JBMR Plus, 2024, **8**, ziae028
https://doi.org/10.1093/jbmrpl/ziae028
Advance access publication: March 5, 2024
Research Article

Identification of an novel genetic variant associated with osteoporosis: insights from the Taiwan Biobank Study

Yi-Ching Liaw[1,2], Koichi Matsuda[1,3,*], Yung-Po Liaw[2,4,5,*]

[1]Department of Computational Biology and Medical Sciences, Laboratory of Clinical Genome Sequencing, Graduate School of Frontier Sciences, The University of Tokyo, Tokyo 108-8639, Japan
[2]Department of Public Health and Institute of Public Health, Chung Shan Medical University, Taichung 40201, Taiwan
[3]Institute of Medical Science, The University of Tokyo, Laboratory of Genome Technology, Human Genome Center, Tokyo 108-8639, Japan
[4]Institute of Medicine, Chung Shan Medical University, Taichung 40201, Taiwan
[5]Department of Medical Imaging, Chung Shan Medical University Hospital, Taichung 40201, Taiwan
*Corresponding authors: Koichi Matsuda, Department of Computational Biology and Medical Sciences, Graduate School of Frontier Sciences, The University of Tokyo, No. 4-6-1, Shirokanedai, Minato, Tokyo 108-8639, Japan (kmatsuda@edu.k.u-tokyo.ac.jp); Yung-Po Liaw, Department of Public Health, Institute of Public Health, Chung Shan Medical University, No. 110, Sec. 1 Jianguo N. Rd., Taichung City 40201, Taiwan (Liawyp@csmu.edu.tw).

Abstract

Purpose: The purpose of this study was to identify new independent significant SNPs associated with osteoporosis using data from the Taiwan Biobank (TWBB).
Material and Methods: The dataset was divided into discovery (60%) and replication (40%) subsets. Following data quality control, genome-wide association study (GWAS) analysis was performed, adjusting for sex, age, and the top 5 principal components, employing the Scalable and Accurate Implementation of the Generalized mixed model approach. This was followed by a meta-analysis of TWBB1 and TWBB2. The Functional Mapping and Annotation (FUMA) platform was used to identify osteoporosis-associated loci. Manhattan and quantile–quantile plots were generated using the FUMA platform to visualize the results. Independent significant SNPs were selected based on genome-wide significance ($P < 5 \times 10^{-8}$) and independence from each other ($r^2 < 0.6$) within a 1 Mb window. Positional, eQTL(expression quantitative trait locus), and Chromatin interaction mapping were used to map SNPs to genes.
Results: A total of 29 084 individuals (3154 osteoporosis cases and 25 930 controls) were used for GWAS analysis (TWBB1 data), and 18 918 individuals (1917 cases and 17 001 controls) were utilized for replication studies (TWBB2 data). We identified a new independent significant SNP for osteoporosis in TWBB1, with the lead SNP rs76140829 (minor allele frequency = 0.055, P-value = 1.15×10^{-08}). Replication of the association was performed in TWBB2, yielding a P-value of 6.56×10^{-3}. The meta-analysis of TWBB1 and TWBB2 data demonstrated a highly significant association for SNP rs76140829 (P-value = 7.52×10^{-10}). In the positional mapping of rs76140829, 6 genes (*HABP2*, *RP11-481H12.1*, *RNU7-165P*, *RP11-139 K1.2*, *RP11-57H14.3*, *and RP11-214 N15.5*) were identified through chromatin interaction mapping in mesenchymal stem cells.
Conclusions: Our GWAS analysis using the Taiwan Biobank dataset unveils rs76140829 in the *VTI1A* gene as a key risk variant associated with osteoporosis. This finding expands our understanding of the genetic basis of osteoporosis and highlights the potential regulatory role of this SNP in mesenchymal stem cells.

Keywords: osteopetrosis < diseases and disorders of/related to bone, genetic research, general population studies < epidemiology.

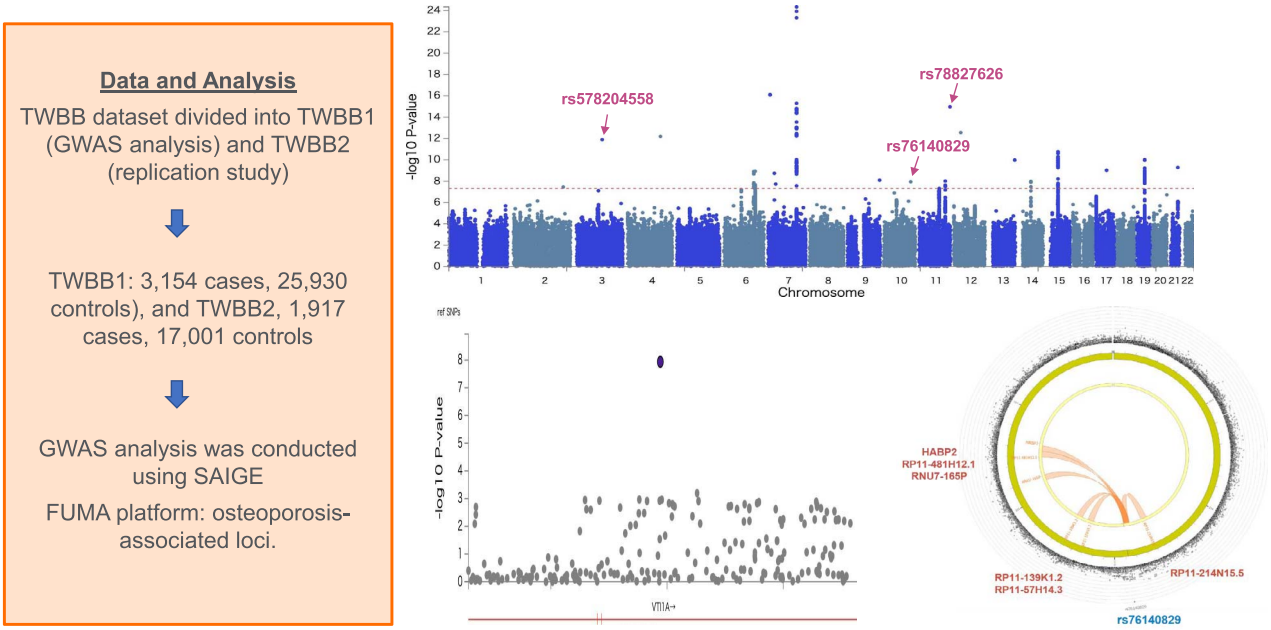
Lay Summary

Our study aimed to identify new genetic factors associated with osteoporosis, a condition characterized by weak and brittle bones. We analyzed data from the Taiwan Biobank, which included information from thousands of individuals. We divided the data into 2 groups: one for discovering potential genetic factors and another for confirming our findings. Using advanced genetic analysis techniques, we identified a specific genetic variant, called rs76140829, that is linked to osteoporosis. This variant is located within a gene called *VTI1A* on chromosome 10. We found that individuals with this genetic variant were more likely to have osteoporosis compared to those without it. Further analysis confirmed our findings, showing a strong association between rs76140829 and osteoporosis across both groups of data. We also investigated how this genetic variant may affect the activity of nearby genes in bone-forming cells called mesenchymal stem cells. Our study sheds light on the genetic factors contributing to osteoporosis and highlights the potential importance of the *VTI1A* gene in this condition. Understanding these genetic factors could lead to personalized treatments and new therapies for osteoporosis in the future.

Received: January 3, 2024. **Revised:** February 18, 2024. **Accepted:** March 1, 2024

Graphical Abstract

Identifying genetic variants associated with osteoporosis using Taiwan Biobank (TWBB) data



Introduction

Osteoporosis, a common skeletal disorder characterized by reduced BMD and increased fracture risk, presents a significant global health burden, particularly in older men and women.[1] Approximately 30% of all osteoporotic fractures occur in men, highlighting that this disease is not exclusive to women.[2]

The development of osteoporosis depends on a combination of genetic, environmental, and lifestyle factors, with genetic factors playing a key role in determining an individual's susceptibility.[3] Clinically, osteoporosis is diagnosed by measuring BMD, which is highly heritable.[4-8] BMD variations have an estimated heritability ranging from 50% to 82%.[8,9]

Substantial progress has been made to identify genetic variants and phenotypes associated with osteoporosis through genome-wide association studies (GWAS).[10] Efforts have been made to map associated variants to osteoporosis-causing genes.[11] GWAS are hypothesis-free and search for associations across all genotyped regions. Thousands of genetic variants have been identified through GWAS.[12,13] Several genetic variants associated with BMD have been identified through previous GWAS and meta-analyses.[14-21] However, most of these studies have focused on common rather than rare genetic variants. Currently, over 500 loci have been associated with bone traits.[22] Despite significant advancements in identifying genes and loci influencing BMD and fractures, many genetic variations contributing to these phenotypes remain unknown.[23] Most GWAS on bone health have predominantly focused on participants of European ethnicities, highlighting the need for such studies in other ethnic groups.[24]

Results from GWAS typically do not directly translate into causal variants because the majorities of hits are in non-coding or intergenic regions. Chromatin interaction mapping is used to map SNPs to genes when there is a significant chromatin interaction between the disease-associated regions and nearby or distant genes. Recent studies of chromatin modification landscapes in a wide range of tissues and cell types have contributed significantly to our understanding of genome function and regulation.[25,26] A study by Schmitt et al.[27] demonstrated that by generating a rich resource of chromatin contact maps across 21 human tissues/ cell types and exploring with integrative analytic methods, the authors were able to catalog 3D genome interactions at various hierarchical levels, and uncovered the highly dynamic nature of local interaction hotspots. These results provide insights into the chromatin organization in mammalian cells.

Moreover, it is critical to investigate the presence of rare or low-frequency genetic variants with substantial effect sizes as this could offer new insights into the pathophysiology of osteoporosis and potentially contribute to the development of targeted therapeutic approaches. In this context, the purpose of this study was to explore the genetic landscape of osteoporosis using data from the Taiwan Biobank (TWBB). We aimed to identify and characterize novel genetic variants associated with osteoporosis in the Taiwanese population. Our study capitalizes on the large sample size and rich genomic diversity within the TWBB to improve the current understanding of osteoporosis genetics and its implications for disease management and prevention.

Materials and methods

GWAS data source and genotyping

GWAS data from TWBB

The TWBB is a prospective cohort that includes individual genotype and detailed clinical data. The DNA samples from TWBB participants were stored at −80°C. Genetic data were obtained using the TWBv2 array (Thermo Fisher Scientific, Inc.). The whole-genome genotyping data were obtained

from the TWBB using the TWBv2.0 genotype identification chip, which genotyped approximately 750 000 SNPs. A total of 16 208 573 SNPs in TWBB had been identified through imputation. This imputation process utilized whole-genome sequencing information obtained from blood samples collected from individuals in the Taiwanese population. Impute 2 software was employed, using the 1KG phase3 EAS (East Asian) dataset as the reference panel. The EAS dataset includes 504 individuals, while the TWBB comprises 1451 individuals. In terms of haplotypes, there were 1008 in the EAS dataset and 2902 in the TWBB dataset, resulting in a total of 3910 haplotypes.

Bone mineral density measurement

BMD was measured at the heel calcaneus in grams per square centimeter (g/cm^2) using the ultrasound (Achilles InSight, GE). We first used 60% of the data as discovery data, and 40% of the data as replication data. Subjects with T-scores between −2.5 and −1 were excluded. Osteoporosis was defined as T-scores less than -2.5, while T-scores above -1 were classified as the control group. It is important to note that T-scores at or below −2.5 are classified as indicative of osteoporosis. This means that the individual has significantly lower bone density, which puts them at a higher risk of fractures and other complications associated with weakened bones. Moreover, T-scores above −1.0 are generally considered normal or indicative of healthy bone density. T-scores between −2.5 and −1 are not used to diagnose osteoporosis because they represent a category known as osteopenia. Although osteopenia indicates lower bone density and an increased risk of fractures compared to normal bone density, it is not classified as osteoporosis, and the management approach for these 2 conditions is distinct. Osteoporosis diagnosis and treatment are generally reserved for individuals with T-scores of −2.5 or lower.

Data quality control

Prior to conducting principal component analysis (PCA), samples were excluded based on the following criteria: (1) sample call rate below 0.98, (2) heterozygosity rate outside 5 SDs from the sample average, and (3) exclusion of one individual from pairs of related samples (duplicates or third-degree relatives) based on pairwise identity-by-decent. PCA was performed using PLINK.[28] The first 5 principal components (PCs) from the PCA were included as covariates in the model to control for population stratification.

Initially, we divided the data into 2 sets: 60% for the discovery phase and 40% for replication purposes. Subjects with T-scores between −2.5 and −1 were excluded. Following quality control measures, we referred to the discovery data as TWBB1 and the replication data as TWBB2. Data from TWBB1 were used for GWAS analysis, while those from TWBB2 were used for the replication study. For the TWBB1 GWAS analysis, logistic regression models under an additive genetic model were employed, adjusting for sex, age, and the top 5 PCs using the Scalable and Accurate Implementation of Generalized mixed model approach.[29] Subsequently, the resulting summary statistics underwent further scrutiny and exploration on Functional Mapping and Annotation (FUMA) platform for comprehensive post-GWAS analysis.

Functional annotation using the FUMA platform

As GWAS results do not necessarily reflect causal variants, this study leveraged the FUMA platform[30] to enhance the functional annotation of GWAS outcomes sourced from multiple datasets. This platform was used to identify genetic loci associated with osteoporosis based on GWAS summary statistics. Functional annotations of SNPs were sourced from RegulomeDB,[31] Combined Annotation Dependent Depletion (CADD),[32] and the core 15-state model of chromatin.[26,33,34] The functional consequences of SNPs on genes were determined through ANNOVAR (annotate variation)[35] using Ensembl genes (build 85).

Additionally, the FUMA platform was employed to visualize the association results through Manhattan and quantile–quantile (Q–Q) plots.[36] Based on the GWAS summary statistics data for osteoporosis in Taiwan, we identified independent significant SNPs, based on their genome-wide significance ($P < 5 \times 10^{-8}$) and their independence from each other ($r^2 < 0.6$) within a 1 Mb window.[37,38]

The main results of a GWAS are generally depicted through a Manhattan plot, a scatterplot illustrating the negative logarithm (base 10) of the P = value (y-axis) against the SNP's association significance, ordered by its chromosomal position (x-axis). Another customary visualization in GWAS is the Q–Q plot, analyzed alongside lambda (λ) values.[36] Q–Q plots compare the distribution of observed P-values (logarithm scale) with the expected P-value distribution under the null hypothesis (ie, no association between genotypes and phenotype), aiding in the detection of genotype–phenotype associations and population substructure control. The degree of deviation from the y = x line is formally measured by the lambda statistic, also called genomic control. A lambda value can be calculated from Z-scores, chi-square statistics, or P-values. A lambda value is also used to detect possible inflation due to population stratification. A value close to 1 suggests that data have been properly adjusted for the population structure.

Positional, eQTL, and Chromatin interaction mapping are used to map SNPs to genes. Cis-eQTL information was gathered from 4 distinct data repositories: GTEx portal v6,[39] Blood eQTL browser,[40] BIOS QTL Browser,[41] and BRAINEAC[42](r10). The mapping of genes was done using Ensembl gene IDs. Utilizing chromatin interaction mapping, we employed a technique to link single nucleotide polymorphisms (SNPs) located within genomic risk loci to specific genes. This was achieved by detecting meaningful interactions within the chromatin structure between regions containing SNPs and regions containing genes. This mapping process relied on capturing 3-dimensional DNA–DNA interactions without imposing a specific distance limit. The FUMA platform currently incorporates Hi-C data from 14 human tissue types and 7 human cell lines, as detailed in the study by Schmitt et al.[27] The 14 human tissues encompass the ovary, bladder, psoas muscle, left ventricle, aorta, right ventricle, lung, spleen, small bowel, adrenal gland, pancreas, prefrontal cortex, liver, and hippocampus. The 7 human cell lines consist of embryonic stem cells (ESCs), mesendoderm cells, mesenchymal stem cells, neural progenitor cells, trophoblast-like cells, fibroblast cells, and lymphoblast cells. Given that chromatin interactions are often defined at a specific resolution (such as 40 kilobases), it is possible for an interaction region to encompass multiple genes. Consequently, all SNPs within these regions were mapped to genes located within the corresponding interaction region. To enhance the prioritization of candidate genes, we integrated data on predicted enhancers and promoters specific to various tissues and cell types from the Roadmap

Table 1. Nineteen independent significant SNPs identified in the Taiwan osteoporosis TWBB1 GWAS.

Ind. Sig. SNP	Chr	Region	Non- effect allele	Effect allele	BETA	SEBETA	*P*-value	Nearest Gene	Gene Distance	ANNOVAR category
rs578204558	**3**	**3q13.13**	**T**	**A**	**51.362**	**7.241**	$\mathbf{1.31 \times 10^{-12}}$	***RP11-457 K10. 2***	**8988**	**Intergenic**
rs9482773	**6**	**6q22.33**	**C**	**G**	**0.177**	**0.029**	$\mathbf{1.24 \times 10^{-9}}$	***RSPO3***	**0**	**Intronic**
rs949767	**6**	**6q23.2**	**C**	**G**	**−0.246**	**0.040**	$\mathbf{1.19 \times 10^{-9}}$	***EYA4***	**0**	**Intronic**
rs211623	6	6q23.2	T	G	−0.233	0.043	4.04×10^{-8}	*EYA4*	0	Intronic
rs10250475	7	7q31.31	C	T	−0.204	0.032	9.71×10^{-11}	*CPED1*	0	Intronic
rs10231005	7	7q31.31	C	A	−0.225	0.030	2.94×10^{-14}	*WNT16*	12 838	Intergenic
rs2908007	**7**	**7q31.31**	**G**	**A**	**0.301**	**0.029**	$\mathbf{4.74 \times 10^{-25}}$	***WNT16***	**3256**	**Intergenic**
rs2536185	7	7q31.31	G	T	−0.307	0.038	5.09×10^{-16}	*WNT16*	2882	Intergenic
rs76140829	**10**	**10q25.2**	**G**	**A**	**−0.375**	**0.066**	$\mathbf{1.15 \times 10^{-8}}$	***VTI1A***	**0**	**Intronic**
rs491895	**11**	**11q14.2**	**C**	**G**	**−0.171**	**0.031**	$\mathbf{4.86 \times 10^{-8}}$	***TMEM135***	**0**	**Intronic**
rs6589302	**11**	**11q23.1**	**T**	**C**	**−0.172**	**0.030**	$\mathbf{9.82 \times 10^{-9}}$	***RP11-65 M17.3 : RP11-65 M17. 1***	**0:0**	**ncRNA_intro nic**
rs78827626	**11**	**11q25**	**G**	**A**	**9.903**	**1.236**	$\mathbf{1.10 \times 10^{-15}}$	***OPCML***	**0**	**Intronic**
rs2761884	**14**	**14q22.2**	**G**	**T**	**0.167**	**0.029**	$\mathbf{1.04 \times 10^{-8}}$	***BMP4***	**0**	**Intronic**
rs10519302	**15**	**15q21.2**	**A**	**G**	**−0.215**	**0.032**	$\mathbf{1.73 \times 10^{-11}}$	***CYP19A1***	**0**	**Intronic**
rs78339028	15	15q21.2	G	A	0.193	0.033	6.07×10^{-9}	*CYP19A1*	0	Intronic
rs1004982	15	15q21.2	T	C	−0.176	0.032	4.19×10^{-8}	*CYP19A1:RP11 -108 K3.2*	0:0	ncRNA_exoni c
rs28840750	19	19q13.11	T	G	−0.229	0.038	1.30×10^{-9}	*RHPN2*	0	Intronic
rs73039433	**19**	**19q13.11**	**G**	**A**	**−0.226**	**0.035**	$\mathbf{9.79 \times 10^{-11}}$	***RHPN2***	**0**	**Intronic**
rs7256899	19	19q13.11	C	G	−0.249	0.042	4.04×10^{-9}	*CTC-461H2.3*	122	Downstream

Epigenomics Project.[25] Using this information, significant interactions were defined based on a false discovery rate of 1×10^{-6}, as recommended by Schmitt et al.[21]

Results

A total of 29 084 subjects (TWBB1, 3154 osteoporosis cases and 25 930 controls) were used for GWAS analysis while 18 918 subjects (TWBB2, 1917 cases and 17 001 controls) were utilized for the replication study.

Identification of significant SNPs

Based on the TWBB1 GWAS summary statistics and the FUMA platform, we identified 19 independent significant SNPs, most of which were intronic (Table 1). Among these SNPs, 11 were lead SNPs (that is, those that had $r^2 < 0.1$ from each other). The Manhattan plot used to visualize the GWAS results is shown in Figure 1, while the Q–Q plot is shown in Figure 2. The lambda value was 1.006.

Identification of key SNPs

In this investigation, we successfully pinpointed 3 previously unidentified SNPs denoted as rs578204558, rs76140829, and rs78827626. These SNPs are situated on autosomal chromosomes 3, 10, and 11, respectively, within the genes *RP11-457 K10.2, VTI1A*, and *OPCML* (refer to Table 1 and Figure 1 for a visual representation of their genomic locations). Notably, the associated *P*-values for these SNPs are 1.31×10^{-12}, 1.15×10^{-8}, and 1.10×10^{-15}. We evaluated the imputation quality of the identified SNPs in our analysis. The info score was used to assess imputation quality, and variants with an info score above 0.4 were considered. Among the 3 SNPs identified in the TWBB1 GWAS, rs78827626 had an info score of only 0.359, indicating lower confidence in its imputation. Additionally, rs578204558 had an info score of 0.423. In contrast, the SNP rs76140829 was derived from genotype data, rather than the imputed data. Therefore, its imputation score is designated as 1.

Meta-analysis and confirmation

Table 2 presents the outcomes of the meta-analysis, incorporating data from both TWBB1 and TWBB2, focusing on SNPs rs76140829 and rs578204558. Notably, the *P*-value associated with SNP rs578204558 in the TWBB2 dataset was 0.883. Furthermore, the combined meta-analysis result for TWBB1 and TWBB2 yielded a *P*-value of .361 for rs578204558. Consequently, based on these findings, we did not categorize SNP rs578204558 as a new independent significant variant. Turning attention to the remaining candidate, SNP rs76140829, it was replicated in TWBB2 with a *P*-value of 6.56×10^{-3}. In the meta-analysis encompassing both TWBB1 (GWAS data) and TWBB2 (replication data), this SNP demonstrated a *P*-value of 7.52×10^{-10}, establishing its candidacy as a new independent significant variant.

For annotation of the newly found SNP, the CADD was 18.28, the RegulomeDB Categorical Scores was 4 (TF binding + DNase peak), and the core 15-state model of chromatin in E006 (ESC Derived) H1-Derived Mesenchymal Stem Cells was 5 (Tx/Wx, ie, weak transcription).

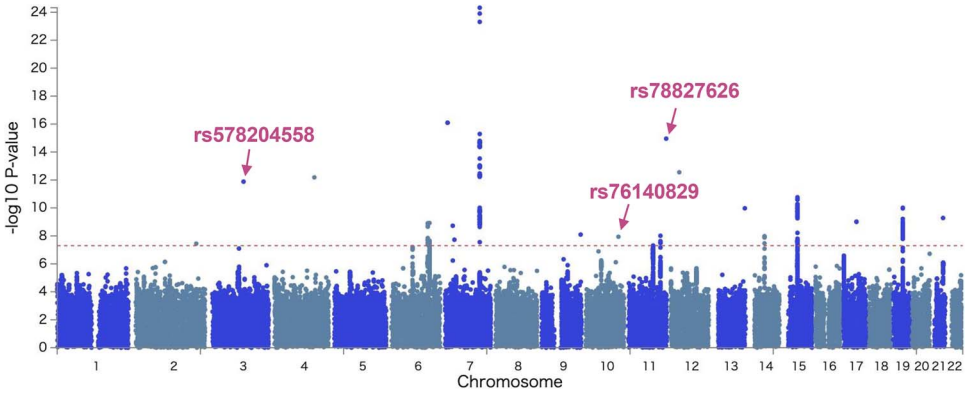


Figure 1. Manhattan plot (GWAS summary statistics) for GWAS of osteoporosis based on TWBB1 data.

Table 2. The meta-analysis (TWBB1 and TWBB2) results for SNP rs76140829 and rs578204558.

SNP	TWBB1 (TWBB GWAS)			TWBB2 (TWBB replication)			Meta-analysis[a]	
	Beta	*P*-value	*N*	Beta	*P*-value	*N*	Effect	*P*-value
rs76140829	−0.375	1.15×10^{-8}	29 084	−0.223	6.56×10^{-3}	18 918	−0.316	7.52×10^{-10}
rs578204558	51.362	1.31×10^{-12}	29 084	−0.161	.883	18 918	0.989	.361

[a]SNP rs76140829 in the meta-analysis (TWBB1 and TWBB2)

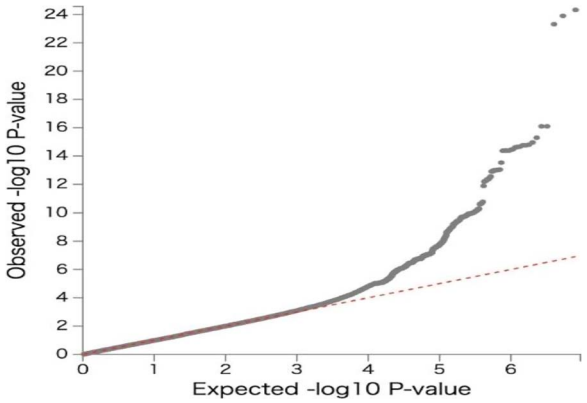


Figure 2. The QQ plot (GWAS summary statistics) derived based on TWBB1 data.

Regional association plot and gene mapping

Through positional mapping, the VTI1A gene and RP11-25C19.3 were mapped. Figure 3 displays the regional association plot of the newly found SNP, rs76140829 within the VTI1A gene in TWBB1. The associated gene, *VTI1A,* is visually represented in red. Situated intronically within the VTI1A gene, rs76140829 is positioned in the 10q25.2 region of chromosome 10, as detailed in Table 1. However, no gene was mapped through eQTL mapping. Moving to Figure 4, this visualization illustrates the genomic risk loci and chromatin interactions associated with SNP rs76140829. Noteworthy, this significant SNP for osteoporosis was found to interact with 6 genes (*HABP2*, *RP11-481H12.1*, *RNU7-165P*, *RP11-139 K1.2*, *RP11-57H14.3*, *and RP11-214 N15.5*) through chromatin interaction mapping in mesenchymal stem cells.

Discussion

In this study, we conducted a GWAS on osteoporosis using the TWBB dataset. Through this analysis, we identified a new independent significant SNP for osteoporosis in TWBB1, with the lead SNP rs76140829 (minor allele frequency [MAF] = 0.055, *P*-value = 1.15×10^{-08}). Nonetheless, this newly found SNP is restricted to TWBB. Its association was further validated in our replication study using TWBB2 data, strengthening its significance in Taiwan. Notably, we opted to bifurcate the TWBB for distinct purposes—one part for the GWAS and another for a replication study. This was because we did not detect the SNP rs76140829 in our osteoporosis GWAS across other Asian countries. Consequently, conducting a meta-analysis and comparing populations was not feasible.

Interestingly, the allele frequency of rs76140829 showed significant variation across different populations. In the 1000 Genomes Project, the MAF of rs76140829 was found to be 0.058 in Asia, 0.055 in TWBB1, and 0.054 in TWBB2. Notably, the MAF of rs76140829 was found to be particularly low among Europeans (0.00008)[43] as shown in Table 3, although no previously published data on this specific variant were available. This highlights the importance of considering population-specific genetic variations when studying the genetic basis of complex traits such as osteoporosis. Noteworthy, we reviewed existing literature to see if this newly found SNP, rs76140829 had been reported in association with osteoporosis but could not find it. To strengthen the evidence for the association, we replicated and validated the initial findings in TWBB2 cohorts (Table 2). It is also important to note that the *VTI1A* gene was previously reported by JA Morris and colleagues.[11] Consequently, we extracted the SNPs documented in their research and compared them with our newly found SNP (rs76140829) using the TWBB (in Table 4). The resulting r^2 values were found to be very small, indicating that the specific SNP is independent. Those SNPs previously reported by JA Morris were not less than 5E-8 in our results in both TWBB1 and TWBB2 (in Table 4). And their r^2 values were very low (less than 0.1) as shown in Table 4 when compared with our newly found SNP (rs76140829). For example, rs4918760 in JA Morris 2019's report had a *P*-value of 2.1E-10, whereas in TWBB1, it was 0.442, and in TWBB2, it was 0.001. And its r^2 value was low (0.021) when compared

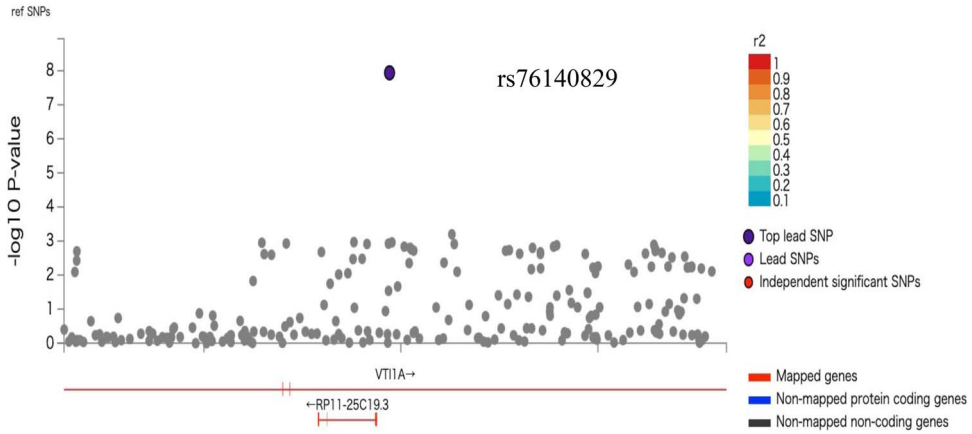


Figure 3. Regional association plot for the newly identified single-nucleotide polymorphisms (SNPs) rs76140829 in TWBB1.

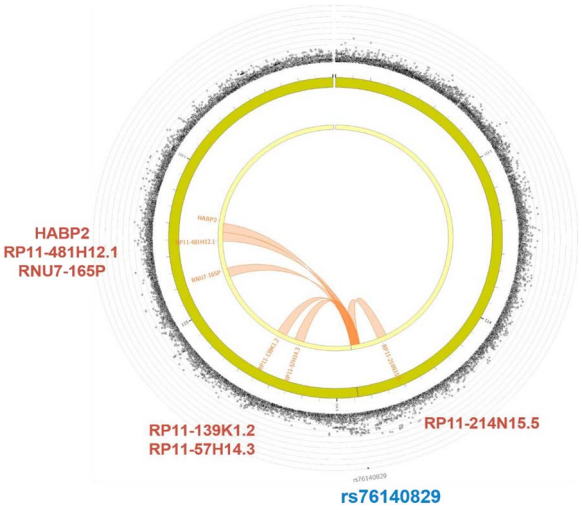


Figure 4. The chromatin interaction mappings for the novel SNP rs76140829 in TWBB1.

with our newly found SNP (rs76140829). Therefore, they are not considered candidate SNPs in the FUMA platform. Our newly found SNP (rs76140829) is currently restricted to Taiwan. It is an independent significant SNP compared to those within the *VT11A* gene published earlier. Their resulting r^2 values are relatively small in the TWBB panel, indicating that the specific SNP is independent. Additionally, we found from the National Center for Biotechnology Information database that the MAF of the rs76140829 SNP was very low among Europeans.

For annotation of the newly found SNP (rs76140829), the CADD was 18.28, a score of 18.28 is relatively high, indicating that this SNP might have a significant functional impact on the genome. CADD scores are generally interpreted in the context of other genomic data, but typically, a score above 10 suggests that the SNP is likely to be among the 10% most deleterious substitutions in the human genome. The RegulomeDB Categorical Scores were 4 (TF binding + DNase peak), a score of 4, indicating "TF binding + DNase peak," suggests that the SNP is located in a region of the genome where transcription factors (TFs) bind, as evidenced by the presence of DNase hypersensitivity sites. This implies that the SNP may influence gene expression by altering transcription factor binding sites. And the core 15-state model of chromatin in E006 (ESC Derived) H1-Derived Mesenchymal Stem Cells was 5 (Tx/Wx, ie, weak transcription). State 5 is characterized by weak transcription, indicating that the region where the SNP is located is likely to be transcriptionally active, though not strongly so, in these specific stem cells. This information can be crucial for understanding the SNP's role in specific cell types, like mesenchymal stem cells derived from ESCs. In summary, these annotations suggest that the newly found SNP is located in a genomic region that is likely to have a regulatory function, potentially affecting gene expression patterns. Its high CADD score indicates a potential for significant impact on genome function, and its location in a transcription factor binding site further supports its potential role in gene regulation. Further experimental validation and study in relevant biological systems would be necessary to fully understand its functional implications.

For Regulatory Motifs Altered by Regulatory Variant rs76140829, allele-specific binding affinity changes based on motifs of important transcriptional regulators. We utilized the Vannoportal platform[44] and discovered that rs76140829 affects the affinity of the following transcription factors, which may be related to osteoporosis: SMC3 (Structural Maintenance of Chromosomes 3), which is involved in maintaining chromosome structure may indirectly affect the expression of genes related to osteoporosis. *TFAP2A* (*Transcription Factor AP-2 Alpha*) is involved in various cellular processes including development and differentiation, which may impact bone development. *TAF1 (TATA-Box Binding Protein Associated Factor 1*) is a fundamental transcription activator, its function might affect the expression of bone-related genes. *DNMT1 (DNA Methyltransferase 1)* is involved in DNA methylation, which could affect the expression of genes related to osteoporosis. *GABPA (GA Binding Protein Transcription Factor Alpha)* might impact bone health by regulating various cellular processes. *SRF (Serum Response Factor)* involved in various cell growth and differentiation processes might impact bone health. *HOXA1 (Homeobox A1)* is a family gene involved in skeletal development and could be related to osteoporosis. *EGR1 (Early Growth Response 1)* involved in cell proliferation and differentiation, may be related to bone health. *CTCF (CCCTC-Binding Factor)* is a chromatin structure regulator that might indirectly affect osteoporosis. *ESRRA (Estrogen-Related Receptor Alpha)* is related to estrogen receptors and

Table 3. Comparing the 19 independent significant SNPs identified in TWBB with those in common databases.

Chr	19 Ind. Sig. SNP	Taiwan Biobank MAF in TWBB1 ($N = 29\,084$)	Taiwan Biobank MAF in TWBB2 ($N = 18\,918$)	Asians			Europeans		
				N	Ref Allele	Alt Allele	*N*	Ref Allele	Alt Allele
3	**rs578204558**	1.72×10^{-5}	1.85×10^{-4}	112	T = 1.000	A = 0.000	9690	T = 1.000	A = 0.000
6	rs9482773	0.447	0.449	112	G = 0.438	C = 0.562	14 286	G = 0.526	C = 0.474
6	rs949767	0.151	0.151	80	C = 1.000	A = 0.000, G = 0.000, T = 0.000	10 062	C = 0.762	A = 0.000, G = 0.238, T = 0.000
6	rs211623	0.136	0.135	112	T = 0.857	G = 0.143	14 286	T = 0.456	G = 0.544
7	rs10250475	0.312	0.316	46	C = 0.98	A = 0.00, T = 0.02	10 120	C = 0.757	A = 0.000, T = 0.243
7	rs10231005	0.415	0.418	426	C = 0.620	A = 0.380	164 612	C = 0.632	A = 0.368
7	rs2908007	0.438	0.436	3844	A = 0.480	G = 0.520	223 388	A = 0.596	G = 0.404
7	rs2536185	0.182	0.183	112	G = 0.821	T = 0.179	14 286	G = 0.526	T = 0.474
10	**rs76140829**	0.055	0.054	156	G = 0.942	A = 0.058	25 908	G = 0.99992	A = 0.00008
11	rs491895	0.325	0.320	112	G = 0.402	C = 0.598, T = 0.000	13 778	G = 0.550	C = 0.450, T = 0.000
11	rs6589302	0.440	0.433	356	C = 0.452	A = 0.000, T = 0.548	26 016	C = 0.433	A = 0.000, T = 0.567
11	**rs78827626**	5.18×10^{-5}	1.06×10^{-4}	112	G = 0.982	A = 0.018	9824	G = 1.000	A = 0.000
14	rs2761884	0.464	0.461	212	G = 0.505	T = 0.495	31 942	G = 0.577	T = 0.423
15	rs10519302	0.292	0.292	3868	A = 0.6861	G = 0.3139	223 912	A = 0.950	G = 0.050
15	rs78339028	0.266	0.267	116	G = 0.802	A = 0.198	14 286	G = 0.99650	A = 0.00350
15	rs1004982	0.296	0.296	80	T = 0.95	C = 0.05, G = 0.00	13 750	T = 0.727	C = 0.273, G = 0.000
19	rs28840750	0.187	0.184	156	T = 0.859	G = 0.141	21 540	T = 0.961	G = 0.039
19	rs73039433	0.223	0.219	112	G = 0.741	A = 0.259	14 286	G = 0.940	A = 0.060
19	rs7256899	0.139	0.141	112	C = 0.786	G = 0.214	14 286	C = 0.893	G = 0.107

might affect bone density. *RORC (Retinoic Acid Receptor-Related Orphan Receptor C)* involved in immune responses and cell differentiation, may indirectly impact bone health. *ATF1 (Activating Transcription Factor 1)* involved in various cellular processes might affect osteoporosis. It is important to note that these transcription factors have broad functions, and their roles in osteoporosis could be very complex. Additionally, they may affect osteoporosis through various mechanisms, including directly regulating the expression of bone-related genes or indirectly impacting bone health by influencing other biological processes. Further research is necessary to determine the exact relationships between these transcription factors and osteoporosis.

We also utilized the Vannoportal platform[44] and discovered that rs76140829 is related to the binding of the ATF3 transcription factor in bone marrow tissue. This SNP might alter the binding ability of ATF3 to its target DNA, thereby affecting its capacity to regulate genes involved in bone metabolism. ATF3 could affect the expression of genes related to bone resorption or bone formation, such as impacting osteoblasts or osteoclasts. Due to these changes in gene expression, there might be an imbalance in the bone formation and resorption processes, leading to osteoporosis. Further research, particularly focusing on how exactly rs76140829 affects the function of ATF3 and how this translates into biological mechanisms affecting bone health, will be crucial. This may include gene expression analysis, protein interaction studies, and functional studies in model organisms. Such research is worth further investigation.

Additionally, we determined the Probabilistic Identification of Causal SNPs (PICS) of this SNP to be 1 using PICS2.[45] Through positional mapping, the *VTI1A* gene and *RP11-25C19.3* were mapped. The *VTI1A* gene is protein coding, PLI is 0.413, and ncRVIS is −1.595. The *VTI1A* gene, known for encoding Vesicle Transport through Interaction with T-SNAREs 1A, is primarily involved in the vesicular transport system within cells. The vesicular transport system is vital for various cellular functions, including those in bone cells (osteoblasts and osteoclasts). Any disruption in this system could potentially affect bone remodeling processes, which are critical in maintaining bone density and preventing osteoporosis. *VTI1A's* role in vesicular transport might influence intracellular signaling pathways involved in bone remodeling. Efficient transport of signaling molecules is essential for the regulation of osteoblast and osteoclast activity, which in turn affects bone formation and resorption. The *RP11-25C19.3* gene is classified as an antisense gene, meaning it has a DNA sequence that is complementary to, and potentially regulates, another gene. Antisense genes typically function by binding to their corresponding sense RNA, affecting its stability and translation, thereby playing a significant role in gene regulation. In the context of osteoporosis, a condition characterized by brittle bones and a higher risk of fractures, this gene could potentially influence the expression of genes related to bone metabolism or bone cell function. Osteoporosis is associated with various genetic and environmental factors, so understanding the potential role of antisense genes like *RP11-25C19.3* in bone health could help reveal genetic mechanisms underlying osteoporosis.

This significant SNP for osteoporosis was found to interact with 6 genes (*HABP2*, *RP11-481H12.1*, *RNU7-165P*, *RP11-139 K1.2*, *RP11-57H14.3*, and *RP11-214 N15.5*) through chromatin interaction mapping in mesenchymal stem cells.

Table 4. The r^2 for those SNPs within VTI1A reported by JA Morris with rs76140829 and *P*-value in Taiwan Biobank.

SNPs within VTI1A reported by JA Morris	r^2 with rs76140829 in Taiwan Biobank	*P*-value in TWBB2 (Replication)	*P*-value in TWBB1 (GWAS)
rs1248383536			
rs2296782	0.001	.113	.036
rs4575195	1.00×10^{-4}	.999	.245
rs4132670	1.00×10^{-4}	.815	.298
rs7099088	0.048	.002	4.65×10^{-4}
rs10736222	0.022	.007	.520
rs4918760	0.021	.001	.442
rs9645551	0.020	.003	.350
rs10787453	0.020	.003	.446
rs7898876	0.020	.003	.372
rs10885366	0.024	.419	.693
rs10885367	0.025	.293	.857
rs10885368	0.025	.293	.857
rs6585166	0.024	.362	.684
rs7917317	0.024	.244	.771
rs11196046	0.085	.510	.001
rs2859885	0.085	.516	.001
rs11196051	0.046	.197	.009
rs10885370	0.050	.204	.003
rs10885371	0.078	.538	.003
rs11196055	0.050	.111	.004

The *HABP2* gene, known for encoding the Hyaluronan Binding Protein 2 or Factor VII activating protease (FSAP), primarily functions in the blood coagulation system and has roles in inflammation and tissue remodeling. Although its direct relationship with osteoporosis is not extensively established in the literature, we can hypothesize potential connections based on its known functions: (1) Tissue Remodeling and Bone Health: *HABP2's* role in tissue remodeling could be relevant to bone health. Bone remodeling is a continuous process involving bone formation and resorption, crucial for maintaining bone strength and density. Any dysregulation in this process can contribute to conditions like osteoporosis. Proteins involved in tissue remodeling, like FSAP, might influence this balance, although the specific mechanisms would require further research. (2) Inflammation and Bone Metabolism: Chronic inflammation is known to negatively impact bone health, potentially leading to osteoporosis. Since HABP2 is implicated in inflammatory processes, it is conceivable that it could indirectly affect bone metabolism. Inflammatory cytokines can stimulate bone resorption and inhibit bone formation, contributing to the development of osteoporosis. (3) Blood Coagulation and Bone Vascularization: The role of *HABP2* in blood coagulation might also be indirectly linked to bone health through the regulation of bone vascularization. Adequate blood supply is essential for bone health, and any alterations in vascularization can affect bone density and strength. (4) Genetic Studies and Associations: Genetic studies exploring associations between various genes and osteoporosis might reveal potential links between *HABP2* and bone health. Such studies could provide insights into whether variations in the *HABP2* gene correlate with osteoporosis risk. (5) Potential Role in Cell Signaling: Since *HABP2* may be involved in cell signaling pathways that regulate cell growth and differentiation, it might have a role in osteoblast (bone-forming cell) and osteoclast (bone-resorbing cell) activities, although this is speculative and would need scientific validation. In summary, although a direct connection between *HABP2* and osteoporosis is not clearly established, its roles in tissue remodeling, inflammation, and possibly in cell signaling and vascularization, suggest potential pathways through which it could influence bone health. Further research, particularly genetic and molecular studies, would be necessary to clarify these potential relationships.

The *RP11-57H14.3* gene, being categorized as "sense intronic," suggests that it is located within the intron of a sense strand of DNA. Intronic regions, though traditionally considered non-coding, have been found to play roles in gene regulation and expression. They can affect how genes are expressed, which in turn could influence bone density and the risk of osteoporosis.

RP11-139 K1.2, *RP11-481H12.1*, *and RP11-214 N15.5* are pseudogenes. Pseudogenes, like *RP11-139 K1.2*, *RP11-481H12.1*, *and RP11-214 N15.5*, are segments of DNA that are similar to normal genes but are non-functional due to mutations or lack of regulatory elements. Traditionally, pseudogenes were considered as "junk DNA" with no significant function. However, recent studies have suggested that some pseudogenes may have regulatory roles in gene expression and could be involved in various biological processes and diseases. In the context of osteoporosis, a condition characterized by weakened bones and an increased risk of fractures, the potential role of pseudogenes is not well-established. However, it is possible that these pseudogenes could be involved in the regulation of genes that are important for bone density and health. For instance, they might influence the expression of genes involved in bone formation, resorption, or mineralization.

The *RNU7-165P* gene, categorized as a snRNA (small nuclear RNA) gene, is part of a class of RNA molecules that play a role in various cellular processes, including RNA splicing, regulation of transcription factors, and maintenance of telomeres. snRNAs are primarily involved in the processing of pre-messenger RNA in the nucleus. Osteoporosis is a condition characterized by weakened bones, increasing the

risk of fractures. Although the direct connection between *RNU7-165P* and osteoporosis is not clear from the available literature, it is possible that this gene could play a role in bone metabolism or bone cell function through its involvement in cellular regulatory processes.

The availability of Hi-C biological resources can aid in the interpretation of GWAS results. Watanabe's study[30] identified putative causal genes by performing chromatin interaction mapping on outcomes from 3 GWAS studies (BMI, CD, and SCZ) and the additionally identified genes based on chromatin interaction information were mostly located outside of the risk loci and were shown to have shared function with known candidates. Several studies have also identified novel candidates from GWAS risk loci by integrating their results with chromatin interactions.[46-49]

We have confirmed the significance of the newly identified SNP rs76140829 using GWAS analysis and meta-analysis in the TWBB. Through chromatin interaction mapping, we discovered its interactions with 6 genes in mesenchymal stem cells: *HABP2*, *RP11-481H12.1*, *RNU7-165P*, *RP11-139 K1.2*, *RP11-57H14.3*, and *RP11-214 N15.5*. The annotation of rs76140829 included a CADD score of 18.28, a RegulomeDB Categorical Score of 4 (indicating TF binding + DNase peak), and a core 15-state chromatin model in E006 (ESC Derived) H1-Derived Mesenchymal Stem Cells with a score of 5 (representing Tx/Wx, ie, weak transcription). Our research also revealed that rs76140829 influences the affinity of numerous transcription factors and is associated with the binding of the ATF3 transcription factor in bone marrow tissue. Moreover, we determined the PICS of this SNP to be 1, suggesting that rs76140829 has potential functional implications. Bone remodeling, a dynamic process involving the coordinated action of osteoclasts and osteoblasts, plays a crucial role in maintaining bone health. Microcracks in aging bones trigger osteocytes to release signaling factors, which recruit osteoclasts to resorb the damaged bone. Subsequently, osteoblasts are activated and differentiate from mesenchymal stem cells, contributing to bone formation. Considering the therapeutic potential of mesenchymal stem cells in treating osteoporosis, the identification of the rs76140829 polymorphism and its association with osteoporosis-related genes in mesenchymal stem cells provides valuable insights for future drug development and personalized treatments, particularly in Asian populations.

Conclusions

In conclusion, our GWAS using data from the TWBB unveiled rs76140829 within the VTI1A gene as a key risk variant associated with osteoporosis. This specific SNP was greatly associated with osteoporosis and was mapped to 6 genes through chromatin interaction mapping in mesenchymal stem cells. The distinct distribution of allele frequencies for rs76140829 among diverse populations underscores the necessity of acknowledging population-specific genetic variations in osteoporosis investigations. The functional significance and potential therapeutic implications of rs76140829 warrant further investigation. Our findings contribute to the understanding of osteoporosis genetics, particularly within the Asian population, and establish a foundation for future endeavors in drug development and personalized treatment strategies.

Acknowledgments

The authors thank Taiwan Biobank for providing the study data and the Ministry of Science and Technology for providing funding for the study.

Author contributions

Yi-Ching Liaw (Conceptualization, Formal analysis, Methodology, Writing—original draft), Koichi Matsuda (Conceptualization, Methodology, Supervision, Writing—review & editing) and Yung-Po Liaw (Conceptualization, Formal analysis, Methodology, Supervision, Writing—review & editing).

Funding

This study was partly funded by the National Science and Technology Council (Project number: Most 110-2121-M-040-002 and Most 111-221-M-040-002).

Conflicts of interest

None declared.

Data availability

Data are available from the authors upon reasonable request and with permission of Taiwan Biobank.

Ethics approval

Approval for this study was provided by the Institutional Review Board of Chung Shan Medical University (IRB: CS1-20009).

Consent to participate

Taiwan Biobank participants had provided written informed consent during enrollment.

Consent to publish

Not applicable.

References

1. Johnell O, Kanis J. An estimate of the worldwide prevalence and disability associated with osteoporotic fractures. *Osteoporos Int*. 2006;17(12):1726–1733. https://doi.org/10.1007/s00198-006-0172-4.
2. De Martinis M, Sirufo MM, Polsinelli M, Placidi G, Di Silvestre D, Ginaldi L. Gender differences in osteoporosis: a single-center observational study. *World J Mens Health*. 2021;39(4):750–759. https://doi.org/10.5534/wjmh.200099.
3. Rivadeneira F, Mäkitie O. Osteoporosis and bone mass disorders: from gene pathways to treatments. *Trends Endocrinol Metab*. 2016;27(5):262–281. https://doi.org/10.1016/j.tem.2016.03.006.
4. Kanis JA. Diagnosis of osteoporosis and assessment of fracture risk. *Lancet*. 2002;359(9321):1929–1936. https://doi.org/10.1016/S0140-6736(02)08761-5.
5. Ioannidis JP, Ng MY, Sham PC, et al. Meta-analysis of genome-wide scans provides evidence for sex-and site-specific regulation of bone mass. *J Bone Miner Res*. 2007;22(2):173–183. https://doi.org/10.1359/jbmr.060806.
6. Arden N, Baker J, Hogg C, Baan K, Spector T. The heritability of bone mineral density, ultrasound of the calcaneus and hip axis length: a study of postmenopausal twins. *J Bone Miner Res*. 1996;11(4):530–534. https://doi.org/10.1002/jbmr.5650110414.

7. Williams FM, Spector TD. The genetics of osteoporosis. *Acta Reumatol Port*. 2007;32(3):231–240.
8. Mendoza N, Quereda F, Presa J, et al. Estrogen-related genes and postmenopausal osteoporosis risk. *Climacteric*. 2012;15(6):587–593. https://doi.org/10.3109/13697137.2012.656160.
9. Howard GM, Nguyen TV, Harris M, Kelly PJ, Eisman JA. Genetic and environmental contributions to the association between quantitative ultrasound and bone mineral density measurements: a twin study. *J Bone Miner Res*. 1998;13(8):1318–1327. https://doi.org/10.1359/jbmr.1998.13.8.1318.
10. Kemp JP, Morris JA, Medina-Gomez C, et al. Identification of 153 new loci associated with heel bone mineral density and functional involvement of GPC6 in osteoporosis. *Nat Genet*. 2017;49(10):1468–1475. https://doi.org/10.1038/ng.3949.
11. Morris JA, Kemp JP, Youlten SE, et al. An atlas of genetic influences on osteoporosis in humans and mice. *Nat Genet*. 2019;51(2):258–266. https://doi.org/10.1038/s41588-018-0302-x.
12. Visscher PM, Wray NR, Zhang Q, et al. 10 years of GWAS discovery: biology, function, and translation. *Am J Hum Genet*. 2017;101(1):5–22. https://doi.org/10.1016/j.ajhg.2017.06.005.
13. Mills MC, Rahal C. A scientometric review of genome-wide association studies. *Commun Biol*. 2019;2(1):9. https://doi.org/10.1038/s42003-018-0261-x.
14. Xiong D-H, Liu X-G, Guo Y-F, et al. Genome-wide association and follow-up replication studies identified ADAMTS18 and TGFBR3 as bone mass candidate genes in different ethnic groups. *Am J Hum Genet*. 2009;84(3):388–398. https://doi.org/10.1016/j.ajhg.2009.01.025.
15. Zheng HF, Forgetta V, Hsu YH, et al. Whole-genome sequencing identifies EN1 as a determinant of bone density and fracture. *Nature*. 2015;526(7571):112–117. https://doi.org/10.1038/nature14878.
16. Rivadeneira F, Styrkársdottir U, Estrada K, et al. Twenty bone-mineral-density loci identified by large-scale meta-analysis of genome-wide association studies. *Nat Genet*. 2009;41(11):1199–1206. https://doi.org/10.1038/ng.446.
17. Estrada K, Styrkarsdottir U, Evangelou E, et al. Genome-wide meta-analysis identifies 56 bone mineral density loci and reveals 14 loci associated with risk of fracture. *Nat Genet*. 2012;44(5):491–501. https://doi.org/10.1038/ng.2249.
18. Moayyeri A, Hsu Y-H, Karasik D, et al. Genetic determinants of heel bone properties: genome-wide association meta-analysis and replication in the GEFOS/GENOMOS consortium. *Hum Mol Genet*. 2014;23(11):3054–3068. https://doi.org/10.1093/hmg/ddt675.
19. Yadav U, Kumar P, Rai V. Vitamin D receptor (VDR) gene FokI, BsmI, ApaI, and TaqI polymorphisms and osteoporosis risk: a meta-analysis. *EJMHG*. 2020;21(1):1–15. https://doi.org/10.1186/s43042-020-00057-5.
20. Zhou Z, Wang M, Yang J, et al. Genome-wide association analysis reveals genetic variations and candidate genes associated with growth-related traits and condition factor in Takifugu bimaculatus. *Reprod Breed*. 2021;1(2):89–99. https://doi.org/10.1016/j.repbre.2021.05.001.
21. Bae J-H, Park D. Effect of dietary calcium on the gender-specific association between polymorphisms in the PTPRD locus and osteoporosis. *Clin Nutr*. 2022;41(3):680–686. https://doi.org/10.1016/j.clnu.2022.01.020.
22. Al-Barghouthi BM, Farber CR. Dissecting the genetics of osteoporosis using systems approaches. *Trends Genet*. 2019;35(1):55–67. https://doi.org/10.1016/j.tig.2018.10.004.
23. Ralston SH, Uitterlinden AG. Genetics of osteoporosis. *Endocr Rev*. 2010;31(5):629–662. https://doi.org/10.1210/er.2009-0044.
24. Zhu X, Bai W, Zheng H. Twelve years of GWAS discoveries for osteoporosis and related traits: advances, challenges and applications. *Bone Res*. 2021;9(1):23. https://doi.org/10.1038/s41413-021-00143-3.
25. Roadmap EC, Kundaje A, Meuleman W, et al. Integrative analysis of 111 reference human epigenomes. *Nature*. 2015;518(7539):317–330. https://doi.org/10.1038/nature14248.
26. Consortium EP. An integrated encyclopedia of DNA elements in the human genome. *Nature*. 2012;489(7414):57–74. https://doi.org/10.1038/nature11247.
27. Schmitt AD, Hu M, Jung I, et al. A compendium of chromatin contact maps reveals spatially active regions in the human genome. *Cell Rep*. 2016;17(8):2042–2059. https://doi.org/10.1016/j.celrep.2016.10.061.
28. Purcell S, Neale B, Todd-Brown K, et al. PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am J Hum Genet*. 2007;81(3):559–575. https://doi.org/10.1086/519795.
29. Zhou W, Nielsen JB, Fritsche LG, et al. Efficiently controlling for case-control imbalance and sample relatedness in large-scale genetic association studies. *Nat Genet*. 2018;50(9):1335–1341. https://doi.org/10.1038/s41588-018-0184-y.
30. Watanabe K, Taskesen E, Van Bochoven A, Posthuma D. Functional mapping and annotation of genetic associations with FUMA. *Nat Commun*. 2017;8(1):1826. https://doi.org/10.1038/s41467-017-01261-5.
31. Boyle AP, Hong EL, Hariharan M, et al. Annotation of functional variation in personal genomes using RegulomeDB. *Genome Res*. 2012;22(9):1790–1797. https://doi.org/10.1101/gr.137323.112.
32. Kircher M, Witten DM, Jain P, O'roak BJ, Cooper GM, Shendure J. A EnAn itwork for estimating the relative pathogenicity of human genetic variants. *Nat Genet*. 2014;46(3):310–315. https://doi.org/10.1038/ng.2892.
33. 5 RECIacKAMWEJBM. 53 SpmCLH, 8 PiBBECJFEJRHMMA-MARB. Integrative analysis of 111 reference human epigenomes. *Nature*. 2015;518(7539):317–330. https://doi.org/10.1038/nature14248.
34. Ernst J, Kellis M. ChromHMM: automating chromatin-state discovery and characterization. *Nat Methods*. 2012;9(3):215–216. https://doi.org/10.1038/nmeth.1906.
35. Wang K, Li M, Hakonarson H. ANNOVAR: thenotation of genetic variants from high-throughput sequencing data. *Nucleic Acids Res*. 2010;38(16):e164–e164. https://doi.org/10.1093/nar/gkq603.
36. Yang J, Weedon MN, Purcell S, et al. Genomic inflation factors under polygenic inheritance. *Eur J Hum Genet*. 2011;19(7):807–812. https://doi.org/10.1038/ejhg.2011.39.
37. Pe'er I, Yelensky R, Altshuler D, Daly MJ. Estimation of the multiple testing burden for genomewide association studies of nearly all common variants. *Genet Epidemiol*. 2008;32(4):381–385. https://doi.org/10.1002/gepi.20303.
38. Fadista J, Manning AK, Florez JC, Groop L. The (in) famous GWAS P-value threshold revisited and updated for low-frequency variants. *Eur J Hum Genet*. 2016;24(8):1202–1205. https://doi.org/10.1038/ejhg.2015.269.
39. Consortium G, Ardlie KG, Deluca DS, et al. The genotype-tissue expression (GTEx) pilot analysis: multitissue gene regulation in humans. *Science*. 2015;348(6235):648–660. https://doi.org/10.1126/science.1262110.
40. Westra H-J, Peters MJ, Esko T, et al. Systematic identification of trans eQTLs as putative drivers of known disease associations. *Nat Genet*. 2013;45(10):1238–1243. https://doi.org/10.1038/ng.2756.
41. Zhernakova DV, Deelen P, Vermaat M, et al. Identification of context-dependent expression quantitative trait loci in whole blood. *Nat Genet*. 2017;49(1):139–145. https://doi.org/10.1038/ng.3737.
42. Ramasamy A, Trabzuni D, Guelfi S, et al. Genetic variability in the regulation of gene expression in ten regions of the human brain. *Nat Neurosci*. 2014;17(10):1418–1428. https://doi.org/10.1038/nn.3801.

43. NCBI. 2022. https://www.ncbi.nlm.nih.gov/snp/rs76140829. Date accessed September 21, 2022.
44. Huang D, Zhou Y, Yi X, et al. VannoPortal: multiscale functional annotation of human genetic variants for interrogating molecular mechanism of traits and diseases. *Nucleic Acids Res*. 2022;50(D1):D1408–D1416. https://doi.org/10.1093/nar/gkab853.
45. Taylor KE, Ansel KM, Marson A, Criswell LA, Farh KK-H. PICS2: next-generation fine mapping via probabilistic identification of causal SNPs. *Bioinformatics*. 2021;37(18):3004–3007. https://doi.org/10.1093/bioinformatics/btab122.
46. Mifsud B, Tavares-Cadete F, Young AN, et al. Mapping long-range promoter contacts in human cells with high-resolution capture Hi-C. *Nat Genet*. 2015;47(6):598–606. https://doi.org/10.1038/ng.3286.
47. Martin P, McGovern A, Orozco G, et al. Capture Hi-C reveals novel candidate genes and complex long-range interactions with related autoimmune risk loci. *Nat Commun*. 2015;6(1):10069. https://doi.org/10.1038/ncomms10069.
48. McGovern A, Schoenfelder S, Martin P, et al. Capture Hi-C identifies a novel causal gene, IL20RA, in the pan-autoimmune genetic susceptibility region 6q23. *Genome Biol*. 2016;17(1):1–15. https://doi.org/10.1186/s13059-016-1078-x.
49. Javierre BM, Burren OS, Wilder SP, et al. Lineage-specific genome architecture links enhancers and non-coding disease variants to target gene promoters. *Cell*. 2016;167(5):1369–1384.e19e19. https://doi.org/10.1016/j.cell.2016.09.037.