated genes more than expected by chance (Supplementary Material, Table S11). Totally, GRAIL identified strong

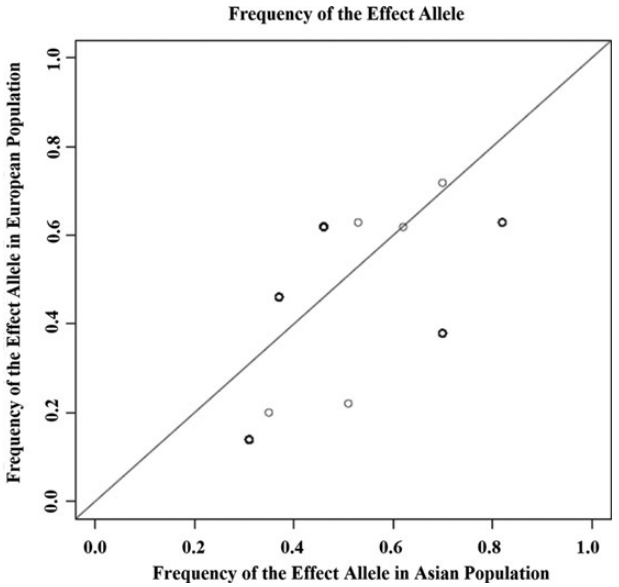

Figure 2. The frequencies of the 10 newly identified SNPs with MAF > 5% in Asians and Europeans.

connections between members of three key biological pathways: insulin/insulin-like growth factor 1 (IGF1) pathway (encoded by *GH1*, *IGF1* and *IGF1R*), the Hedgehog signaling pathway (encoded by *IHH* and *HHIP*) and the transforming growth factor (TGF)-β signaling pathway (encoded by *LTBP1*, *TGFβ* and *BMP6*). Among them, two biological pathways of the insulin/IGF1 pathway and the Hedgehog signaling pathway were consistent with the results in the MAGENTA pathway analysis. The Hedgehog signaling pathway and the TGFβ signaling pathway were reported in the GIANT study (8).

Discussion

In thus far the largest GWA meta-analysis of height in Asians, we identified 98 loci (MAF ranges from 0.08 to 0.49), including 17 novel and 81 previously reported loci. These variants together explained ~8.89% of the phenotypic variation in height in East Asians. Our findings considerably expand knowledge of the genetic architecture of human height in Asians.

For the 17 novel height loci in Asians, 7 newly identified SNPs are with low frequency (MAF < 5%) in Europeans. Marigorta *et al.* observed that GWAS with larger sample sizes might be more powerful in detection of variants with weaker effects than with lower frequencies (9). This may partly explain why these seven SNPs were not detected in GIANT and highlight the importance of GWAS in various ethnic populations in identifying novel genetic variants. We have compared the associations of the 17 loci with height between our study and GIANT; and found that the effect sizes of the common SNPs observed in GIANT were consistently smaller than those in the present study (Supplementary Material, Table S12).

Except for the 17 novel loci for height in Asians, we replicated 81 previously reported loci, mostly identified in European populations. In addition, 75 reported loci showed nominally significant association with height in the present study. These observations are in line with recent findings of high trans-ethnic replicability of GWA associations between Europeans and East Asians (9), and lend further support to the hypothesis that these two ethnic populations may share causal variants for numerous loci determining phenotypes such as adult height. Of note, six confirmed loci including *ATP13A2*, *NUDT3*, *ATF7*, *ATP5G2*, *NUP37* and *IGF1* were only identified in Asian population. For example, rs7313075 at *IGF1*, a gene known to be involved in skeletal development(16,17), was reported in previous GWAS in Asians (11) and showed a robust association with adult height in the present study ($P = 4.00 \times 10^{-47}$), but was not reported in the GIANT consortium. The index SNP is less common in Europeans (effect allele frequency, EAF = 0.10) (18) than in East Asians (EAF = 0.32), although GIANT would have >80% power to detect similarly sized genetic effect with sample size of >130 thousands subjects. Some other previously reported loci from GWAS in Europeans or Asians failed to be validated in the present study (Supplementary Material, Tables S3 and S4), probably due to sampling error or insufficient power of the current analysis.

Among the 17 novel loci, 2 loci (*PTPN14* at 1q32.2 and *PTPN9* at 15q24.2) belong to the same protein tyrosine phosphatase (PTP) family. The *PTPN9* locus has been suggestively associated with adult height in Japanese (11). In addition, another PTP family member *PTPN11* (12q24.1) was previously associated with Noonan syndrome (19), a Mendelian disorder characterized by proportionately short stature and other symptoms such as dysmorphic facial features, heart disease (20) and acute myeloid leukemia (21).

In addition, several newly identified loci might have pleiotropic effects. For example, the newly identified SNP rs174547 in *FADS1-FADS2* gene cluster, a member of the *fatty acid desaturase* (*FADS*) gene family involved in the metabolism of polyunsaturated fatty acids, shows a perfect correlation ($r^2 = 1.0$; Fig. 3) with SNPs associated with phospholipid levels (22,23), carotid intima media thickness (24), lipid metabolism (25,26), liver enzyme levels (27) and coronary heart disease (25). The index SNPs of seven other newly identified loci also strongly correlate with SNPs associated with various traits such as circulating metabolites, cardiovascular disease and certain cancers (Supplementary Material, Table S8). Our observations suggest that these genes may code for products involved in biological function on various targets.

Our pathway analyses found that seven important pathways and six molecular functions are likely to be involved in regulation of height. Among them, the Hedgehog signaling pathway, histone, transcription factor and extracellular matrix glycoprotein were found to be associated with height in previous studies (8,28,29). Another six novel pathways (sulfur metabolism, glycosylphosphatidylinositol anchor biosynthesis, fructose and mannose metabolism, the insulin signaling pathway, glutamate receptor signaling and wnt signaling pathway) and three molecular functions (nuclear hormone receptor, kinase modulator and select regulatory molecule) were nominally significant (8). Notably, the newly identified wnt signaling pathway, which includes two novel height loci (*CTBP2* and *MYC*), has been involved in body axis specification (30), appendicular skeletal development (31) and bone mineral density regulation (32,33).

In conclusion, our study identified 98 loci associated with adult height in East Asians at $P < 5 \times 10^{-8}$, including 17 novel loci. These variants together explained ~8.89% of the phenotypic variation in height in East Asians. We have shown that the majority of loci encompassing common variants related to adult height are shared between Asians and Europeans, and we have provided evidence for novel variants related to adult height in Asians. For example, according to a new study by GIANT (34) that was published when our paper was under review, the majority of the novel loci identified in our study were also detected in the

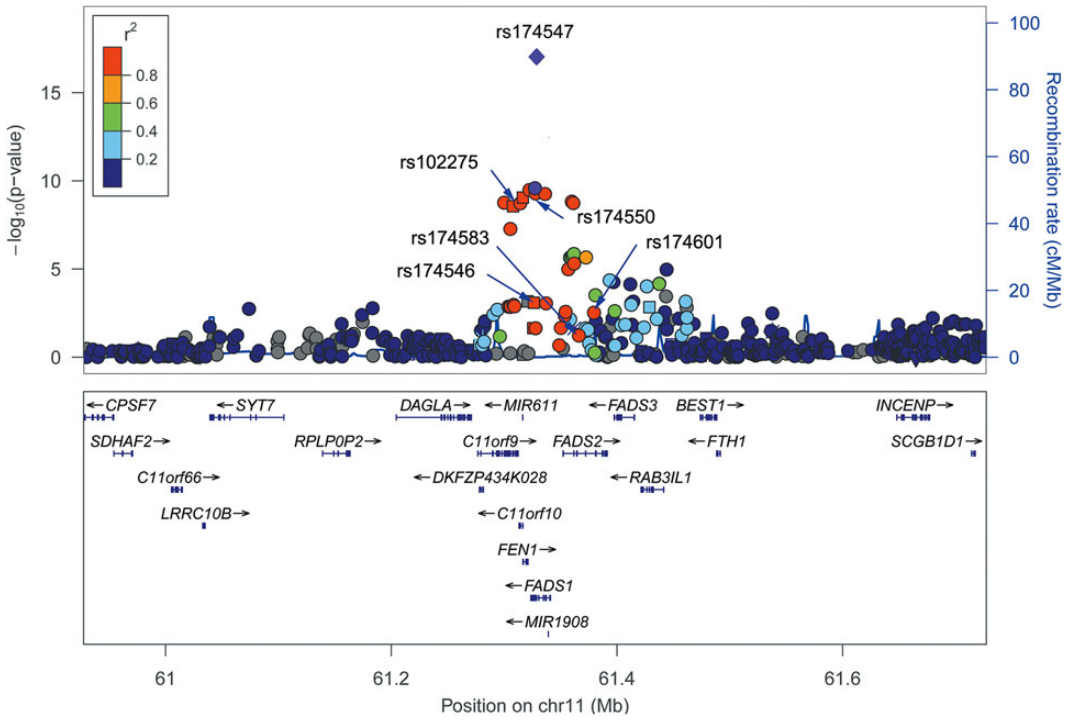

Figure 3. Regional association plots for newly identified locus *FADS1-FADS2* cluster. SNPs are plotted by their position on the chromosome against their association ($-\log_{10} P$) with adult height using Stage 1 (GWAS meta-analysis) data. The names are shown for the top SNPs (purple triangle), which were selected on the basis of combined data from all the studies (Table 1). The P-value in Stage 1 for the same SNP is denoted by a purple circle. Estimated recombination rates (from HapMap) are plotted in cyan to reflect the local LD structure. The SNPs surrounding the top SNP are colored to reflect their LD with the top SNP [using pairwise r^2 values from HapMap Han Chinese in Beijing (CHB) and Japanese in Tokyo (JPT) data]. Blue arrows indicate SNPs that were highly linkage disequilibrium with the top height SNP rs174547 ($r^2 = 1.0$) and their association with other biomedical traits or diseases. The positions of genes and exons, as well as the direction of transcription, are shown below the plots (using data from the UCSC Genome Browser). Plots were generated using LocusZoom.

Europeans, though the index variants were not identical. These observations suggest that the height-associated loci are largely shared by Asians and Europeans. Even though several loci such as *EXTL3*, *MYC* and *FADS1* were not found in GIANT, suggesting that these loci are likely to affect height only in Asians. Taken together, our results offer new insight into the genetic architecture of height in Asians and suggest novel biological pathways that may determine adult height.

Materials and Methods

Study design

This study included two stages. Stage 1 was meta-analysis of study-specific results on the association between SNPs and height from 11 GWAS that participated in the consortium and included a total of 36 227 individuals of East Asian ancestry. The promising SNPs selected from the Stage 1 meta-analysis were further examined by *in silico* (Stage 2) replication analyses. Basic information for all the participating studies is summarized in Supplementary Material, Fig. S1, Tables S1 and S2 and Note. Each participating study was approved by its local institutional review board, and informed consent was obtained from participants.

Stage 1 samples and genotyping

The sample sizes of the 11 GWAS in Stage 1 varied from 742 to 9297, and in total of 36 227 individuals were included. Approximately 41.3% of the participants were male, and the ages ranged from 18 to 95 years (all ≥18 years). For genotyping, four studies used commercially available Affymetrix arrays, six studies used Illumina platforms and one study of SGWAS used both platforms (detailed information is provided in the Supplementary Material, Table S3 and Note). To allow for the combination of data derived from different genotyping platforms and to improve coverage of the genome, genotype imputation was performed by each participating study using MACH http://www.sph.umich.edu/csg/abecasis/MACH/index.html (35), IMPUTE http://mathgen.stats.ox.ac.uk/impute/impute.html (36) or BEAGLE (37). For each SNP, the imputation software estimates its overall imputation quality score. Study-specific details are presented in Supplementary Material, Table S3.

Stage 1 statistical analysis

A uniform statistical analysis protocol was followed by each participating study. Each study measured height. To improve the normality of the height distribution and alleviate the impact of outliers, the age-adjusted z-score was calculated separately for height values for each gender in each study. The association between the SNPs and the z-score-transformed height values was analyzed with a linear additive regression model. Stratified analyses by gender were also performed by each study.

A total of 2 704 730 autosomal SNPs were meta-analyzed across 22 input files (many of the 11 cohorts had separate male–female files). We used a cutoff of $N \times \text{MAF} > 3$ to exclude the extremely rare variants present in only one or two samples that might be due to genotyping/imputation errors or private mutations. We used effects weighted by the inverse variance that

was implemented in the METAL software package (see http://www.sph.umich.edu/csg/abecasis/Metal/). The meta-analyses were carried out with all data combined and were also stratified by gender. The presence of heterogeneity across studies, populations and between genders was tested with Cochran's Q statistics (38) and I^2 statistic (39).

To correct each study for residual population stratification or cryptic relatedness, the meta-analyses were performed with genomic control correction (40) by adjusting for the study-specific inflation factor (λ), which ranged from 1.00 to 1.08 in Stage 1 (Supplementary Material, Table S3). After study-specific genomic control adjustment, the estimated inflation factor for the Stage 1 meta-analysis statistic was 1.112. The inflation factor for males and females was 1.051 and 1.050, respectively (Supplementary Material, Fig. S4).

On the basis of the Stage 1 meta-analysis on the association between SNPs and height in all participants, we selected for Stage 2 replication a total of 2162 SNPs, which included 1783 SNPs with $P < 1.0 \times 10^{-5}$ and 379 SNPs located in previously reported height-related loci that had $P > 1.0 \times 10^{-5}$ in the Stage 1 meta-analysis.

Stage 2 *in silico* replication

The 2162 SNPs selected for replication were investigated in an independent set of 57 699 individuals (53.0% males, 100% of the samples were ≥18 years old) of East Asian ancestry from five additional GWAS. The sample sizes of the five additional studies varied between 1881 and 43 386 subjects. The RIKEN study was the main source of the replication data and included 43 386 individuals. Three studies (SCES, CAGE and RIKEN) used the Illumina platform and two studies of HEXA and BGWAS used an Affymetrix array. Genotype imputation was also performed by each study using MACH http://www.sph.umich.edu/csg/abecasis/MACH/index.html (35), IMPUTE http://mathgen.stats.ox.ac.uk/impute/impute.html (36) or BEAGLE (37), as for studies included in Stage 1.

Each study individually conducted a similar analysis of the SNPs selected from Stage 1, using the same protocol as in Stage 1. The Stage 2 data were combined using effects weighted by the inverse variance meta-analysis methods as in Stage 1 without genomic control correction. The results from Stages 1 and 2 were combined and analyzed using meta-analysis methods. Quanto 1.2.4 was used for power calculation (http://biostats.usc.edu/software). The genome-wide set of association results from the GWAS meta-analysis stage of the present study could be found on AGEN website, http://www.agenconsortium.org/about.php.

Quality control procedures

The following quality control procedures were recommended for each participating study. SNPs were excluded, either in the primary analysis conducted by each participating study or at the meta-analysis stage (Supplementary Material, Table S3), if they had (i) SNPs with a call rate <95%, (ii) $P < 1.0 \times 10^{-4}$ for Hardy–Weinberg violation, (iii) SNPs with a concordance rate <95% among the duplicated QC samples, (iv) SNPs with MAF < 1%, (v) low imputation quality (for imputed SNPs; r-hat < 0.3 for MACH http://www.sph.umich.edu/csg/abecasis/MACH or proper-info <0.5 for IMPUTE http://mathgen.stats.ox.ac.uk/impute/impute.html) and (vi) samples that were potentially contaminated. Samples from individuals were removed if they had a call rate of <90%, if they showed first-degree cryptic relationships in an identity-by-descent (IBD) analysis or if they were potentially contaminated. The specific quality control procedures adopted by each study are summarized in Supplementary Material, Table S3.

Conditional analysis

To investigate the independent association of SNPs in the same locus, the conditional and joint analyses were conducted using the combined summary estimates of 1750 SNPs in the combined two-stage analysis and a reference population (SGWAS, $n = 9297$) with genotyping information, using GCTA software (http://www.complextraitgenomics.com/software/gcta/) according to the corresponding instruction. Regional association plot of the 17 novel height loci were generated using the LocusZoom (http://csg.sph.umich.edu/locuszoom/) with data from the UCSC Genome Browser (http://www.genome.ucsc.edu/).

Estimation of the explained variance

The fraction of variance explained by an individual height-associated SNP was estimated by $2\beta^2 f(1-f)$ (41), where f is the frequency of the variant and β is its additive effect estimated from the combined data. We subsequently estimated the overall fraction of variance that can be explained by the 17 novel loci and the total 98 loci identified in the current study, respectively, by using multiple linear regression model in 9297 participants from the largest individual sample (SGWAS) of our meta-analysis. The adjusted multiple regression Rsq is an estimate of variance explained.

MAGENTA pathway analysis

MAGENTA (15) is (GSEA)-like statistical test. By calculating a gene-set enrichment P-value and a FDR for a given biological pathway or gene set of interest, MAGENTA can evaluate prespecified gene sets for enrichment of modest associations with a complex disease or trait. This method does not require genotype data, making it relevant to our GWAS meta-analyses.

In our study, the pathway analysis was based on the data of meta-analysis results of Stage 1 studies, and we used gene-set annotations from the KEGG, PANTHER, INGENUITY and Gene Ontology databases, to find out whether our height loci belong to specific biological pathways, and potentially to discover new pathways that may be enriched for modest height associations not yet identified. KEGG, PANTHER and INGENUITY pathways were downloaded from the Molecular Signatures Database (MsigDB, http://www.broad.mit.edu/gsea/msigdb/collections.jsp), PANTHER molecular function gene sets were downloaded from the PANTHER website (http://www.pantherdb.org/) and Gene Ontology biological process and molecular function categories were downloaded from the Gene Ontology website (http://cvsweb.geneontology.org/).

GRAIL analysis

GRAIL (42) evaluates each gene within a confirmed height-associated locus for relatedness with all other implicated genes through text-mining in PubMed abstracts, and the degree of relatedness is assessed by text-based similarity. We conducted the unconventional model for GRAIL analysis using the hg18 build of the human genome and the December 2012 PubMed data, although bias may be generated from the identification of novel associations stimulated by GWAS. Nevertheless, the December 2012 PubMed abstracts identified 80.9% of the significant genes revealed by its 2006 equivalence. We conducted

simulations to 100 sets of randomly selected SNPs, each containing 89 SNPs and implicating 434 ± 22 (5%) genes. The median number of $P < 0.01$ hits was 6 ($P < 0.001$, the Wilcoxon signed rank test, power >99.9%, power calculated with GPower 3.1 (43) with a range of 0–29, suggesting that these height loci were functionally related with high confidence.

Supplementary Material

Supplementary Material is available at *HMG* online.

Acknowledgement

We acknowledge the participants who participated the present study and the staff who helped in recruitment.

Conflict of Interest statement. None declared.

Funding

BGWAS (Beijing Genome-wide Association Studies): This work was supported by the National Basic Research Program of China (973 Plan) (2011CB503901), the High-Tech Research and Development Program of China (863 Plan) (2012AA02A516), grants (2011BAI11B03, 2011BAI09B03) from the Ministry of Science and Technology of China. CLHNS (the Cebu Longitudinal Health and Nutrition Survey): The study was supported by National Institutes of Health grants DK078150, TW05596 and HL085144 and pilot funds from RR20649, ES10126 and DK56350. We thank the Office of Population Studies Foundation research and data collection teams and the study participants who generously provided their time for this study. CRC (the Cardiometabolic Risk in Chinese study): This study was supported by grants from the National Heart, Lung, and Blood Institute (HL071981), the National Institute of Diabetes and Digestive and Kidney Diseases (DK091718 and DK078616), the Boston Obesity Nutrition Research Center (DK46200) and United States–Israel Binational Science Foundation Grant 2011036. FAMHES (Fangchenggang Area Male Health and Examination Survey): This study was supported by the Guangxi Provincial Department of Finance and Education (2009GJCJ150); intramural funding from the Fudan-VARI Center for Genetic Epidemiology and Fudan University Institute of Urology to Z. Mo. HEXA (Health Examinee) shared control study, KARE (the Korea Association Resource study): The work was supported by a grant from the Korea Center for Disease Control and Prevention (4845-301, 4851-302, 4851-307), and intramural grant from the Korea National Institute of Health (2012-N73002-00). NHAPC (Nutrition and Health of Aging Population in China study): This study is supported by research grants including the National High Technology Research and Development Program (2009AA022704), Knowledge Innovation Program (KSCX2-EW-R-10), the National Natural Science Foundation of China (30930081, 81021002, 81170734) and the and National Key Basic Research Program of China (2012CB524900). NJ study (Nanjing GWA study): This work is funded by the China National High-Tech Research and Development Program Grant (2009AA022705), the National Key Basic Research Program Grant (2011CB503805), the National Natural Science Foundation of China (30730080, 30972541, 30901233 and 30872178). RJ study (RuiJin T2D study): the Key Laboratory for Endocrine and Metabolic Diseases of Ministry of Health (No.1994DP131044), the Sector Funds of Ministry of Health (No. 201002002), the National Key New Drug Creation and Manufacturing Program of Ministry of Science and Technology (No. 2012ZX09303006-001) and the National Natural Science Foundation of China (No. 81270877). SGWAS (Shanghai Genome-Wide Association Studies): The content of this paper is solely the responsibility of the authors and does not necessarily represent the official views of the funding agencies. The authors thank the study participants and research staff for their contributions and commitment to this project, Regina Courtney for DNA preparation, Jing He for data processing and analyses and Bethanie Rammer for editing and preparing the manuscript. This research was supported in part by U.S. National Institutes of Health grants R37CA070867, R01CA082729, R01CA124558, R01CA148667, and R01CA122364, as well as Ingram Professorship and Research Reward funds from the Vanderbilt University School of Medicine. Participating studies (grant support) in the Shanghai Genome Wide Association Studies are as follows: Shanghai Women's Health Study (R37CA070867), Shanghai Men's Health Study (R01CA82729) and Shanghai Breast and Endometrial Cancer Studies (R01CA064277 and R01CA092585). SP2 (Singapore Prospective Study Program), SCES (Singapore Chinese Eye Study) and SiMES (Singapore Malay Eye Study): the studies were supported by the Singapore Ministry of Health's National Medical Research Council under its Individual Research Grant funding scheme, the Singapore National Research Foundation under its Clinician Scientist Award and Singapore Translational Research Investigator Award funding schemes which are administered by the Singapore Ministry of Health's National Medical Research Council and the Singapore Biomedical Research Council (BMRC) individual research grant funding scheme. WHBJ (Wuhan-Beijing GWA study): This work was supported by grants from the National Basic Research Program grant (2011CB503800) and the Programme of Introducing Talents of Discipline to Universities to T. Wu and the Program for New Century Excellent Talents in University and the General Program of the National Natural Science Foundation of China (81172751) to M.H.

References

1. Perola, M., Sammalisto, S., Hiekkalinna, T., Martin, N.G., Visscher, P.M., Montgomery, G.W., Benyamin, B., Harris, J.R., Boomsma, D., Willemsen, G. *et al.* (2007) Combined genome scans for body stature in 6,602 European twins: evidence for common Caucasian loci. *PLoS Genet.*, **3**, e97.
2. Carmichael, C.M. and McGue, M. (1995) A cross-sectional examination of height, weight, and body mass index in adult twins. *J. Gerontol. A Biol. Sci. Med. Sci.*, **50**, B237–B244.
3. Kabat, G.C., Heo, M., Kamensky, V., Miller, A.B. and Rohan, T.E. (2013) Adult height in relation to risk of cancer in a cohort of Canadian women. *Int. J. Cancer*, **132**, 1125–1132.
4. Green, J., Cairns, B.J., Casabonne, D., Wright, F.L., Reeves, G. and Beral, V. (2011) Height and cancer incidence in the Million Women Study: prospective cohort, and meta-analysis of prospective studies of height and total cancer risk. *Lancet Oncol.*, **12**, 785–794.
5. Shimizu, Y., Imano, H., Ohira, T., Kitamura, A., Kiyama, M., Okada, T., Ishikawa, Y., Shimamoto, T., Yamagishi, K., Tanigawa, T. *et al.* (2013) Adult height and body mass index in relation to risk of total stroke and its subtypes: the circulatory risk in communities study. *J. Stroke Cerebrovasc. Dis.*, **23**, 667–674.
6. Rosenberg, M.A., Patton, K.K., Sotoodehnia, N., Karas, M.G., Kizer, J.R., Zimetbaum, P.J., Chang, J.D., Siscovick, D., Gottdiener, J.S., Kronmal, R.A. *et al.* (2012) The impact of height on the risk of atrial fibrillation: the Cardiovascular Health Study. *Eur. Heart J.*, **33**, 2709–2717.

7. Tripaldi, R., Stuppia, L. and Alberti, S. (2013) Human height genes and cancer. *Biochim. Biophys. Acta*, **1836**, 27–e41.
8. Lango Allen, H., Estrada, K., Lettre, G., Berndt, S.I., Weedon, M.N., Rivadeneira, F., Willer, C.J., Jackson, A.U., Vedantam, S., Raychaudhuri, S. *et al.* (2010) Hundreds of variants clustered in genomic loci and biological pathways affect human height. *Nature*, **467**, 832–838.
9. Marigorta, U.M. and Navarro, A. (2013) High trans-ethnic replicability of GWAS results implies common causal variants. *PLoS Genet.*, **9**, e1003566.
10. Kim, J.-J., Lee, H.-I., Park, T., Kim, K., Lee, J.-E., Cho, N.H., Shin, C., Cho, Y.S., Lee, J.-Y., Han, B.-G. *et al.* (2009) Identification of 15 loci influencing height in a Korean population. *J. Hum. Genet.*, **55**, 27–31.
11. Okada, Y., Kamatani, Y., Takahashi, A., Matsuda, K., Hosono, N., Ohmiya, H., Daigo, Y., Yamamoto, K., Kubo, M., Nakamura, Y. *et al.* (2010) A genome-wide association study in 19 633 Japanese subjects identified LHX3-QSOX2 and IGF1 as adult height loci. *Hum. Mol. Genet.*, **19**, 2303–2312.
12. Lei, S.-F., Yang, T.-L., Tan, L.-J., Chen, X.-D., Guo, Y., Guo, Y.-F., Zhang, L., Liu, X.-G., Yan, H., Pan, F. *et al.* (2008) Genome-wide association scan for stature in Chinese: evidence for ethnic specific loci. *Hum. Genet.*, **125**, 1–9.
13. Yang, J., Ferreira, T., Morris, A.P., Medland, S.E., Madden, P.A., Heath, A.C., Martin, N.G., Montgomery, G.W., Weedon, M.N., Loos, R.J. *et al.* (2012) Conditional and joint multiple-SNP analysis of GWAS summary statistics identifies additional variants influencing complex traits. *Nat. Genet.*, **44**, 369–375, S361–S363.
14. Wang, K., Li, M. and Bucan, M. (2007) Pathway-based approaches for analysis of genomewide association studies. *Am. J. Hum. Genet.*, **81**, 1278–1283.
15. Segre, A.V., Groop, L., Mootha, V.K., Daly, M.J. and Altshuler, D. (2010) Common inherited variation in mitochondrial genes is not enriched for associations with type 2 diabetes or related glycemic traits. *PLoS Genet.*, **6**, e1001058.
16. Clemmons, D.R. (2007) Modifying IGF1 activity: an approach to treat endocrine disorders, atherosclerosis and cancer. *Nat. Rev. Drug Discov.*, **6**, 821–833.
17. Kemp, S.F. (2009) Insulin-like growth factor-I deficiency in children with growth hormone insensitivity: current and future treatment options. *BioDrugs*, **23**, 155–163.
18. The International HapMap Consortium. (2003) The International HapMap Project. *Nature*, **426**, 789–796.
19. Tartaglia, M., Mehler, E.L., Goldberg, R., Zampino, G., Brunner, H.G., Kremer, H., van der Burgt, I., Crosby, A.H., Ion, A., Jeffery, S. *et al.* (2001) Mutations in PTPN11, encoding the protein tyrosine phosphatase SHP-2, cause Noonan syndrome. *Nat. Genet.*, **29**, 465–468.
20. Noonan, J.A. (1968) Hypertelorism with Turner phenotype. A new syndrome with associated congenital heart disease. *Am. J. Dis. Child*, **116**, 373–380.
21. Tartaglia, M., Niemeyer, C.M., Fragale, A., Song, X., Buechner, J., Jung, A., Hahlen, K., Hasle, H., Licht, J.D. and Gelb, B.D. (2003) Somatic mutations in PTPN11 in juvenile myelomonocytic leukemia, myelodysplastic syndromes and acute myeloid leukemia. *Nat. Genet.*, **34**, 148–150.
22. Schaeffer, L., Gohlke, H., Muller, M., Heid, I.M., Palmer, L.J., Kompauer, I., Demmelmair, H., Illig, T., Koletzko, B. and Heinrich, J. (2006) Common genetic variants of the FADS1 FADS2 gene cluster and their reconstructed haplotypes are associated with the fatty acid composition in phospholipids. *Hum. Mol. Genet.*, **15**, 1745–1756.
23. Tanaka, T., Shen, J., Abecasis, G.R., Kisialiou, A., Ordovas, J.M., Guralnik, J.M., Singleton, A., Bandinelli, S., Cherubini, A., Arnett, D. *et al.* (2009) Genome-wide association study of plasma polyunsaturated fatty acids in the InCHIANTI Study. *PLoS Genet.*, **5**, e1000338.
24. Demirkan, A., van Duijn, C.M., Ugocsai, P., Isaacs, A., Pramstaller, P.P., Liebisch, G., Wilson, J.F., Johansson, A., Rudan, I., Aulchenko, Y.S. *et al.* (2012) Genome-wide association study identifies novel loci associated with circulating phospho- and sphingolipid concentrations. *PLoS Genet.*, **8**, e1002490.
25. Kwak, J.H., Paik, J.K., Kim, O.Y., Jang, Y., Lee, S.H., Ordovas, J.M. and Lee, J.H. (2011) FADS gene polymorphisms in Koreans: association with omega6 polyunsaturated fatty acids in serum phospholipids, lipid peroxides, and coronary artery disease. *Atherosclerosis*, **214**, 94–100.
26. Waterworth, D.M., Ricketts, S.L., Song, K., Chen, L., Zhao, J.H., Ripatti, S., Aulchenko, Y.S., Zhang, W., Yuan, X., Lim, N. *et al.* (2010) Genetic variants influencing circulating lipid levels and risk of coronary artery disease. *Arterioscler. Thromb. Vasc. Biol.*, **30**, 2264–2276.
27. Chambers, J.C., Zhang, W., Sehmi, J., Li, X., Wass, M.N., Van der Harst, P., Holm, H., Sanna, S., Kavousi, M., Baumeister, S.E. *et al.* (2011) Genome-wide association study identifies loci influencing concentrations of liver enzymes in plasma. *Nat. Genet.*, **43**, 1131–1138.
28. Weedon, M.N., Lango, H., Lindgren, C.M., Wallace, C., Evans, D.M., Mangino, M., Freathy, R.M., Perry, J.R., Stevens, S., Hall, A.S. *et al.* (2008) Genome-wide association analysis identifies 20 loci that influence adult height. *Nat. Genet.*, **40**, 575–583.
29. Lettre, G., Jackson, A.U., Gieger, C., Schumacher, F.R., Berndt, S.I., Sanna, S., Eyheramendy, S., Voight, B.F., Butler, J.L., Guiducci, C. *et al.* (2008) Identification of ten loci associated with height highlights new biological pathways in human growth. *Nat. Genet.*, **40**, 584–591.
30. Sokol, S.Y. (1999) Wnt signaling and dorso-ventral axis specification in vertebrates. *Curr. Opin. Genet. Dev.*, **9**, 405–410.
31. Zhou, Z.Q., Shung, C.Y., Ota, S., Akiyama, H., Keene, D.R. and Hurlin, P.J. (2011) Sequential and coordinated actions of c-Myc and N-Myc control appendicular skeletal development. *PLoS ONE*, **6**, e18795.
32. Kung, A.W., Xiao, S.M., Cherny, S., Li, G.H., Gao, Y., Tso, G., Lau, K.S., Luk, K.D., Liu, J.M., Cui, B. *et al.* (2010) Association of JAG1 with bone mineral density and osteoporotic fractures: a genome-wide association study and follow-up replication studies. *Am. J. Hum. Genet.*, **86**, 229–239.
33. Piek, E., Sleumer, L.S., van Someren, E.P., Heuver, L., de Haan, J.R., de Grijs, I., Gilissen, C., Hendriks, J.M., van Ravestein-van Os, R.I., Bauerschmidt, S. *et al.* (2010) Osteo-transcriptomics of human mesenchymal stem cells: accelerated gene expression and osteoblast differentiation induced by vitamin D reveals c-MYC as an enhancer of BMP2-induced osteogenesis. *Bone*, **46**, 613–627.
34. Wood, A.R., Esko, T., Yang, J., Vedantam, S., Pers, T.H., Gustafsson, S., Chu, A.Y., Estrada, K., Luan, J., Kutalik, Z. *et al.* (2014) Defining the role of common variation in the genomic and biological architecture of adult human height. *Nat. Genet.*, **46**, 1173–1186.
35. Li, Y., Willer, C., Sanna, S. and Abecasis, G. (2009) Genotype imputation. *Annu. Rev. Genomics Hum. Genet.*, **10**, 387–406.
36. Marchini, J., Howie, B., Myers, S., McVean, G. and Donnelly, P. (2007) A new multipoint method for genome-wide association studies by imputation of genotypes. *Nat. Genet.*, **39**, 906–913.

37. Browning, B.L. and Browning, S.R. (2009) A unified approach to genotype imputation and haplotype-phase inference for large data sets of trios and unrelated individuals. *Am. J. Hum. Genet.*, **84**, 210–223.
38. Cochran, W.G. (1954) The combination of estimates from different experiments. *Biometrics*, **10**, 101–129.
39. Higgins, J.P. and Thompson, S.G. (2002) Quantifying heterogeneity in a meta-analysis. *Stat. Med.*, **21**, 1539–1558.
40. Devlin, B. and Roeder, K. (1999) Genomic control for association studies. *Biometrics*, **55**, 997–1004.
41. Park, J.H., Wacholder, S., Gail, M.H., Peters, U., Jacobs, K.B., Chanock, S.J. and Chatterjee, N. (2010) Estimation of effect size distribution from genome-wide association studies and implications for future discoveries. *Nat. Genet.*, **42**, 570–575.
42. Raychaudhuri, S., Plenge, R.M., Rossin, E.J., Ng, A.C., Purcell, S. M., Sklar, P., Scolnick, E.M., Xavier, R.J., Altshuler, D. and Daly, M.J. (2009) Identifying relationships among genomic disease regions: predicting genes at pathogenic SNP associations and rare deletions. *PLoS Genet.*, **5**, e1000534.
43. Faul, F., Erdfelder, E., Lang, A.G. and Buchner, A. (2007) G*Power 3: a flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behav. Res. Methods*, **39**, 175–191.