

SCIENTIFIC REPORTS



OPEN

Association of human height-related genetic variants with familial short stature in Han Chinese in Taiwan

Ying-Ju Lin^{1,2}, Wen-Ling Liao^{3,4}, Chung-Hsing Wang⁵, Li-Ping Tsai⁶, Chih-Hsin Tang⁷, Chien-Hsiun Chen^{2,8}, Jer-Yuarn Wu^{2,8}, Wen-Miin Liang⁹, Ai-Ru Hsieh⁹, Chi-Fung Cheng⁹, Jin-Hua Chen¹⁰, Wen-Kuei Chien¹¹, Ting-Hsu Lin¹, Chia-Ming Wu¹, Chiu-Chu Liao¹, Shao-Mei Huang¹ & Fuu-Jen Tsai^{1,2,5,12}

Human height can be described as a classical and inherited trait model. Genome-wide association studies (GWAS) have revealed susceptible loci and provided insights into the polygenic nature of human height. Familial short stature (FSS) represents a suitable trait for investigating short stature genetics because disease associations with short stature have been ruled out in this case. In addition, FSS is caused only by genetically inherited factors. In this study, we explored the correlations of FSS risk with the genetic loci associated with human height in previous GWAS, alone and cumulatively. We systematically evaluated 34 known human height single nucleotide polymorphisms (SNPs) in relation to FSS in the additive model ($p < 0.00005$). A cumulative effect was observed: the odds ratios gradually increased with increasing genetic risk score quartiles ($p < 0.001$; Cochran-Armitage trend test). Six affected genes—*ZBTB38*, *ZNF638*, *LCORL*, *CABLES1*, *CDK10*, and *TSEN15*—are located in the nucleus and have been implicated in embryonic, organismal, and tissue development. In conclusion, our study suggests that 13 human height GWAS-identified SNPs are associated with FSS risk both alone and cumulatively.

Human height can be described as a classical, inherited, and polygenic trait model. The ultimate phenotype represents the outcome of genetics and developmental biology, including the enlargement of bone length and the sizes of tissues and organ sizes. Undoubtedly, the developmental processes may be affected by environmental factors, including nutrition¹, exercise², diseases^{3–5}, and infections^{6,7}. However, ~80% of the variance in height among individuals can also be explained by genetic variation^{8–10}. Modern high-density genotyping platforms have now made it possible to identify candidate genes across the whole genome in relation to complex traits and diseases^{11–13}.

Genetic factors that contribute to the genetic architecture of human height are widely identified by means of genome-wide screening methods^{14–24}. Human height-related susceptibility loci related to the protein tyrosine phosphatase family, insulin-like growth factor, and skeletal growth have been reported in Asian populations^{18, 19, 21, 22, 25}. In addition, predisposing loci determining height have also been discovered in European populations in genes related to skeletal development, mitosis, fibroblast growth factors, the WNT/ β -catenin pathway, chondroitin sulfate-related genes, mammalian target of rapamycin, osteoglycin, binding of hyaluronic acid, Hedgehog

¹Genetic Center, Department of Medical Research, China Medical University Hospital, Taichung, Taiwan. ²School of Chinese Medicine, China Medical University, Taichung, Taiwan. ³Graduate Institute of Integrated Medicine, China Medical University, Taichung, Taiwan. ⁴Center for Personalized Medicine, China Medical University Hospital, Taichung, Taiwan. ⁵Children's Hospital of China Medical University, Taichung, Taiwan. ⁶Department of Pediatrics, Buddhist Tzu Chi General Hospital, Taipei Branch, Taipei, Taiwan. ⁷Graduate Institute of Biomedical Sciences, China Medical University, Taichung, Taiwan. ⁸Institute of Biomedical Sciences, Academia Sinica, Taipei, Taiwan. ⁹Graduate Institute of Biostatistics, School of Public Health, China Medical University, Taichung, Taiwan. ¹⁰Biostatistics Center and School of Public Health, Taipei Medical University, Taipei, Taiwan. ¹¹National Applied Research Laboratories, National Center for High-performance Computing, Hsinchu, Taiwan. ¹²Department of Biotechnology and Bioinformatics, Asia University, Taichung, Taiwan. Ying-Ju Lin and Wen-Ling Liao contributed equally to this work. Correspondence and requests for materials should be addressed to F.-J.T. (email: d0704@mail.cmuh.org.tw)

signaling, the extracellular matrix, and cancer-associated pathways^{14–17}. The results of genome-wide association studies (GWAS) have thus contributed new information on the susceptibility loci related to human height and provide insights into the extreme polygenic nature of human height.

Short stature is defined as a body height below the 3rd percentile for the corresponding age and gender of the population. Among people affected by short stature, 37% have familial short stature (FSS), 27% have a constitutional growth delay, 17% have both FSS and a constitutional growth delay, 9% have a systemic disease, 5% have an endocrine condition, and the cause is unknown in the remaining 5% of cases^{26–28}. Representing the majority of cases of short stature, FSS is characterized by a height below the 3rd percentile in the population, a normal annual growth rate, normal bone age, family history of short stature, normal onset of puberty, and normal results of biochemical analyses. FSS is also called “genetic short stature” because the growth conditions are normal but the body height is below the 3rd percentile and there is a family history of short stature. FSS represents a very suitable study object for research on short stature genetics because disease association has been ruled out for the short stature in FSS. FSS is caused only by genetically inherited factors. Thus, we initiated a pilot study to search for genetic loci that may predispose individuals to FSS. We explored the associations of FSS risk in Taiwan with genetic loci that have been related to human height according to previous GWAS, both alone and cumulatively.

Results

GWAS-determined genetic variants related to human height and FSS risk in the Han Chinese population of Taiwan. To identify the susceptible single nucleotide polymorphisms (SNPs) associated with FSS risk, we chose genetic SNPs that have been related to human height according to published GWAS, which were considered to be the closest genes potentially associated with FSS. There were 1,033 genetic SNPs reported in 14 human height GWAS conducted to date (Table S1). Of these, 34 SNPs were ultimately identified as candidate susceptibility SNPs for FSS according to the selection criteria [call rates of genotyping for both FSS and controls >95%, genotypes for genetic variants in controls conforming to Hardy-Weinberg equilibrium ($p > 0.05$), and p -value for the additive model $< 5.00E-05$ (0.05/1,033)]. The 34 SNPs in 15 closest genes identified in human height GWAS are summarized in Table 1 and Supplementary Table S2. As shown in Table 1, the minor allele frequencies (MAFs) of these 34 genetic SNPs were similar to the MAFs in Han Chinese from Beijing (CHB) NCBI GRCh38.p2 assembly) from the online database of the dbSNP website (HAPMAP-CHB; <http://www.ncbi.nlm.nih.gov/projects/SNP/index.html>). The pair-wise linkage disequilibrium (LD) between the 34 SNPs was also estimated (Supplementary Tables S2–S5 and Figure S1).

The excluded SNPs were those with strong LD ($D' > 0.8$). According to the LD results, we then re-selected these SNPs for calculation of the cumulative effect. The re-selected SNPs included 13 SNPs in the 13 closest genes: rs1046934 in *TSEN15*, rs3791679 in *EFEMP1*, rs3771381 in *ZNF638*, rs10935120 in *CEP63*, rs7632381 in *ZBTB38*, rs13131350 in *LCORL*, rs6845999 in *HHIP*, rs4240326 in *ANAPC10*, rs6470764 in *GSDMC*, rs12338076 in *QSOX2*, rs4842838 in *ADAMTSL3*, rs258324 in *CDK10*, and rs4308051 in *CABLES1*. The association was statistically evaluated by the odds ratio (OR) and its corresponding 95% confidence interval (CI) in the additive genetic model. As shown in Table 2, these 13 genetic SNPs showed a statistically significant association with FSS risk (p -value < 0.00005).

Cumulative effects of the 13 human height genetic SNPs on FSS risk. In accordance with the potential additive inheritance model for the individual SNP association analysis (Table 2), we subdivided the three genotypes of each SNP into three groups: risk allele homozygote (the risk genotype is coded as “2”), risk allele heterozygote (the risk genotype is coded as “1”), and Non-risk allele homozygote (the non-risk genotype is coded as “0”) (Table S6, Figures S2 and S3). We calculated the multi-locus genetic risk score (GRS) for each individual by summing the number of risk alleles (0/1/2) for each of the 13 SNPs weighted by their estimated effect sizes (log OR) (Table S6 and Figure S3). Estimates of the association between the weighted GRS divided into quartiles (Q1–Q4) and FSS were calculated using a logistic regression model; individuals in genetic risk score Q1 served as the reference. As shown in Table 3, there was a continuous increase in FSS risk with the cumulative weighted GRS quartile numbers ($p < 0.001$; Cochran-Armitage trend test). That is, compared with individuals in Q1, those identified in Q2, Q3, and Q4 had a significant association with increased FSS risk, with an increasing risk for each quartile (Table 3). These results suggested a cumulative effect of these 13 SNPs on FSS risk.

Functional analysis. We applied these 13 SNPs re-selected from the LD results to possible pathway mapping and further functional validation to explore possible functional relations among the genes affected by these 13 SNPs. The network analysis identified a single cluster of 35 genes that included 11 associated genes discovered in this study (Figure S4) related to functions of embryonic development, organismal development, and tissue development. Six of these genes are all located in the nucleus: *ZBTB38*, *ZNF638*, *LCORL*, *CABLES1*, *CDK10*, and *TSEN15*.

Discussion

Human height GWAS have led to the discovery of thousands of novel genetic variants^{14–24}. In this work, we initiated our FSS study on a single ethnic group (Han Chinese) and investigated the associations of FSS risk with genetic variants previously related to human height in GWAS, both alone and cumulatively. We identified a continuous increase in FSS risk with increasing genetic risk score quartiles, suggesting a cumulative effect of these 13 SNPs on FSS risk. These data suggest that FSS represents a suitable research object for investigating short stature genetics. To our knowledge, this is the first study reporting the genetic profile of FSS.

The LD analysis revealed several SNPs in strong LD that have been linked to adult height in previous genetic studies, mapping to genes such as *COLGALT2* and *TSEN15*, *EFEMP1* (implicated in bone and cartilage development pathways, including the constitutions of the extracellular matrix), *ZBTB38*, *LCORL*, *HHIP*, *ANAPC10*,

No.	rs ID	Gene	Chr.	Position	MAF (Allele) ^a	MAF (%) in CHB ^a	MAF (%) in controls	Call rate (%)	P for controls HWE
1	rs1926872	COLGALT2	1	184049341	G	47.1%	51.5%	100.0%	0.34
2	rs1046934	TSEN15	1	184054395	C	47.1%	51.8%	100.0%	0.35
3	rs3791679	EFEMP1	2	55869757	T	20.7%	24.9%	99.8%	0.08
4	rs3791675	EFEMP1	2	55884174	G	22.2%	25.3%	100.0%	0.26
5	rs3771381	ZNF638	2	71333535	T	42.7%	39.7%	100.0%	0.52
6	rs10935120	CEP63	3	134514250	A	18.9%	15.0%	99.8%	0.85
7	rs6440003	ZBTB38	3	141375367	A	35.6%	34.8%	100.0%	0.21
8	rs7632381	ZBTB38	3	141387221	C	34.4%	35.6%	100.0%	0.24
9	rs1344672	ZBTB38	3	141406863	C	33.3%	35.5%	100.0%	0.29
10	rs9825379	ZBTB38	3	141418193	A	17.1%	19.9%	99.9%	0.40
11	rs7678436	DCAF16	4	17796343	A	34.1%	32.4%	99.6%	0.69
12	rs16895802	NCAPG	4	17814266	G	12.2%	12.0%	100.0%	0.60
13	rs6842303	LCORL	4	17852432	T	45.0%	47.4%	100.0%	0.86
14	rs13131350	LCORL	4	17875864	G	22.2%	24.6%	100.0%	0.69
15	rs16896276	LCORL	4	18013533	A	46.7%	47.4%	100.0%	0.51
16	rs2011603	LCORL	4	18023861	G	45.1%	47.9%	100.0%	0.57
17	rs17720281	NA	4	144622624	T	20.0%	24.9%	100.0%	0.64
18	rs6845999	HHIP	4	144644674	T	26.8%	24.7%	99.9%	0.66
19	rs4240326	ANAPC10	4	144918112	A	36.7%	30.0%	100.0%	0.55
20	rs6823268	ANAPC10	4	145061411	G	17.1%	17.3%	100.0%	0.31
21	rs4733724	GSDMC	8	129711482	A	30.5%	28.2%	100.0%	0.10
22	rs6470764	GSDMC	8	129713419	C	33.0%	28.2%	100.0%	0.10
23	rs10858250	QSOX2	9	136227369	G	24.7%	23.4%	100.0%	0.90
24	rs12338076	QSOX2	9	136229894	C	32.4%	31.4%	99.9%	0.44
25	rs2401171	ADAMTSL3	15	83888924	T	29.1%	23.5%	100.0%	0.09
26	rs10906982	ADAMTSL3	15	83899406	T	28.9%	23.7%	99.8%	0.07
27	rs7183263	ADAMTSL3	15	83904289	T	29.1%	23.4%	100.0%	0.10
28	rs11259936	ADAMTSL3	15	83911830	A	29.1%	23.4%	100.0%	0.10
29	rs4842838	ADAMTSL3	15	83913372	G	28.9%	23.3%	99.8%	0.08
30	rs258324	CDK10	16	89687847	A	34.4%	32.5%	100.0%	0.50
31	rs4800452	CABLES1	18	23147647	C	17.6%	16.9%	100.0%	0.50
32	rs4369779	CABLES1	18	23155444	T	15.8%	14.4%	99.9%	0.53
33	rs4308051	CABLES1	18	23155497	T	15.8%	14.4%	100.0%	0.50
34	rs8094261	CABLES1	18	23166764	G	13.4%	14.8%	100.0%	0.44

Table 1. Information of genetic SNPs identified from GWAS of human height. Abbreviations: SNP, single nucleotide polymorphism; GWAS, genome-wide association studies; chr., chromosome; MAF, minor allele frequency; CHB, Han Chinese in Beijing, China; HWE, Hardy-Weinberg equilibrium; NA, not available. MAF (Allele) and MAF (%) in CHB (NCBI GRCh38.p2 assembly) were downloaded from the online database of dbSNP website (HAPMAP-CHB; <http://www.ncbi.nlm.nih.gov/projects/SNP/index.html>).

GSDMC, QSOX2, ADAMTSL3, and CABLES1^{14, 16–20, 25}. Other SNPs were found to be independent, located near the ZNF638 and CEP63 genes. CEP63 recruits the cell cycle regulatory protein CDK1 to the centrosome, and has been associated with regulation of mitotic entry, centrosome amplification, and genome maintenance, thereby affecting cell proliferation. Only one SNP was identified for chromosome 16, related to CDK10.

Network cluster analysis revealed functions of these 13 genes that are mainly related to embryonic development, organismal development, and tissue development. Two of the genes found to be associated with FSS belong to the zinc finger protein family: ZBTB38 and ZNF638. There were seven relevant SNPs in ZBTB38: rs6440003, rs6763931, rs724016, rs7632381, rs1344672, rs9825379, and rs10513137 that have previously been reported in human height GWAS^{12, 16, 18, 19, 25}. Moreover, genetic variants in ZBTB38, including rs724016, rs1582874, rs11919556, rs6440006, rs7612543, rs62282002, and rs18651435, have also been implicated in idiopathic short stature (ISS)²⁹. ZBTB38 encodes a zinc finger- and BTB domain-containing protein 38. It belongs to the family of zinc finger proteins and thus serves as a zinc finger transcriptional activator that binds methylated DNA; such potential suppression of the transcription of methylated regions might affect adult stature by regulating the production of insulin-like growth factor-II³⁰. Therefore, ZBTB38 may affect adult height, ISS, and FSS by regulating the protein expression of IGF-II. In addition, ZBTB38 also regulates the transcription of tyrosine hydroxylase, the rate-limiting enzyme in catecholamine synthesis³¹.

Moreover, the SNP rs3771381 in ZNF638, encoding zinc finger protein 68, was also associated with FSS risk. ZNF638 genetic variants have been reported in human height GWAS^{25, 32}. ZNF638 is a nucleoplasmic protein

rs ID	Gene	Chr.	Risk allele	Homozygous dominant/Heterozygous/ Homozygous recessive genotype frequency		Additive model	
				Familial short stature cases (N = 978)	Controls (N = 1129)	OR (95%CI)	P
rs1046934	<i>TSEN15</i>	1	A	287/500/189	270/548/311	1.32 (1.16–1.49)	1.06E-05
rs3791679	<i>EFEMP1</i>	2	C	629/311/34	625/443/59	1.39 (1.19–1.61)	1.96E-05
rs3771381	<i>ZNF638</i>	2	A	278/495/205	416/530/183	1.31 (1.16–1.48)	1.83E-05
rs10935120	<i>CEP63</i>	3	A	623/317/38	816/285/26	1.43 (1.22–1.68)	1.27E-05
rs7632381	<i>ZBTB38</i>	3	T	479/424/75	459/536/134	1.35 (1.18–1.54)	1.17E-05
rs13131350	<i>LCORL</i>	4	G	432/433/112	639/424/66	1.55 (1.36–1.78)	2.06E-10
rs6845999	<i>HHIP</i>	4	C	639/306/32	637/425/66	1.41 (1.22–1.64)	6.86E-06
rs4240326	<i>ANAPC10</i>	4	G	571/350/57	557/466/106	1.37 (1.19–1.59)	5.95E-06
rs6470764	<i>GSDMC</i>	8	T	590/336/51	593/435/101	1.35 (1.18–1.54)	2.57E-05
rs12338076	<i>QSOX2</i>	9	A	564/330/83	536/475/117	1.33 (1.16–1.52)	2.92E-05
rs4842838	<i>ADAMTSL3</i>	15	G	484/400/94	652/424/51	1.42 (1.23–1.63)	8.51E-07
rs258324	<i>CDK10</i>	16	C	528/388/62	510/505/114	1.37 (1.19–1.56)	7.32E-06
rs4308051	<i>CABLES1</i>	18	T	636/304/38	831/272/26	1.43 (1.22–1.69)	1.39E-05

Table 2. Association between human height related genetic SNPs and familial short stature occurrence. Abbreviations: SNP, single nucleotide polymorphism; Chr., chromosome; FSS, familial short stature; OR, odds ratio; CI, confidence interval. OR calculation was conducted according to the defined risk alleles.

Weighted genetic risk score (wGRS) quartile	FSS (%)	Controls (%)	OR	95% CI	P value
	N = 978	N = 1129			
	N (%)	N (%)			
Q1	144 (14.7)	412 (36.5)	Ref.	Ref.	Ref.
Q2	199 (20.3)	282 (25.0)	2.02	1.55–2.63	<0.001
Q3	254 (26.0)	278 (24.6)	2.61	2.03–3.37	<0.001
Q4	381 (39.0)	157 (13.9)	6.94	5.32–9.06	<0.001
Cochran-Armitage Trend Test					<0.001

Table 3. Association between genetic risk score from the 13 human height genetic loci and FSS risk. Abbreviations: FSS, familial short stature; OR, odds ratio; CI, confidence interval; Ref., reference. The 13 human height genetic loci and the respective risk genotypes were shown in Table S3. The controls were the individuals whose body height was 75% above the general population in Taiwan. Weighted genetic risk score quartile 1 (Q1): $0 < \text{wGRS} \leq 11.88$; Weighted genetic risk score quartile 2 (Q2): $11.88 < \text{wGRS} \leq 13.50$; Weighted genetic risk score quartile 3 (Q3): $13.50 < \text{wGRS} \leq 15.05$; Weighted genetic risk score quartile 4 (Q4): $15.05 < \text{wGRS}$.

that binds to cytidine-rich sequences in double-stranded DNA³³. This zinc finger protein is a member of the large class of transcription factors that participates in skeletal development³⁴ and regulates adipocyte differentiation³⁵. *ZNF638* has been reported to be a transcriptional coregulator of the early regulator of adipogenesis via induction of PPAR γ in cooperation with CCAAT/enhancer binding proteins (C/EBPs)^{35,36}.

LCORL encodes a ligand-dependent nuclear receptor corepressor-like protein and is a transcription factor. Polymorphisms in this gene are associated with measures of skeletal frame size and adult height in human and in livestock animals^{24,25,37–40}. SNPs affecting the transcription or translation of *LCORL* may result in up- or down-regulation of the expression of genes involved in growth³⁸, and may affect bone formation, skeletal frame size, as well as body height; thus, further functional investigations of these variants are warranted. One SNP (rs13131350) within this gene was associated with FSS in our study. This SNP was also reported in a meta-analysis study of GWAS of adult height in East Asians²⁵. Regulation of *LCORL* protein expression may affect bone formation, skeletal frame size, as well as body height; thus, further functional investigations are warranted.

There were two cell cycle regulation-related genes—*CABLES1* and *CDK10*—associated with FSS in our study. According to the present study, four SNPs in *CABLES1* are associated with FSS in Han Chinese in Taiwan: rs4800452, rs4369779, rs4308051, and rs8094261, which have also been reported in human height GWAS^{14,18,23,25}. *CABLES1* encodes a Cdk5 and Abl enzyme substrate protein 1, and serves as a candidate tumor suppressor that negatively regulates cell growth by inhibiting cyclin-dependent kinases^{41,42}. Loss of *CABLES1* protein expression has been observed at high frequency in human colorectal, lung, ovarian, and endometrial hyperplasia and in endometrial cancers^{42–45}. *CABLES1* protein overexpression induces apoptosis and inhibits cell growth by stabilizing p21 and decreasing cyclin-dependent kinase 2 kinase activity, thereby affecting cell proliferation⁴⁶.

The other cell cycle-regulatory gene is *CDK10*. We found one SNP (rs258324) located in *CDK10* associated with FSS, which was also reported in a human height GWAS²⁵. *CDK10* encodes cyclin-dependent kinase 10 and belongs to the CDK subfamily of the family of Ser/Thr protein kinases, which are necessary for cell cycle

progression. *CDK10* plays important roles in cellular proliferation and in regulation of the G2-M phase of the cell cycle^{47, 48}. Therefore, *CABLES1* and *CDK10* may affect adult height and FSS by regulating cell survival, cell cycle, and cellular proliferation.

We identified one SNP (rs1046934) in *TSEN15* associated with FSS. This SNP has been reported in a human height GWAS study¹⁴. *TSEN15* encodes a highly conserved subunit of a tRNA-splicing endonuclease that catalyzes the removal of introns from tRNA precursors^{49, 50}. tRNA splicing is a fundamental process for cell growth and division and is highly conserved among vertebrates. Mutations in *TSEN15* cause neurogenetic disorders, including pontocerebellar hypoplasia and progressive microcephaly^{50–52}. These findings are suggestive of *TSEN15*'s importance in brain development. However, the roles of *TSEN15* in adult height and FSS remain unclear and further functional characterization is needed.

In conclusion, our study suggests that 13 SNPs related to human height (according to previous GWAS) are associated with FSS risk alone and cumulatively. This is the first report on the genetic profile of FSS. Such information may be helpful for further research on short stature in the Han Chinese population, and for future functional studies dealing with growth and development. Confirmatory studies with a larger sample size and with functional characterization of these candidate genes would be warranted in the near future.

Methods

Study participants. The case population consisted of 978 individuals with FSS sequentially enrolled from the Children's Hospital, China Medical University, Taichung, Taiwan, from August 1999 to September 2014. All participants were unrelated Han Chinese children who were residents of central Taiwan and fulfilled the diagnostic criteria of FSS^{26, 28}. The criteria for FSS in our association study were (1) height less than the 3rd percentile of the corresponding population, (2) normal annual growth rate, (3) bone age appropriate for chronologic age, (4) a family history of short stature, (5) normal onset of puberty, and (6) normal results of physical examination. The controls consisted of 1,129 subjects who were randomly selected from the Taiwan Biobank (http://www.twbiobank.org.tw/new_web/index.php). The criteria for controls for inclusion in our association study were (1) no history of FSS diagnosis, (2) body height above that of the top 75th percentile of the general population in Taiwan; and (3) age <61 years. All the participating FSS cases and controls were of Han Chinese origin, which constitutes 98% of the Taiwanese population. The ethnic background was assigned in accordance with the results of self-reported questionnaires. All research was performed in accordance with the relevant guidelines and regulations. The study protocol was approved by the institutional review board and the ethics committee of Human Studies Committee of China Medical University Hospital. Written informed consent was obtained from the participants, their parent, or legal guardian in accordance with the institutional requirements and the Declaration of Helsinki principles.

SNP selection, genotyping, and quality control. We initially included genetic SNPs listed as supplementary files from previous human height GWAS^{14–24}. The genetic resources for the human height GWAS are shown in Table S7. The repeated SNPs were removed and the resulting 1,033 SNPs from 14 GWAS of human height are listed in Table S1. Genotyping of these 1,033 SNPs was performed in both cases and controls. Genomic DNA was extracted from peripheral blood leukocytes according to standard protocols (Qiagen Genomic DNA Isolation Kit; Qiagen, Taichung, Taiwan). The DNA concentration and quality were confirmed on a NanoDrop 1000 spectrophotometer (Thermo Fisher Scientific, Taichung, Taiwan). SNPs of each sample were genotyped using a custom-designed VeraCode GoldenGate Genotyping Assay System (Illumina; <http://www.illumina.com/>)^{53–55}. Genotypic data were quality controlled, and SNPs were excluded from further analysis if (1) the MAF was less than 5% in the controls (the Han Chinese population); (2) the total call rate was less than 95% for both cases and controls; or (3) the SNP significantly departed from Hardy-Weinberg equilibrium proportions for the controls ($p < 0.05$). Then SNPs that passed the correction test were considered to be significant ($P = 0.05/1,033 = 0.00005$).

Statistical analysis. Hardy-Weinberg equilibrium for genotype distributions of controls in this study was evaluated by the goodness-of-fit χ^2 test (Table 1).

Alleles were identified by genotyping and direct counting. The difference of allelic frequency in the additive model between the cases and controls were measured by ORs with 95% CIs using logistical regression models (Table 2). Deviation from the additive model was also tested for 34 selected SNPs and a p -value for the DOMDEV test $< 1.47 \times 10^{-3}$ was considered significant ($p = 0.05/34$) (Table S2). All statistical analyses were performed in the SPSS software, v12.0 for Windows (IBM, Armonk, NY, USA).

For the analysis of haplotype blocks, we used the Lewontin D' and R^2 measure to evaluate the intermarker coefficient of linkage disequilibrium in both controls and FSS cases⁵⁶. The confidence interval for LD was estimated using a resampling procedure and was used to construct the haplotype blocks^{57, 58}.

References

- Paciorek, C. J., Stevens, G. A., Finucane, M. M. & Ezzati, M. Nutrition Impact Model Study, G. Children's height and weight in rural and urban populations in low-income and middle-income countries: a systematic analysis of population-representative data. *Lancet Glob Health* **1**, e300–309, doi:10.1016/S2214-109X(13)70109-8 (2013).
- Izard, R. M., Fraser, W. D., Negus, C., Sale, C. & Greeves, J. P. Increased density and periosteal expansion of the tibia in young adult men following short-term arduous training. *Bone* **88**, 13–19, doi:10.1016/j.bone.2016.03.015 (2016).
- Tarquinio, D. C. *et al.* Growth failure and outcome in Rett syndrome: specific growth references. *Neurology* **79**, 1653–1661, doi:10.1212/WNL.0b013e31826e9a70 (2012).
- Bali, D. S., Chen, Y. T., Austin, S. & Goldstein, J. L. In *GeneReviews(R)* (eds R. A. Pagon *et al.*) (1993).
- Gunnell, D. *et al.* Height, leg length, and cancer risk: a systematic review. *Epidemiol Rev* **23**, 313–342 (2001).

6. Stawerska, R. *et al.* Prevalence of autoantibodies against some selected growth and appetite-regulating neuropeptides in serum of short children exposed to *Candida albicans* colonization and/or *Helicobacter pylori* infection: the molecular mimicry phenomenon. *Neuro Endocrinol Lett* **36**, 458–464 (2015).
7. Bejerman, N. *et al.* Complete genome sequence of a new enamovirus from Argentina infecting alfalfa plants showing dwarfism symptoms. *Arch Virol* **161**, 2029–2032, doi:10.1007/s00705-016-2854-3 (2016).
8. Lettre, G. Recent progress in the study of the genetics of height. *Hum Genet* **129**, 465–472, doi:10.1007/s00439-011-0969-x (2011).
9. Lanktree, M. B. *et al.* Meta-analysis of Dense Genecentric Association Studies Reveals Common and Uncommon Variants Associated with Height. *Am J Hum Genet* **88**, 6–18, doi:10.1016/j.ajhg.2010.11.007 (2011).
10. Visscher, P. M., Hill, W. G. & Wray, N. R. Heritability in the genomics era—concepts and misconceptions. *Nat Rev Genet* **9**, 255–266, doi:10.1038/nrg2322 (2008).
11. Plenge, R. M. *et al.* TRAF1-C5 as a risk locus for rheumatoid arthritis—a genomewide study. *N Engl J Med* **357**, 1199–1209, doi:10.1056/NEJMoa073491 (2007).
12. Weeks, D. E. *et al.* Age-related maculopathy: a genomewide scan with continued evidence of susceptibility loci within the 1q31, 10q26, and 17q25 regions. *Am J Hum Genet* **75**, 174–189, doi:10.1086/422476 (2004).
13. Samani, N. J. *et al.* Genomewide association analysis of coronary artery disease. *N Engl J Med* **357**, 443–453, doi:10.1056/NEJMoa072366 (2007).
14. Lango Allen, H. *et al.* Hundreds of variants clustered in genomic loci and biological pathways affect human height. *Nature* **467**, 832–838, doi:10.1038/nature09410 (2010).
15. Lettre, G. *et al.* Identification of ten loci associated with height highlights new biological pathways in human growth. *Nat Genet* **40**, 584–591, doi:10.1038/ng.125 (2008).
16. Wood, A. R. *et al.* Defining the role of common variation in the genomic and biological architecture of adult human height. *Nat Genet* **46**, 1173–1186, doi:10.1038/ng.3097 (2014).
17. Weedon, M. N. *et al.* Genome-wide association analysis identifies 20 loci that influence adult height. *Nat Genet* **40**, 575–583, doi:10.1038/ng.121 (2008).
18. Kim, J. J. *et al.* Identification of 15 loci influencing height in a Korean population. *J Hum Genet* **55**, 27–31, doi:10.1038/jhg.2009.116 (2010).
19. Okada, Y. *et al.* 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, doi:10.1093/hmg/ddq091 (2010).
20. Gudbjartsson, D. F. *et al.* Many sequence variants affecting diversity of adult human height. *Nat Genet* **40**, 609–615, doi:10.1038/ng.122 (2008).
21. Yang, J. *et al.* Conditional and joint multiple-SNP analysis of GWAS summary statistics identifies additional variants influencing complex traits. *Nat Genet* **44**, 369–375, S361–363, doi:10.1038/ng.2213 (2012).
22. Cho, Y. S. *et al.* A large-scale genome-wide association study of Asian populations uncovers genetic factors influencing eight quantitative traits. *Nat Genet* **41**, 527–534, doi:10.1038/ng.357 (2009).
23. Chan, Y. *et al.* Genome-wide Analysis of Body Proportion Classifies Height-Associated Variants by Mechanism of Action and Implicates Genes Important for Skeletal Development. *Am J Hum Genet* **96**, 695–708, doi:10.1016/j.ajhg.2015.02.018 (2015).
24. Soranzo, N. *et al.* Meta-analysis of genome-wide scans for human adult stature identifies novel Loci and associations with measures of skeletal frame size. *PLoS Genet* **5**, e1000445, doi:10.1371/journal.pgen.1000445 (2009).
25. He, M. *et al.* Meta-analysis of genome-wide association studies of adult height in East Asians identifies 17 novel loci. *Hum Mol Genet* **24**, 1791–1800, doi:10.1093/hmg/ddu583 (2015).
26. Lindsay, R., Feldkamp, M., Harris, D., Robertson, J. & Rallison, M. Utah Growth Study: growth standards and the prevalence of growth hormone deficiency. *J Pediatr* **125**, 29–35, doi:S0022-3476(94)70117-2 [pii] (1994).
27. Moore, K. C., Donaldson, D. L., Ideus, P. L., Gifford, R. A. & Moore, W. V. Clinical diagnoses of children with extremely short stature and their response to growth hormone. *J Pediatr* **122**, 687–692 (1993).
28. Sisley, S., Trujillo, M. V., Khoury, J. & Backeljauw, P. Low incidence of pathology detection and high cost of screening in the evaluation of asymptomatic short children. *J Pediatr* **163**, 1045–1051, doi:10.1016/j.jpeds.2013.04.002 (2013).
29. Wang, Y. *et al.* An SNP of the ZBTB38 gene is associated with idiopathic short stature in the Chinese Han population. *Clin Endocrinol (Oxf)* **79**, 402–408, doi:10.1111/cen.12145 (2013).
30. Filion, G. J. *et al.* A family of human zinc finger proteins that bind methylated DNA and repress transcription. *Mol Cell Biol* **26**, 169–181, doi:10.1128/MCB.26.1.169-181.2006 (2006).
31. Kiefer, H. *et al.* ZENON, a novel POZ Kruppel-like DNA binding protein associated with differentiation and/or survival of late postmitotic neurons. *Mol Cell Biol* **25**, 1713–1729, doi:10.1128/MCB.25.5.1713-1729.2005 (2005).
32. Hao, Y. *et al.* Genome-wide association study in Han Chinese identifies three novel loci for human height. *Hum Genet* **132**, 681–689, doi:10.1007/s00439-013-1280-9 (2013).
33. Inagaki, H. *et al.* A large DNA-binding nuclear protein with RNA recognition motif and serine/arginine-rich domain. *J Biol Chem* **271**, 12525–12531 (1996).
34. Ganss, B. & Jheon, A. Zinc finger transcription factors in skeletal development. *Crit Rev Oral Biol Med* **15**, 282–297 (2004).
35. Meruvu, S., Hugendubler, L. & Mueller, E. Regulation of adipocyte differentiation by the zinc finger protein ZNF638. *J Biol Chem* **286**, 26516–26523, doi:10.1074/jbc.M110.212506 (2011).
36. Du, C., Ma, X., Meruvu, S., Hugendubler, L. & Mueller, E. The adipogenic transcriptional cofactor ZNF638 interacts with splicing regulators and influences alternative splicing. *J Lipid Res* **55**, 1886–1896, doi:10.1194/jlr.M047555 (2014).
37. Hirschhorn, J. N. & Lettre, G. Progress in genome-wide association studies of human height. *Horm Res* **71**(Suppl 2), 5–13, doi:10.1159/000192430 (2009).
38. Lindholm-Perry, A. K. *et al.* Association, effects and validation of polymorphisms within the NCAPG - LCORL locus located on BTA6 with feed intake, gain, meat and carcass traits in beef cattle. *BMC Genet* **12**, 103, doi:10.1186/1471-2156-12-103 (2011).
39. Takasuga, A. PLAG1 and NCAPG-LCORL in livestock. *Anim Sci J* **87**, 159–167, doi:10.1111/asj.12417 (2016).
40. Tozaki, T. *et al.* Sequence variants of BIEC2-808543 near LCORL are associated with body composition in Thoroughbreds under training. *J Equine Sci* **27**, 107–114, doi:10.1294/jes.27.107 (2016).
41. Wu, C. L. *et al.* Cables enhances cdk2 tyrosine 15 phosphorylation by Wee1, inhibits cell growth, and is lost in many human colon and squamous cancers. *Cancer Res* **61**, 7325–7332 (2001).
42. Zukerberg, L. R. *et al.* Loss of cables, a cyclin-dependent kinase regulatory protein, is associated with the development of endometrial hyperplasia and endometrial cancer. *Cancer Res* **64**, 202–208 (2004).
43. Dong, Q. *et al.* Loss of cables, a novel gene on chromosome 18q, in ovarian cancer. *Mod Pathol* **16**, 863–868, doi:10.1097/01.MP.0000084434.88269.0A (2003).
44. Park, D. Y. *et al.* The Cables gene on chromosome 18q is silenced by promoter hypermethylation and allelic loss in human colorectal cancer. *Am J Pathol* **171**, 1509–1519, doi:10.2353/ajpath.2007.070331 (2007).
45. Tan, D. *et al.* Loss of cables protein expression in human non-small cell lung cancer: a tissue microarray study. *Hum Pathol* **34**, 143–149, doi:10.1053/hupa.2003.26 (2003).
46. Shi, Z. *et al.* Cables1 complex couples survival signaling to the cell death machinery. *Cancer Res* **75**, 147–158, doi:10.1158/0008-5472.CAN-14-0036 (2015).

47. Kasten, M. & Giordano, A. Cdk10, a Cdc2-related kinase, associates with the Ets2 transcription factor and modulates its transactivation activity. *Oncogene* **20**, 1832–1838, doi:10.1038/sj.onc.1204295 (2001).
48. Sergere, J. C., Thuret, J. Y., Le Roux, G., Carosella, E. D. & Leteurtre, F. Human CDK10 gene isoforms. *Biochem Biophys Res Commun* **276**, 271–277, doi:10.1006/bbrc.2000.3395 (2000).
49. Paushkin, S. V., Patel, M., Furia, B. S., Peltz, S. W. & Trotta, C. R. Identification of a human endonuclease complex reveals a link between tRNA splicing and pre-mRNA 3' end formation. *Cell* **117**, 311–321 (2004).
50. Cassandrini, D. *et al.* Pontocerebellar hypoplasia: clinical, pathologic, and genetic studies. *Neurology* **75**, 1459–1464, doi:10.1212/WNL.0b013e3181f88173 (2010).
51. Breuss, M. W. *et al.* Autosomal-Recessive Mutations in the tRNA Splicing Endonuclease Subunit TSEN15 Cause Pontocerebellar Hypoplasia and Progressive Microcephaly. *Am J Hum Genet* **99**, 228–235, doi:10.1016/j.ajhg.2016.05.023 (2016).
52. Alazami, A. M. *et al.* Accelerating novel candidate gene discovery in neurogenetic disorders via whole-exome sequencing of prescreened multiplex consanguineous families. *Cell Rep* **10**, 148–161, doi:10.1016/j.celrep.2014.12.015 (2015).
53. Tindall, E. A. *et al.* Interpretation of custom designed Illumina genotype cluster plots for targeted association studies and next-generation sequence validation. *BMC Res Notes* **3**, 39, doi:10.1186/1756-0500-3-39 (2010).
54. Lin, Y. J. *et al.* Association between GRIN3A gene polymorphism in Kawasaki disease and coronary artery aneurysms in Taiwanese children. *PLoS One* **8**, e81384, doi:10.1371/journal.pone.0081384 (2013).
55. Lin, Y. J. *et al.* Coronary artery aneurysms occurrence risk analysis between Kawasaki disease and LRP1B gene in Taiwanese children. *Biomedicine (Taipei)* **4**, 10, doi:10.7603/s40681-014-0010-5 (2014).
56. Barrett, J. C., Fry, B., Maller, J. & Daly, M. J. Haploview: analysis and visualization of LD and haplotype maps. *Bioinformatics* **21**, 263–265, doi:10.1093/bioinformatics/bth457 (2005).
57. Gabriel, S. B. *et al.* The structure of haplotype blocks in the human genome. *Science* **296**, 2225–2229, doi:10.1126/science.1069424 (2002).
58. Lin, Y. J. *et al.* HLA-E gene polymorphism associated with susceptibility to Kawasaki disease and formation of coronary artery aneurysms. *Arthritis Rheum* **60**, 604–610, doi:10.1002/art.24261 (2009).

Acknowledgements

This study was supported by grants from the China Medical University (CMU102-PH-01), the China Medical University Hospital (DMR-105-031 and DMR-105-098), the National Science Council, the Ministry of Science and Technology, Taiwan (NSC 102-2314-B-039 -011 -MY3, MOST 103-2320-B-039 -006 -MY3, and MOST 105-2314-B-039 -037 -MY3), and China Medical University under the Aim for Top University Plan of the Ministry of Education, Taiwan. We thank the National Center for Genome Medicine of the National Core Facility Program for Biotechnology, Ministry of Science and Technology, for technical and bioinformatics support. We also thank Drs. Kuan-Teh Jeang and Willy W.L. Hong for technical help and suggestions.

Author Contributions

Y.J.L., W.L.L., and F.J.T. conceived of and designed the experiments. W.L.L., C.H.T., J.H.C., W.K.C., T.H.L., C.M.W., C.C.L., and S.M.H. conducted the experiments. W.L.L., W.M.L., A.R.H., C.F.C., J.H.C., and W.K.C. analyzed the data. C.H.W., L.P.T., C.H.C., and J.Y.W. contributed reagents, materials, and analytical tools. Y.J.L. wrote the manuscript. All the coauthors have read and approved the final version of the manuscript.

Additional Information

Supplementary information accompanies this paper at doi:10.1038/s41598-017-06766-z

Competing Interests: The authors declare that they have no competing interests.

Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2017