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

Ying-Ju Lin, Wen-Ling Liao, Chung-Hsing Wang, Li-Ping Tsai, Chih-Hsin Tang, Chien-Hsiun Chen, Jer-Yuarn Wu, Wen-Miin Liang, Ai-Ru Hsieh, Chi-Fung Cheng, Jin-Hua Chen, Wen-Kuei Chien, Ting-Hsu Lin, Chia-Ming Wu, Chiu-Chu Liao, Shao-Mei Huang, & Fuu-Jen Tsai

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 nutrition, exercise, diseases, and infections. However, ~80% of the variance in height among individuals can also be explained by genetic variation. 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.

Genetic factors that contribute to the genetic architecture of human height are widely identified by means of genome-wide screening methods. 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. Genetic factors that contribute to the genetic architecture of human height are widely identified by means of genome-wide screening methods. 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. Additionally, 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.
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 <5.0E-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, rs37791679 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, organisal 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, and TSEN15.
rs1926872, COLGALT2
rs10406934, TSEN15
rs3791679, EFEMP1
rs3791675, EFEMP1
rs3771381, ZNF638
rs10935120, CEPI3
rs6440003, ZBTB38
rs7632381, ZBTB38
rs1344672, ZBTB38
rs9825379, ZBTB38
rs7678436, DCAF16
rs16895802, NCAPG
rs6842303, LCORL
rs13131350, LCORL
rs16896276, LCORL
rs2011603, LCORL
rs17720281, NA
rs6845999, HHIP
rs4240326, ANAPC10
rs6823268, ANAPC10
rs4733724, GSDMC
rs6470764, GSDMC
rs10858250, QSOX2
rs12338076, QSOX2
rs2401171, ADAMTSL3
rs10906982, ADAMTSL3
rs7183263, ADAMTSL3
rs11259936, ADAMTSL3
rs4842838, ADAMTSL3
rs2583224, CDK10
rs4800452, CABLES1
rs4369779, CABLES1
rs4308051, CABLES1
rs8094261, CABLES1

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 are genes that are mainly related to embryonic development, organ development, and tissue development. The two genes found to be associated with FSS belong to the zinc finger protein family: ZBTB38 and ZNF638. ZBTB38 is associated with FSS risk.

Network cluster analysis revealed functions of these 13 genes that are mainly related to embryonic development, organ development, and tissue development. Two of the genes found to be associated with FSS belong to the zinc finger protein family: ZBTB38 and ZNF638. ZBTB38 is associated with FSS risk.

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. ZNF638 is a nucleoplasmic protein.
that binds to cytidine-rich sequences in double-stranded DNA. This zinc finger protein is a member of the large class of transcription factors that participates in skeletal development and regulates adipocyte differentiation. ZNF638 has been reported to be a transcriptional coregulator of the early regulator of adipogenesis via induction of PPARγ expression. LCORL encodes a lipodystrophy-related nuclear receptor coactivator-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 livestock animals. SNPs affecting the transcription or translation of LCORL may result in up- or down-regulation of the expression of genes involved in growth, 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 of GWAS of adult height in East Asians. 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. 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. Loss of CABLES1 protein expression may affect bone formation, skeletal frame size, as well as body height; thus, further functional investigations are warranted. 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 GWAS. 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 and cell cycle control. CDK10 is a tumor suppressor gene that negatively regulates cell growth by inhibiting cyclin-dependent kinases and decreasing cell proliferation.

### Table 2: Association between genetic risk score from the 13 human height genetic loci and FSS risk

| 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. |
|---|

| rs ID  | Gene       | Chr. | Risk allele | Homozygous dominant/Heterozygous/ Homozygous recessive genotype frequency | Additive model |
|--------|------------|------|-------------|-----------------------------------|---------------|
| rs1046934 | TSEN15     | 1    | A           | 287/50/189                       | 1.32 (1.16–1.49) | 1.06E-05 |
| rs3791679  | EEFM1      | 2    | C           | 629/31/34                       | 1.39 (1.19–1.61) | 1.96E-05 |
| rs3771381  | ZNF638     | 2    | A           | 278/49/205                       | 1.31 (1.16–1.48) | 1.83E-05 |
| rs10935120 | CEP3       | 3    | A           | 623/31/38                       | 1.43 (1.22–1.68) | 1.27E-05 |
| rs7632381  | ZBTB38     | 3    | T           | 479/42/74                       | 1.35 (1.18–1.54) | 1.17E-05 |
| rs13131350 | LCORL      | 4    | G           | 422/43/112                       | 1.55 (1.36–1.78) | 2.06E-10 |
| rs6845999  | HHIP       | 4    | G           | 569/30/62                       | 1.41 (1.22–1.64) | 6.86E-06 |
| rs4240326  | ANAPC10    | 4    | G           | 571/35/57                       | 1.37 (1.19–1.59) | 5.95E-06 |
| rs6470764  | GSDMC      | 8    | T           | 590/33/51                       | 1.35 (1.18–1.54) | 2.57E-05 |
| rs12338076 | QSOX2      | 9    | A           | 564/33/83                       | 1.33 (1.16–1.52) | 2.92E-05 |
| rs4842388  | ADAMTS3    | 15   | G           | 484/40/94                       | 1.42 (1.23–1.63) | 8.51E-07 |
| rs258324   | CDK10      | 16   | C           | 528/34/62                       | 1.37 (1.19–1.56) | 7.32E-06 |
| rs4308051  | CABLES1    | 18   | T           | 636/30/38                       | 1.43 (1.22–1.69) | 1.39E-05 |

### 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 3. The controls were the individuals whose body height was 75% above the general population in Taiwan. Weighted genetic risk score quartile 1 (Q1): 0 < wGRS < 11.88; Weighted genetic risk score quartile 2 (Q2): 11.88 < wGRS < 13.50; Weighted genetic risk score quartile 3 (Q3): 13.50 < wGRS < 15.05; Weighted genetic risk score quartile 4 (Q4): 15.05 < wGRS. |
|---|

| Weighted genetic risk score (wGRS) quartile | FSS (%) | Controls (%) | OR | 95% CI | P value |
|---|---|---|---|---|---|
| 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 |
progression. CDK10 plays important roles in cellular proliferation and in regulation of the G2-M phase of the cell cycle. 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 study. TSEN15 encodes a highly conserved subunit of a tRNA-splicing endonuclease that catalyzes the removal of introns from tRNA precursors. 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. 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. 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. 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). 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 cases. The confidence interval for LD was estimated using a resampling procedure and was used to construct the haplotype blocks.

References

1. 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).
2. Izard, R. M., Fray, W. D., Négus, 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).
3. Tarquinio, D. C. et al. Growth failure and outcome in Rett syndrome: specific growth references. Neurology 79, 1653–1661, doi:10.1212/WNL.0b013e3182e9e70 (2012).
4. Bali, D. S., Chen, Y. T., Austin, S. & Goldstein, J. L. In GeneReview(R) (eds R. A. Pagon et al.) (1993).
5. 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. Lango Allen, H. et al. Meta-analysis of Dense Genecentric Association 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.

12. Weeks, D. E.

13. Wood, A. R.

14. Lango Allen, H.

15. Lettre, G.

16. Weedon, M. N.

17. Weedon, M. N.

18. Kim, J. J.

19. Filion, G. J.

20. Kim, J. J.

21. Yang, J.

22. Cho, Y. S.

23. Chan, Y.

24. Ganss, B. & Jheon, A.

25. He, M.

26. Lindsay, R., Feldkamp, M., Harris, D., Robertson, J. & Rallison, M.

27. Utah Growth Study: growth standards and the prevalence of short children exposed to Candida albicans colonization and/or Helicobacter pylori infection: the molecular mimicry phenomenon.

28. Sisley, S., Trujillo, M. V., Khoury, J. & Backeljauw, P.

29. Filion, G. J.

30. Filion, G. J.

31. Filion, G. J.

32. Filion, G. J.

33. Filion, G. J.

34. Filion, G. J.

35. Filion, G. J.

36. Du, C., Ma, X., Meruvu, S., Hugendubler, L. & Mueller, E.

37. Du, C., Ma, X., Meruvu, S., Hugendubler, L. & Mueller, E.

38. Du, C., Ma, X., Meruvu, S., Hugendubler, L. & Mueller, E.

39. Du, C., Ma, X., Meruvu, S., Hugendubler, L. & Mueller, E.

40. Du, C., Ma, X., Meruvu, S., Hugendubler, L. & Mueller, E.

41. Du, C., Ma, X., Meruvu, S., Hugendubler, L. & Mueller, E.

42. Zukerberg, L. R.

43. Tang, D.

44. Dong, Q.

45. Tan, D.

46. Shigehara, Y. et al. The 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. Tan, D. et al. Loss of cancer protein expression in human non-small cell lung cancer: a tissue microarray study. *Hum Pathol* **34**, 143–149, doi:10.1016/j.humpath.2003.02.003 (2003).

48. 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, V. 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. Haplovie: analysis and visualization of LD and haplotype maps. Bioinformatics 21, 263–265, doi:10.1093/bioinformatics/btl457 (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.

© The Author(s) 2017