

SCIENTIFIC REPORTS



OPEN

Genetic association study of common variants in *TGFB1* and *IL-6* with developmental dysplasia of the hip in Han Chinese population

Wenlong Ma¹, Zhuqing Zha³, Ke Chen¹, Honggan Chen¹, Yixin Wu¹, Jianbing Ma², Sixiang Zeng⁴, Liqiang Zhi² & Shuxin Yao²

Developmental dysplasia of the hip (DDH) is a congenital or developmental deformation or misalignment of the hip joint that is affected by environmental and genetic factors. Recently, polymorphisms in both *TGFB1* and *IL-6* have been identified as being significantly associated with hip osteoarthritis in Caucasians. In this study, we conducted a case-control study involving 4,206 Han Chinese individuals to investigate the effects of *TGFB1* and *IL-6* on the disease status and severity of DDH. A total of 32 single-nucleotide polymorphisms (SNPs) were selected to ensure coverage of the two genetic loci. We found SNP rs1800470 in *TGFB1* (OR = 1.255, $P = 0.0004$) and rs1800796 (OR = 0.84, $P = 0.0228$) in *IL-6* to be significantly associated with DDH in this cohort. Further haplotype-based analysis replicated this significant result. Another SNP in *IL-6*, rs1800796, showed a marginally significant association with DDH. As a non-synonymous SNP, rs1800470 alters the amino acid sequence of the polypeptide encoded by *TGFB1*; however, bioinformatics analyses revealed that this SNP has limited functional significance. No significant results were obtained in an association study focusing on the severity of DDH and epistasis analysis. Our findings support an important role for *TGFB1* in the risk of DDH. Further research is needed to validate the weak association between rs1800796 in *IL-6* and DDH.

Developmental dysplasia of the hip (DDH) is a congenital or developmental deformation or misalignment of the hip joint that endangers skeletal development in children¹. As one of the most common skeletal congenital anomalies, DDH affects 1 in 1000 live births in the Caucasian population². In China, the incidence rate of DDH ranges from 1 to 10 in 1000 live births^{3,4}. Overall, this condition is five times more frequent in females than in males⁵. The severity of DDH ranges from a mild form of acetabular dysplasia to a moderate form of subluxation of the hips and even to the severe form of complete dislocation of the hips¹. As a complex developmental disorder, DDH is affected by both environmental and genetic factors. Early gene-association mapping studies have identified several susceptibility loci for DDH. For example, a genome-wide linkage scan based on a Japanese family with acetabular dysplasia (a mild form of DDH) and osteoarthritis identified a peak region at chromosome 13q22⁶. Although no genome-wide association study focusing on DDH has been reported to date, but several well-designed candidate gene-based association mapping studies have identified several genes, including *CX3CR1*⁷, *DKK1*⁸, *HOXD9*⁹ and *PAPPA2*¹⁰. In addition to studies focusing on disease status, some early analyses have focused on the severity of DDH, identifying polymorphism markers in such genes as *ASPN*¹¹, *GDF5*¹², *HOXB9*¹³, and *TBX4*¹⁴ as being significantly associated with DDH severity.

TGFB1 encodes transforming growth factor beta 1 (TGF- β 1), a member of the transforming growth factor beta superfamily of cytokines¹⁵. As a secreted protein, TGF- β 1 performs many cellular functions, including the control of cell growth, proliferation, differentiation and apoptosis¹⁵. The gene *IL-6* encodes an interleukin that

¹Department of Hip Injury and Disease, Orthopedic Hospital of Henan Province, Qiming South Road No.82, Luoyang, Henan, China. ²Department of Joint Surgery, Honghui Hospital, Xi'an Jiaotong University Health Science Center, Xi'an, Shaanxi, China. ³The Third Clinical Medical College of Guangzhou, University of Traditional Chinese Medicine, Guangzhou, Guangdong, China. ⁴Department of Orthopedics, the First Affiliated Hospital, Xi'an Jiaotong University, Xi'an, Shaanxi, China. Wenlong Ma and Zhuqing Zha contributed equally to this work. Correspondence and requests for materials should be addressed to L.Z. (email: zhiliqiang2011@126.com) or S.Y. (email: popgold@163.com)

Characteristics	Discovery (N = 1,488)		Validation (N = 2,718)	
	DDH patients	Healthy controls	DDH patients	Healthy controls
Number	373	1,115	691	2,027
Age, mean \pm SD (months)	23.6 \pm 21.4	70.3 \pm 44.4	20.8 \pm 15.6	66.7 \pm 31.6
Male/Female	79/294	232/883	108/583	326/1,701
Acetabular dysplasia	18		42	
Unilateral	12	NA	28	NA
Bilateral	6	NA	14	NA
Subluxation of the hips	82		131	
Unilateral	62	NA	102	NA
Bilateral	20	NA	29	NA
Dislocation of the hips	293		518	
Unilateral	186	NA	378	NA
Bilateral	87	NA	140	NA

Table 1. Characteristics of DDH patients and healthy controls. DDH: developmental dysplasia of the hips; SD: standard deviation; NA: not available.

can act as both a pro-inflammatory cytokine and an anti-inflammatory myokine¹⁶. Both TGF- β 1 and IL-6 are pro-inflammatory cytokines involved in the pathogenesis of hip osteoarthritis¹⁷, which is a major consequence of DDH in adulthood, and both proteins are also involved in the bone remodelling process¹⁸. Dickinson *et al.* investigated the mRNA levels of *TGF β 1* in different tissues, and the findings suggested a role in growth and differentiation in cartilage, endochondral and membrane bone, and skin¹⁹. Furthermore, polymorphisms in *TGF β 1* and *IL-6* have been identified as significantly associated with hip osteoarthritis in studies based on Caucasian populations^{20,21}. However, most of these early studies suffered from the lacks of replication and insufficient statistical power which may be due to the sample size and small effect size or a combination of both.

In this study, we conducted a case-control study based on 4,206 Han Chinese subjects to investigate the potential effects of *TGF β 1* and *IL-6* on the disease status and severity of DDH in this population. Thirty-two single-nucleotide polymorphisms (SNPs) were selected to ensure coverage of the two genetic loci. The main purpose of our study was to replicate some existing results in Han Chinese population. Our study explored specific linkage disequilibrium (LD) structure of Han Chinese population on *TGF β 1* and *IL-6*, and identified Han Chinese specific connection between the polymorphisms within *TGF β 1* and *IL-6* and DDH.

Methods

Subjects. A total of 373 patients (294 females and 79 males) with DDH and 1,115 controls (883 females and 232 males) recruited from Luoyang Orthopaedic Hospital were included as discovery samples, and 691 patients (583 females and 108 males) with DDH and 2,027 controls (1,701 females and 326 males) enrolled from Xi'an Honghui Hospital and the First Affiliated Hospital of Xi'an Jiaotong University were included as validation samples. The study subjects recruited in this study were unrelated Han Chinese individuals from the city of Luoyang in Henan Province and Xi'an in Shaanxi Province. Patients were diagnosed by medical examination with radiographic evidence, and all had unilateral or bilateral DDH. The severity of DDH was defined according to three grades: instability; subluxation; dislocation. The controls were identified by a detailed inquiry of their medical history and physical examination, and none had a history or symptoms of DDH. Subjects with any systemic syndrome were excluded from our study. Written informed consent was obtained from the study subjects. Related clinical characteristics and demographic data for the subjects are summarized in Table 1. The study protocol conformed to the ethical guidelines of the 1975 Declaration of Helsinki and was approved by the Ethics Committee of Luoyang Orthopaedic Hospital, Xi'an Honghui Hospital and the First Affiliated Hospital of Xi'an Jiaotong University.

SNP Selection and Genotyping. Using 1000 Genomes CHB (Han Chinese in Beijing) data²², we selected tagging SNPs for all SNPs in the *IL6* and *TGF β 1* gene regions with minor allele frequencies (MAFs) \geq 0.01. Haploview was utilized for the SNP selection process; SNP tagging was performed by Tagger, a built-in algorithm of Haploview²³. The threshold we used for tagging was $r^2 \leq$ 0.8, which is also the default setting of Tagger. Finally, 32 tag SNPs covering the *IL6* (8 tag SNPs) and *TGF β 1* (24 tag SNPs) gene regions were included in the study (Supplemental Table S1). Haploview was also utilized to visualize LD structure^{24,25}. The LD structure based on the data for the discovery stage is shown in Fig. 1 and Supplemental Figure S1.

Genomic DNA was isolated from peripheral blood using a Tiangen DNA extraction kit (Tiangen Biotech Co. Ltd, Beijing, China) according to the manufacturer's protocol. SNP genotyping was performed using a Sequenom MassARRAY platform with iPLEX GOLD chemistry (Sequenom, San Diego, CA, USA) based on the manufacturer's protocols. The results were processed using Sequenom Typer 4.0 software²⁶, and genotype data were generated from the samples. Genotyping was conducted by laboratory personnel blinded to the case-control status, and the genotyping results, data entry and statistical analyses were independently reviewed by two authors. We randomly re-performed the analysis on 5% of the sample, with a concordance of 100%.

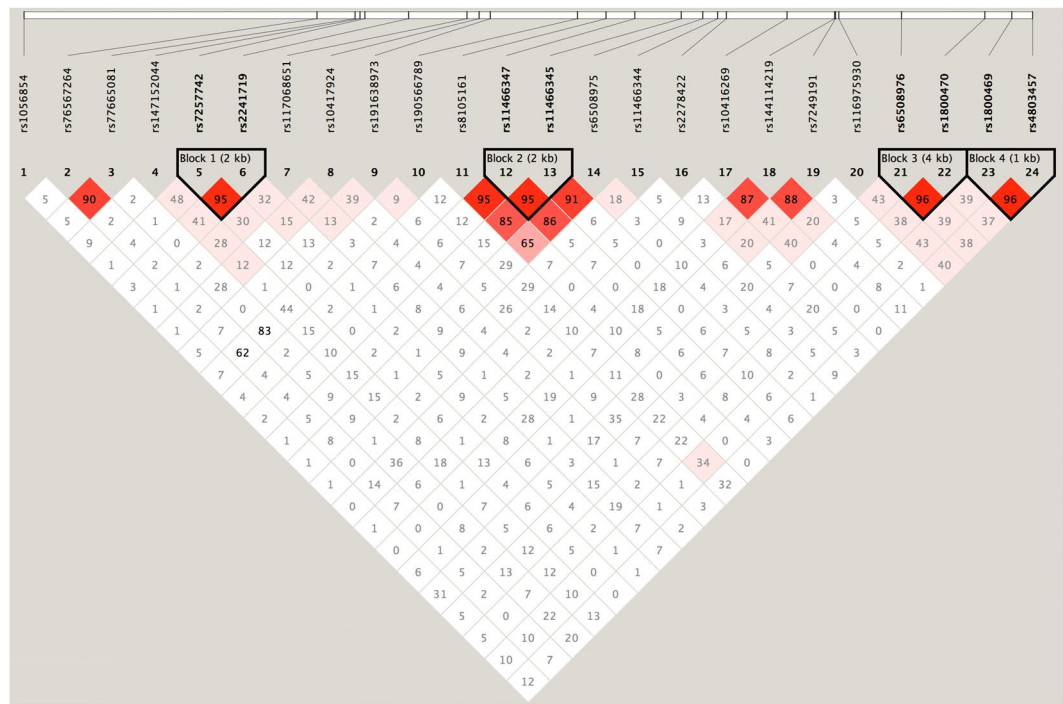


Figure 1. Linkage disequilibrium (LD) patterns of the region around the *TGFBI* gene in the Han Chinese population. The LD blocks are indicated as shaded matrices, and four LD blocks were found.

Statistical Analyses. MAFs and Hardy-Weinberg equilibrium of the genotyped SNPs were calculated and tested in the discovery stage using Plink²⁷. Logistic models were fit to assess potential associations between the 32 selected SNPs and DDH disease status in the discovery stage, including age and sex as covariates for each model. Genotypes of the SNPs were coded in three models: additive, dominant and recessive. For the additive model, the genotype was coded as 0, 1, or 2 for the three genotypes AA, Aa, and aa (where a is the minor allele), respectively. For the dominant model, the genotype was coded as 0, 1, or 1; for the recessive model, the genotype was coded as 0, 0, or 1. For SNPs that showed nominal significance in the discovery stage, genotyping of subjects was conducted in the validation stage. In addition, we also included those SNPs that were in strong LD ($r^2 > 0.9$) with nominally significant SNPs in the validation stage. To enhance the statistical power, we tested the selected markers using the combined sample from both the discovery and validation stages. Logistic models were also fit to evaluate the potential effects of selected markers, as in the discovery stage; in addition to age and sex, we included the batch number of sample recruitment as a covariate to adjust for this potential confounding factor. Logistic models were fit and tested using Plink²⁷. In addition to the single-marker-based analysis, we conducted haplotype-based analysis using data from the discovery stage. LD blocks based on the 32 SNPs with the sample from the discovery stage were constructed using methods proposed by Gabriel *et al.*²⁸. Essentially, 95% confidence bounds on D' are generated and each comparison is considered “strong LD” when the confidence bounds have an upper bound ≥ 0.98 and a lower bound ≥ 0.7 ; a block is created if 95% of informative comparisons are “strong LD”. Haplotype based analysis can be more powerful than single marker based association analysis especially when the haplotypes were formed by several highly correlated SNPs. Therefore we tested the haplotypes of the LD blocks their significance using Plink²⁷.

To evaluate the potential effects of different types of DDH on our significant findings, we separately conducted single-marker-based analyses using DDH samples of different types, i.e., acetabular dysplasia, subluxation of hips and dislocation of hip. In addition, to test whether significant SNPs are also indicators of the severity of the disorder, we dichotomized our DDH cases as two categories: the severe form, which included samples with dislocation of the hips only, and the mild to moderate form, which included samples of acetabular dysplasia and subluxation of the hips. We then performed logistic regression for SNPs that were significantly associated with the disease status of DDH. To assess potential gene-by-gene interaction between *IL6* and *TGFBI*, we also conducted interaction analyses using logistic models by adding a multiple term (sex, age and stage number were also included in the models as covariates). Both the stratification and case-only severity analyses were performed using Plink²⁷, and interaction analyses were performed using R project²⁹. Bonferroni corrections were applied to all analyses described above to adjust for multiple comparisons.

Bioinformatics Analyses. Several bioinformatic databases and tools were utilized to evaluate the functional significance of the targeted SNPs. The open bioinformatics platform Galaxy was utilized to extract basic information for selected SNPs³⁰. RegulomeDB was employed to evaluate the potential functional significance of selected SNPs³¹. Analyses using Polyphen2³² and SIFT³³ were performed to examine potential effects of non-synonymous

CHR	SNP	BP	Effect allele	Other allele	Discovery (N = 1,488)			Combined Sample (N = 4,206)		
					OR	95%CI	P	OR	95%CI	P
7	rs1800796	22726627	G	C	0.79	[0.62,0.99]	0.0427	0.84	[0.73,0.98]	0.0228
7	rs2069837	22728408	G	T	1.08	[0.83,1.39]	0.5652	1.02	[0.87,1.20]	0.7794
19	rs6508976	41348769	G	T	1.03	[0.84,1.26]	0.7539	0.99	[0.88,1.13]	0.9292
19	rs1800470	41353016	G	A	1.37	[1.12,1.68]	0.0023	1.26	[1.11,1.42]	0.0004

Table 2. Single-marker-based analyses of 4 SNPs that were genotyped in samples (additive model). The significant SNP is indicated in bold. The *P* value threshold used in the combined sample was $0.05/2 = 0.025$.

CHR	SNP	BP	A1	Dislocation of the Hips (N = 3,933)		Acetabular Dysplasia (N = 3,202)		Subluxation of the Hips (N = 3,355)	
				OR	P	OR	P	OR	P
7	rs1800796	22726627	G	0.82	0.0138	0.67	0.0963	1.01	0.9234
7	rs2069837	22728408	G	0.99	0.8965	0.81	0.4142	1.22	0.1426
19	rs6508976	41348769	G	0.99	0.8886	1.04	0.8413	1.08	0.5063
19	rs1800470	41353016	G	1.28	0.0004	1.29	0.2030	1.24	0.0570

Table 3. Stratification analyses based on different types of DDH for SNPs genotyped in combined samples. Significant results are highlighted in bold.

SNPs on protein function. In addition, we used the protein-protein interaction database STRING³⁴ to explore features of the gene-gene networks surrounding our candidate genes.

Results

Association Analyses. Single-marker-based association analysis identified 2 of the 32 SNPs as nominally significant in samples during the discovery stage (rs1800470, OR = 1.37, $P = 0.0023$, *TGFBI*; rs1800796, OR = 0.79, $P = 0.0427$, *IL6*). Four SNPs, including two other in strong LD with the significant SNPs, were genotyped in the validation sample. Analyses based on the combined sample from both the discovery and validation stages showed that two SNPs, rs1800470 (OR = 1.255, $P = 0.0004$) and rs1800796 (OR = 0.84, $P = 0.0228$), were significantly associated with the disease status of DDH. The results of the 4 SNPs genotyped based on the additive model in the combined sample are shown in Table 2. The complete results of the single-marker-based analyses in the discovery stage are summarized in Supplemental Table S2. Four LD blocks were constructed based on the 32 selected markers in the discovery stage (Fig. 1 and Supplemental Figure S1). Haplotype analysis identified a 2-SNP haplotype as significantly associated with the disease status of DDH (rs6508976-rs1800470, $P = 1.65 \times 10^{-26}$), and this finding confirmed the results of the single-marker-based analyses (Supplemental Table S3).

We performed association analyses for rs1800470 stratified by different DDH types. Our results showed that the significance of rs1800470 was retained only in the severe type, i.e., dislocation of the hips. Conversely, this SNP did not show any significant association with the disease status of DDH for the mild to moderate types (acetabular dysplasia and subluxation of the hips) (Table 3). Furthermore, in case-only analyses, rs1800470 did not show a significant association with the severity of DDH (Supplemental Table S4), and no significant signals were identified in pair-wise gene-by-gene interaction analyses after Bonferroni corrections were applied (Supplemental Table S5).

Bioinformatic Analyses. rs1800470 is a common non-synonymous SNP (P [Pro] - L [Leu]) located in exon 7 of *TGFBI*. Population-based data from ExAc³⁵ and the 1000 Genomes Project²² show that the MAF of the C allele is > 0.4 in both the Chinese population and aggregated populations (similar to our MAF data from the discovery stage). The score of Polyphen2 for this SNP was 0.64, and its biological effect was considered “tolerated”. Functional prediction by SIFT (Sorting Intolerant From Tolerant) classifies this SNP as a “benign” variant. The RegulomeDB score for this SNP is 4, suggesting that this SNP has very limited functional significance (the score system for RegulomeDB ranges from 1 to 6; the higher the score, the lower the functional significance of an SNP).

The sub-network structure of *IL6* and *TGFBI* were explored using STRING, and the interacting genes are shown in Supplemental Figure S2. For *TGFBI*, two other genes from the transforming growth factor gene family, *TGFBR1* and *TGFBR2*, were identified as highly related to our candidate gene. In addition, several genes from the mothers against decapentaplegic homolog (SMAD) gene family, including *SMAD2*, *SMAD3* and *SMAD7*, were also found to interact with *TGFBI*. For *IL-6*, the most related gene identified was the gene encoding its receptor, *IL6R*.

Discussion

Our study established significant associations between the disease status of DDH and polymorphisms in *TGFBI* (rs1800470) and *IL-6* (rs1800796). Two previous studies attempted to identify relationships between *TGFBI* and *IL-6* with the severe form of DDH^{20,21}. Using data from 28 patients and 20 healthy controls, Kolundžić *et al.* identified association signals of rs1800470 in *TGFBI* and rs1800796 in *IL-6* with severe hip osteoarthritis in adulthood²⁰, and using data from 68 cases and 152 controls, Čengić *et al.* reported significant associations between

these two SNPs and severe hip osteoarthritis²¹. To some extent, the present study can be considered a replication of these previous analyses; however, our study has several advantages. First, both of these previous studies focused on the severe type of DDH, whereas we recruited DDH patients with different forms of severity. Thus, our study design enabled us to extend our results to a more generalized DDH population instead of only the severe type. In addition, our study design allowed for investigating whether the identified significant SNPs could also serve as an indicator of DDH severity. Our stratification analyses showed that a significant signal of rs1800470 could only be identified in the severe type of DDH (dislocation of the hips), and subsequent analysis focusing on the effects of this SNP and DDH severity showed negative results (rs1800470, OR = 0.99, $P = 0.9021$). This contradiction can be partly explained by differences in sample sizes for the different types of DDH. We recruited a total of 1,064 DDH cases in both stages, and among these subjects, there were 60 cases of acetabular dysplasia, 213 cases of subluxation of hips and 811 cases of dislocation of hip. In stratification analyses, the effect of an SNP might be missed in groups of the less severe types (acetabular dysplasia and subluxation of the hips) due to insufficient statistical power, and the post-hoc statistical power for subgroup of acetabular dysplasia and subluxation of hips were 38% and 55%, respectively. Second, both former studies had very limited sample sizes and suffered from low statistical power, and the ORs of the targeted SNPs were estimated to be very high (6–10). In contrast, our estimation of the OR of rs1800470 was 1.255. This difference can be explained, at least partly, by the different sample sizes used in the two previous studies compared to our study. Given that it is not sufficient to draw conclusions from analyses of only some SNPs^{36–40}, our haplotype analyses further corroborated that the SNP rs1800470 was strongly associated with the risk of DDH. Third, both previous studies only considered two SNPs, rs1800470 and rs1800796. Moreover, we selected targeted SNPs using population-based large-scale genotyping data, and the genomic coverage of our study is superior to those of the previous studies.

The significant SNP rs1800470 is a non-synonymous SNP with a high MAF. Although this SNP alters the amino acid sequence of the polypeptide encoded by *TGFBI*, the potential functional consequence remains unclear. One previous study hypothesized that this SNP might affect the peptide export efficiency of TGF- β 1¹⁹. However, this assumption has not been confirmed experimentally. Our bioinformatic analyses based on large-scale databases (SIFT, Polyphen2 and RegulomeDB) indicated that this SNP has very limited functional significance. A possible explanation for the statistical association signal identified in this study for rs1800470 is that this signal might be a surrogate for several rare or low-frequency variants that are enriched around the genomic position of this SNP. Regardless, this hypothesis cannot be directly tested in this study, which was mainly based on selected common SNPs. The other significant SNP rs1800796 which located at the 5' near gene region also had very limited functional significance according to RegulomeDB (with score of 4). Future studies based on sequencing technology and populations from different ethnic groups with larger sample sizes are needed to validate this result.

Population stratification is a challenge for our study. As a candidate gene-based study, it is unrealistic for us to implement standard procedures, such as genomic control or principle component analysis, requiring genome-wide genotyping data to adjust this potential confounder. Nonetheless, this might be a problem, especially considering that the samples for the discovery and validation stages were recruited from different sites. To adjust for this potential confounder, we restricted our sample recruitment process to ensure that the study subjects had relatively similar genetic backgrounds, even though this might bring a new problem of generalization. In addition, for combined analysis using samples from both the discovery and validation stages, we included a covariate indicating the sample batch to adjust for potential batch effects. Another potential issue is that SNP rs1800470 was not significant in the validation stage (OR = 1.18, $P = 0.04$). We performed a post-hoc power analysis using the sample size ($N = 2,781$) and the OR information (OR = 1.18) of the validation stage for rs1800470, revealing that the statistical power for detecting this SNP was only approximately 40%. Therefore, insufficient statistical power might have had a strong contribution.

The TGF- β /SMAD signalling pathway plays a very critical role in the expression of genes controlling differentiation in embryonic stem cells⁴¹. *SMAD3* encodes a transcriptional modulator that binds to the promoter region of many genes that are regulated by TGF- β . As a candidate-gene-based study, we only focused on two genes: *TGFBI* and *IL-6*. However, according to the results of protein-protein interaction analysis, several other genes appear to strongly interact with our candidate genes (Supplemental Figure S2). Among these, *SMAD3*, which encodes a transcriptional modulator that binds to the promoter region of many genes that are regulated by TGF- β , is particularly interesting. The onset and development of DDH might involve additional genes in a systematic manner. In the future, it will be important to include more *TGFBI*-related genes, perhaps an entire pathway, in a genetics study, and the TGF- β /SMAD signalling pathway is a good candidate in this regard.

Conclusion

In this study, we investigated the potential association of *TGFBI* and *IL-6* with DDH. Our findings support an important role for *TGFBI* and *IL-6* in the onset and development of DDH, though this significant SNP cannot serve as an indicator of DDH severity. Further research is needed to validate the association identified between *IL-6* and DDH.

References

1. Kokavec, M. & Bialik, V. Developmental dysplasia of the hip. *Prevention and real incidence. Bratislavské lekárske listy* **108**, 251–254 (2007).
2. Sollazzo, V., Bertolani, G., Calzolari, E., Atti, G. & Scapoli, C. A two-locus model for non-syndromic congenital dysplasia of the hip (CDH). *Annals of human genetics* **64**, 51–59, doi:10.1017/S000348000007958 (2000).
3. Laurence, M., Harper, P. S., Harris, R., Nevin, N. C. & Roberts, D. F. Report of the delegation of clinical geneticists to China, Spring 1986. *Biology and society: the journal of the Eugenics Society* **4**, 61–77 (1987).
4. Lau, E. M., Lin, F., Lam, D., Silman, A. & Croft, P. Hip osteoarthritis and dysplasia in Chinese men. *Annals of the rheumatic diseases* **54**, 965–969 (1995).

5. Weinstein, S. L. Natural history of congenital hip dislocation (CDH) and hip dysplasia. *Clinical orthopaedics and related research*, 62–76 (1987).
6. Mabuchi, A. *et al.* Osteoarthritis of the Hip Joint Associated with Acetabular Dysplasia Maps to Chromosome 13q. *Am J Hum Genet.* **79**(1), 163–168 (2006).
7. Li, L. *et al.* CX3CR1 polymorphisms associated with an increased risk of developmental dysplasia of the hip in human. *Journal of orthopaedic research: official publication of the Orthopaedic Research Society* **35**, 377–380, doi:10.1002/jor.23294 (2017).
8. Liu, S. *et al.* Two single-nucleotide polymorphisms in the DKK1 gene are associated with developmental dysplasia of the hip in the Chinese Han female population. *Genet Test Mol Biomarkers* **18**, 557–561, doi:10.1089/gtmb.2014.0044 (2014).
9. Tian, W. *et al.* Association analysis between HOXD9 genes and the development of developmental dysplasia of the hip in Chinese female Han population. *Bmc Musculoskeletal Disorders* **13**, 59 (2012).
10. Jia, J. *et al.* Association of a single nucleotide polymorphism in pregnancy-associated plasma protein-A2 with developmental dysplasia of the hip: a case-control study. *Osteoarthritis and cartilage* **20**, 60–63, doi:10.1016/j.joca.2011.10.004 (2012).
11. Shi, D. *et al.* Association of the D repeat polymorphism in the ASPN gene with developmental dysplasia of the hip: a case-control study in Han Chinese. *Arthritis research & therapy* **13**, R27, doi:10.1186/ar3252 (2011).
12. Dai, J. *et al.* Association of a single nucleotide polymorphism in growth differentiate factor 5 with congenital dysplasia of the hip: a case-control study. *Arthritis research & therapy* **10**, R126, doi:10.1186/ar2540 (2008).
13. Hao, Z. *et al.* Association of a single nucleotide polymorphism in HOXB9 with developmental dysplasia of the hip: a case-control study. *Journal of orthopaedic research: official publication of the Orthopaedic Research Society* **32**, 179–182, doi:10.1002/jor.22507 (2014).
14. Wang, K. *et al.* Association of a single nucleotide polymorphism in Tbx4 with developmental dysplasia of the hip: a case-control study. *Osteoarthritis and cartilage* **18**, 1592–1595, doi:10.1016/j.joca.2010.09.008 (2010).
15. Vaughn, S. P. *et al.* Confirmation of the mapping of the Camurati-Englemann locus to 19q13.2 and refinement to a 3.2-cM region. *Genomics* **66**, 119–121, doi:10.1006/geno.2000.6192 (2000).
16. Zilberstein, A., Ruggieri, R., Korn, J. H. & Revel, M. Structure and expression of cDNA and genes for human interferon-beta-2, a distinct species inducible by growth-stimulatory cytokines. *The EMBO journal* **5**, 2529–2537 (1986).
17. Malesud, C. J. Anticytokine therapy for osteoarthritis: evidence to date. *Drugs & aging* **27**, 95–115, doi:10.2165/11319950-000000000-00000 (2010).
18. Kwan Tat, S., Padrines, M., Theoleyre, S., Heymann, D. & Fortun, Y. IL-6, RANKL, TNF-alpha/IL-1: interrelations in bone resorption pathophysiology. *Cytokine & growth factor reviews* **15**, 49–60 (2004).
19. Dickinson, M. E. *et al.* Chromosomal localization of seven members of the murine TGF- β superfamily suggests close linkage to several morphogenetic mutant loci. *Genomics* **6**, 505 (1990).
20. Kolundzic, R. *et al.* Association of interleukin-6 and transforming growth factor-beta1 gene polymorphisms with developmental hip dysplasia and severe adult hip osteoarthritis: a preliminary study. *Cytokine* **54**, 125–128, doi:10.1016/j.cyto.2011.02.004 (2011).
21. Cengic, T. *et al.* Association of TGF β 1 29C/T and IL6 -572G/C polymorphisms with developmental hip dysplasia: a case-control study in adults with severe osteoarthritis. *International orthopaedics* **39**, 793–798, doi:10.1007/s00264-015-2675-0 (2015).
22. Genomes Project, C. *et al.* A global reference for human genetic variation. *Nature* **526**, 68–74, doi:10.1038/nature15393 (2015).
23. 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).
24. Guan, F. *et al.* MIR137 gene and target gene CACNA1C of miR-137 contribute to schizophrenia susceptibility in Han Chinese. *Schizophr Res.* **152**(1), 97–104 (2014).
25. Guan, F. *et al.* Evaluation of association of common variants in HTR1A and HTR5A with schizophrenia and executive function. *Sci Rep* **6**, 38048 (2016).
26. Guan, F. *et al.* Two-stage association study to identify the genetic susceptibility of a novel common variant of rs2075290 in ZPR1 to type 2 diabetes. *Sci Rep* **6**, 29586 (2016).
27. Purcell, S. *et al.* PLINK: a tool set for whole-genome association and population-based linkage analyses. *American journal of human genetics* **81**, 559–575 (2007).
28. Gabriel, S. B. *et al.* The structure of haplotype blocks in the human genome. *Science* **296**, 2225–2229, doi:10.1126/science.1069424 (2002).
29. Team, R. D. C. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. *Computing* **1**, 12–21 (2013).
30. Afgan, E. *et al.* The Galaxy platform for accessible, reproducible and collaborative biomedical analyses: 2016 update. *Nucleic Acids Res* **44**, W3–W10, doi:10.1093/nar/gkw343 (2016).
31. Boyle, A. P. *et al.* Annotation of functional variation in personal genomes using RegulomeDB. *Genome Research* **22**, 1790–1797 (2012).
32. Adzhubei, I. A. *et al.* A method and server for predicting damaging missense mutations. *Nature methods* **7**, 248–249, doi:10.1038/nmeth0410-248 (2010).
33. Kumar, P., Henikoff, S. & Ng, P. C. Predicting the effects of coding non-synonymous variants on protein function using the SIFT algorithm. *Nature Protocols* **4**, 1073 (2009).
34. Szklarczyk, D. *et al.* STRING v10: protein-protein interaction networks, integrated over the tree of life. *Nucleic Acids Res* **43**, D447–452, doi:10.1093/nar/gku1003 (2015).
35. Lek, M. *et al.* Analysis of protein-coding genetic variation in 60,706 humans. *Nature* **536**, 285–291, doi:10.1038/nature19057 (2016).
36. Chen, G. *et al.* Genetic analysis of common variants in the HDAC2 gene with schizophrenia susceptibility in Han Chinese. *J Hum Genet.* **60**, 479–484 (2015).
37. Guan, F. *et al.* Two-stage replication of previous genome-wide association studies of AS3MT-CNNM2-NT5C2 gene cluster region in a large schizophrenia case-control sample from Han Chinese population. *Schizophr Res.* **176**(2–3), 125–130 (2016).
38. Jia, X. *et al.* Two-stage additional evidence support association of common variants in the HDAC3 with the increasing risk of schizophrenia susceptibility. *Am J Med Genet B Neuropsychiatr Genet.* **171**, 1105–1111 (2016).
39. Yang, H. *et al.* 4q22.1 contributes to bone mineral density and osteoporosis susceptibility in postmenopausal women of Chinese Han population. *PLoS One.* **8**, e80165 (2013).
40. Zhang, B. *et al.* Common variants in SLC1A2 and schizophrenia: Association and cognitive function in patients with schizophrenia and healthy individuals. *Schizophr Res.* **169**, 128–34 (2015).
41. Massague, J. & Xi, Q. TGF-beta control of stem cell differentiation genes. *FEBS Lett* **586**, 1953–1958, doi:10.1016/j.febslet.2012.03.023 (2012).

Acknowledgements

This research was supported by the Project of Shaanxi Key Scientific and Technological Innovation Team (No:2013KCT-26). The funding sources had no role in the design of this study, the collection, analysis and interpretation of data, the writing of the report, or the decision to submit the paper for publication.

Author Contributions

Authors Zhi L. and Yao S. conceived and designed the study. Ma W., Chen K. and Chen H. carried out candidate SNPs selection and statistical analyses. Wu Y., Ma J. and Zeng S. conducted subject screening and contributed to the collection and preparation of control DNA samples. Ma W. and Zha Z. wrote the first draft of the paper, and Yao S. revised the current paper.

Additional Information

Supplementary information accompanies this paper at doi:[10.1038/s41598-017-11185-1](https://doi.org/10.1038/s41598-017-11185-1)

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