Journal of Bone and Mineral Metabolism (2019) 37:43–52
https://doi.org/10.1007/s00774-017-0892-x

ORIGINAL ARTICLE

Identification of pleiotropic genetic variants affecting osteoporosis risk in a Korean elderly cohort

Eun Pyo Hong[1] · Ka Hyun Rhee[2] · Dong Hyun Kim[3,4] · Ji Wan Park[1]

Received: 1 June 2017 / Accepted: 29 November 2017 / Published online: 22 December 2017

Abstract

Pleiotropy has important implications for understanding the genetic basis and risk assessment of osteoporosis. Our aim was to identify pleiotropic genetic variants associated with the development of osteoporosis and predict osteoporosis risk by leveraging pleiotropic variants. We evaluated the effects of 21 conventional risk factors and 185 single-nucleotide polymorphisms (SNPs) in 63 inflammation- and metabolism-related genes on osteoporosis risk in a community-based Korean cohort study of 1025 participants, the Hallym Aging Study. Ten nongenetic factors, including sex (female) and hematocrit level, and 12 SNPs across ten genes showed evidence of association with incident osteoporosis in 270 initially osteoporosis-free subjects who completed a 6-year follow up. Three gene variants, rs1801282 (*PPARG*-Pro12Ala, hazard ratio (HR) = 3.26, $P = 0.008$), rs1408282 (near *EPHA7*, HR = 1.87, $P = 0.002$), and rs2076212 (*PNPLA3*-Gly115Cys, HR = 2.24, $P = 0.024$), were associated with significant differences in survival among the three genotype groups ($P_{diff} = 0.042$, 0.003, and 0.048, respectively). Individuals in the highest polygenic risk score tertile were 27.9 fold more likely to develop osteoporosis than those in the lowest tertile ($P = 0.004$). The *PPARG* gene in particular was a hub pleiotropic gene in the epistasis network. Our findings highlight pleiotropic modulations of metabolism- and inflammation-related genes in the development of osteoporosis and demonstrate the contribution of pleiotropic genetic variants in prediction of osteoporosis risk.

Keywords Osteoporosis · Pleiotropy · Polygenic risk score · Risk assessment · Survival analysis

Introduction

Osteoporosis, which is the most common metabolic bone disease, is caused by the downregulation of osteoblastic proliferation and the abnormal activation of osteoclasts, which accelerate with age. Osteoporosis is diagnosed when bone mineral density (BMD) decreases to greater than 2.5 standard deviations (SDs) below the average value for young, normal adults (*T* score <−2.5 SD) [1]. Worldwide, more than 200 million women suffer from osteoporosis; and one in three women and one in five men over 50 years of age will have an osteoporotic fracture in their lifetime. Morbidity and mortality associated with osteoporosis increase with the aging of a population, and osteoporotic fractures cause enormous healthcare costs and reduce quality of life [2]. The prevalence of osteoporosis is higher in Asian and European populations than in African populations; in particular, it is much higher in Korean women more than 50 years old (i.e., 38%) than in other ethnic groups [3].

Both genetic and environmental factors have been shown to affect osteoporosis risk; however, family studies have

Electronic supplementary material The online version of this article (https://doi.org/10.1007/s00774-017-0892-x) contains supplementary material, which is available to authorized users.

✉ Ji Wan Park
jwpark@hallym.ac.kr

1 Department of Medical Genetics, College of Medicine, Hallym University, 1 Hallymdaehak-gil, Chuncheon, Gangwon-do 24252, Republic of Korea

2 Department of Medicine, College of Medicine, Hallym University, Chuncheon, Republic of Korea

3 Department of Social and Preventive Medicine, College of Medicine, Hallym University, Chuncheon, Republic of Korea

4 Hallym Research Institute of Clinical Epidemiology, Hallym University Sacred Heart Hospital, Anyang, Republic of Korea

demonstrated that BMD is a highly heritable trait, with an estimated heritability of 0.6–0.9 [4]. Nongenetic factors, such as family history, low body weight, low calcium and vitamin D intake, excessive tobacco use and alcohol consumption, and a sedentary lifestyle, as well as aging and postmenopausal changes in women, have been identified as the predictors of developing osteoporosis [5].

Osteoporosis is a polygenic disorder influenced by the combined effects of multiple genes, each with a small effect. Previous candidate gene studies and genome-wide association studies (GWASs), mostly conducted using a case–control design, have suggested that more than 60 genes, including *VDR*, *ESR1*, *LRP5*, *RUNX2*, and *TNFSF11*, are associated with osteoporosis susceptibility [6, 7]. However, a lack of replication of these associations is common, and a large fraction of genetic heritability is missing from GWAS results for most complex traits, such as osteoporosis [8]. Recent European studies identified rare missense mutations in two genes, *LGR4* and *COL1A2,* and noncoding variants located near the *EN1* gene that were associated with BMD and fractures [9–11].

Accumulating evidence suggests that genetic pleiotropy, the effect of a genetic variant on multiple traits, is prevalent; therefore, identifying pleiotropic genetic variants has important implications for understanding the genetic basis of osteoporosis and the clinical relationship between osteoporosis and other diseases. Previous studies have suggested that cardiovascular disease (CVD) and osteoporosis may be independent predictors of each other [12]. Metabolic traits such as low high-density lipoprotein (HDL) cholesterol levels have long been shown to be associated with risks of both CVD and osteoporosis [13]. Two independent studies have demonstrated pleiotropic effects of the *ITGA1* and *METTL21C* genes on BMD-related traits, type 2 diabetes mellitus (T2DM), and total fat mass (TFM) [14, 15]. Investigating how genetic variants affect multiple traits simultaneously is important for the early detection of at-risk individuals and for identifying novel functional pathways to help develop effective treatments for osteoporosis patients [16]. Incorporating pleiotropic variants into assessment models may improve genetic risk prediction for osteoporosis, although this is not well studied [17].

Prospective cohort studies that measure both genotypes and environmental exposures at baseline and follow-up participants for the development of a disease over time are extremely valuable for identifying individuals at high risk of developing the disease. Longitudinal approaches allow for investigations of the predictive power of genetic variants in assessing individual disease risks [18].

The aims of this study were to investigate whether candidate genes for metabolic traits and inflammation have pleiotropic effects related to osteoporosis and to evaluate the prognostic values of risk prediction models that include pleiotropic genetic variants for incident osteoporosis in a Korean community-based prospective cohort study, the Hallym Aging Study (HAS).

Materials and methods

Study populations and data collection

In 2003, 1510 elderly Koreans aged 45 years or older were recruited by systematic random sampling from 200 areas in Chuncheon City, located in the mid-eastern part of South Korea. The initial survey, in 2004, collected detailed information on epidemiological and clinical variables for 1025 participants (437 men and 588 women) in the Hallym University Chuncheon Sacred Heart Hospital. The second and third waves of data were collected from 702 and 382 participants in 2007 and 2010, respectively. Face-to-face interview surveys were used with self-administered questionnaires after written informed consent was obtained from all research participants. The BMD of the lumbar spine (L1–L4) was measured at baseline by dual-energy X-ray absorptiometry (Hologic QDR 2000 DXA; Hologic, Bedford, MA, USA). Osteoporosis was defined by a BMD *T* score ≤ -2.5 according to the World Health Organization diagnostic criteria for osteoporosis (http://www.4bonehealth.org/education/world-health-organization-criteria-diagnosis-osteoporosis/) [19]. The methods used for recruitment, survey data collection, and genotyping have been described in detail elsewhere [20, 21]. We excluded those subjects who were lost to follow-up or were without adequate genotyping data and 52 patients with osteoporosis at baseline from a total of 1025 participants. Among the remaining 270 osteoporosis-free subjects (97 men, 173 women; mean age, 67.4 years), we identified 45 incident cases of osteoporosis based on their medical history obtained through follow-up interviews. This research protocol was approved by the Institutional Review Board of Hallym University.

Candidate gene selection and genotyping

We initially genotyped 960 participants for 192 SNPs near or within 64 candidate genes mostly reported to be associated with aging-related traits, such as anthropometric, metabolic, and inflammatory traits, using the GoldenGate Assay (Illumina, San Diego, CA, USA). More specifically, we extracted genomic DNA using the FlexiGene DNA Extraction Kit (Qiagen, Valencia, CA, USA). We selected one to eight SNP markers located near or in each candidate gene by searching literature databases, such as 'HuGE Navigator Phenopedia' (https://phgkb.cdc.gov/HuGENavigator/startPagePhenoPedia.do), and by using a web-based SNP selection tool, 'LD Tag SNP Selection (TagSNP)' (http://

manticore.niehs.nih.gov/snpinfo/snptag.php). After excluding 7 SNPs with genotype call rates less than 95%, a final set of 185 SNPs in Hardy–Weinberg equilibrium (HWE) ($P > 0.05$) and with minor allele frequency (MAF) greater than or equal to 1% were included in subsequent analyses. We identified SNPs in strong linkage disequilibrium (LD) ($r^2 \geq 0.8$) using Haploview v. 4.2 (http://www.broadinstitute.org/haploview/haploview).

Statistical analysis

We evaluated the effects of 21 variables, including age, sex, cigarette smoking, alcohol consumption, four anthropometric measures, and 13 clinical variables, on the occurrence of osteoporosis in 270 subjects at risk for osteoporosis. A total of 45 new events were observed: 32 and 13 occurred during the second and third follow-up periods, respectively. Three variables, sex, and plasma hemoglobin (Hb) and hematocrit (HCT) levels, remained significant in the Cox proportional hazard (CPH) multivariate regression analysis after backward elimination ($P < 0.05$). We initially performed CPH analyses to estimate hazard ratios (HRs) and corresponding 95% confidence intervals (CIs) for each SNP marker after adjusting for the effects of four covariates: age, sex, Hb, and HCT (Model 1). We further examined whether the association between a SNP and osteoporosis is maintained even after adjustment for conventional risk factors, such as body mass index (BMI), cigarette smoking, and alcohol consumption (Model 2) [2, 5]. We drew a Manhattan plot to depict the distribution of $-\log_{10}$ (P values) of the associations between 185 SNPs and incident osteoporosis under Model 1 using the R package "*qqman*" (http://cran.r-project.org/web/packages/qqman). We plotted Kaplan–Meier (KM) survival curves and performed log-rank tests to determine whether the survival curves were significantly different among genotypes. The vertical axis of the KM plot, showing survival probabilities, represents the estimated probabilities of not developing osteoporosis. Univariate and multivariate CPH analyses were conducted using STATA version 11.2 (Stata, College Station, TX, USA).

Risk prediction for incident osteoporosis

We developed a weighted polygenic risk score (PRS) model by summing the values obtained from the number of risk alleles at each locus (0, 1, or 2) of the 12 associated SNPs multiplied by their natural log-transformed HR. We categorized the summed PRSs into low-, middle-, and high-risk tertiles as follows: T1, 0.47 to 3.00; T2, 3.01 to 3.70; and T3, 3.71 to 8.40. We further explored potential associations between risk groups and osteoporotic events and generated the KM survival plot along with the corresponding log-rank test P values to compare survival across PRS tertiles. To evaluate the predictability of the PRS model, we estimated Harrell's C index and the symmetrical 95% CI of the model based on the Somers' D nonparametric method using the STATA module 'somersd'.

Functional enrichment and gene network analyses

We analyzed the enrichment of gene ontology (GO) terms to group the 12 SNPs that were identified in the CPH analysis into functionally annotated gene sets using the Database for Annotation, Visualization and Integrated Discovery (DAVID) v. 6.7 (http://david.ncifcrf.gov/). GO terms with Fisher's Exact P values equal to or less than 0.05 were considered to be strongly enriched in the annotation categories. We displayed the functional association network of the identified gene variants in the pathogenesis of osteoporosis using STRING v10 (http://string-db.org/).

Results

Nongenetic risk assessment in the prediction of incident osteoporosis

Of the 270 osteoporosis-free subjects at the initial visit, 45 were newly diagnosed with osteoporosis during the 6-year follow-up. The univariate analyses revealed that 10 of the 21 variables were significantly associated with incident osteoporosis ($P < 0.05$). Specifically, women (HR = 5.66, $P = 3.4 \times 10^{-6}$), as well as slim and short people, had significantly higher osteoporosis risk. High triglyceride and adiponectin levels slightly increased osteoporosis risk. However, high serum creatinine (HR = 0.07, $P = 0.003$), Hb (HR = 0.78, $P = 0.03$), and HCT (HR = 0.90, $P = 0.004$) levels appeared to be associated with reduced osteoporosis risk. Past smoking and current drinking were significantly associated with longer survival (Table 1). Adjustment for sex led to reduced survival in ex-smokers (OR = 2.7), but this association was not significant (data not shown). Among these variables, sex, Hb, and HCT were selected in the stepwise procedure to fit a multivariate CPH model.

Pleiotropic associations of genetic variants related to inflammation and metabolism with incident osteoporosis

Among the 185 SNPs located in or near 63 metabolism-related genes, we identified 15 SNPs in strong LD with each other ($r^2 \geq 0.8$). Twelve SNPs (MAF >5%) located in or near 10 candidate genes yielded promising association signals with incident osteoporosis after adjusting for age, sex, Hb, and HCT (Model 1, $P < 0.05$; Supplementary Fig. 1). These associations were somewhat weakened

Table 1 Anthropometric and clinical data at baseline and two follow-up surveys for 270 initially osteoporosis-free subjects and the results from multivariate Cox proportional hazard models for 45 events that occurred within intervals of 2–3 years

Characteristics	Survey year, number of subjects (%)/event or mean ± SE			CPH[a]		
	2004	2007	2010	HR	95% CI	*P*
Women, *n* (%)[b]	173 (53.9)	173 (53.9)	173 (53.9)	5.66	2.72, 11.75	3.4×10^{-6}
Age, years	67.4 ± 0.4	70.4 ± 0.4	73.4 ± 0.4	1.01	0.98, 1.05	0.531
Weight, kg	61.6 ± 0.6	60.1 ± 0.5	59.6 ± 0.5	0.95	0.92, 0.98	0.002
Height, cm	156.2 ± 0.5	156.0 ± 0.5	155.1 ± 0.5	0.94	0.90, 0.97	1.0×10^{-4}
Body mass index, kg/m^2	25.2 ± 0.2	24.7 ± 0.2	24.7 ± 0.2	0.99	0.90, 1.08	0.754
Waist circumference, cm	86.5 ± 0.5	85.3 ± 0.5	85.9 ± 0.5	0.97	0.94, 1.01	0.109
Systolic blood pressure, mmHg	135.7 ± 1.1	143.4 ± 1.0	137.3 ± 0.9	0.99	0.98, 1.01	0.267
Diastolic blood pressure, mmHg	80.8 ± 0.7	84.4 ± 0.6	78.5 ± 0.5	0.98	0.95, 1.00	0.079
Total cholesterol, mg/dl	200.6 ± 1.9	191.9 ± 2.1	188.6 ± 2.0	1.00	0.99, 1.00	0.979
Triglycerides, mg/dl	168.3 ± 5.5	146.6 ± 4.9	148.7 ± 6.2	1.01	1.00, 1.02	0.040
Aspartate aminotransferase, IU/l	24.6 ± 0.6	23.9 ± 0.7	22.8 ± 0.5	1.01	0.99, 1.02	0.576
Alanine transaminase, IU/l	21.6 ± 0.8	20.0 ± 0.7	18.7 ± 0.6	1.01	0.99, 1.03	0.507
Adiponectin, μg/ml	8.1 ± 0.2	11.3 ± 0.3	10.4 ± 0.3	1.07	1.00, 1.15	0.037
Creatinine, mg/dl	0.9 ± 0.0	0.8 ± 0.0	0.8 ± 0.0	0.07	0.01, 0.41	0.003
Albumin, g/dl	4.6 ± 0.0	4.6 ± 0.0	4.5 ± 0.0	0.60	0.22, 1.64	0.323
C-reactive protein, mg/l	2.0 ± 0.2	2.5 ± 0.4	2.5 ± 0.5	0.89	0.75, 1.07	0.210
Hemoglobin, g/dl[b]	13.6 ± 0.1	14.0 ± 0.1	13.7 ± 0.1	0.78	0.63, 0.98	0.030
Hematocrit, %[b]	40.5 ± 0.2	41.2 ± 0.2	41.6 ± 0.2	0.90	0.84, 0.97	0.004
Fasting glucose, mg/dl	102.2 ± 1.6	103.1 ± 1.4	99.1 ± 1.6	0.99	0.98, 1.01	0.399
Cigarette smoking, *n* (%)						
Never-smoker	195 (60.8)	193 (60.1)	193 (60.1)	Reference		
Ex-smoker	82 (25.6)	91 (28.4)	94 (29.3)	0.35	0.16, 0.79	0.011
Current smoker	44 (13.7)	37 (11.5)	34 (10.6)	0.41	0.14, 1.15	0.089
Alcohol consumption, *n* (%)						
Never-drinker	167 (52.0)	185 (57.6)	175 (54.5)	Reference		
Ex-drinker	35 (10.9)	36 (11.2)	41 (12.8)	0.24	0.06, 0.99	0.048
Current drinker	118 (36.8)	100 (31.2)	105 (32.7)	0.34	0.17, 0.69	0.003
Osteoporosis, *n* (%)						
No	270 (84.1)	256 (79.8)	237 (73.8)	–	–	–
Yes	51 (15.9)	65 (20.3)	84 (26.2)	–	–	–

[a] Hazard ratios (HRs), 95% confidence intervals (CIs), and *P* values were estimated from the Cox proportional hazard (CPH) regression analyses using 45 new osteoporotic events and 225 non-events in subjects followed for 6 years

[b] Three variables—sex, hemoglobin, and hematocrit—were selected with a stepwise backward elimination method in a multivariate Cox proportional hazard regression analysis

after further adjustment for known risk factors, such as BMI, cigarette smoking, and alcohol consumption; however, 4 SNPs located in or near *PPARG* and *EPHA7* genes yielded even stronger associations in Model 2 (Table 2). Furthermore, three SNPs, rs1801282, rs1408282, and rs2076212, were associated with significantly different survival among the three genotype groups based on log-rank tests ($P_{diff} = 0.042$, 0.003, and 0.048, respectively). Specifically, the intergenic SNP, rs1408282, located approximately 97 kb apart from the 3′-end of the ephrin receptor A7 gene (*EPHA7*, 6q16.1) was most significantly associated with incident osteoporosis (HR = 1.87, $P = 0.002$ in Model 1; HR = 1.90, $P = 0.002$ in Model 2). Particularly, AA homozygotes showed a significant association with increased osteoporosis risk when compared with both GA heterozygotes and the reference genotype, GG. The survival probabilities (with the numbers of events and non-events) for GG, GA, and AA genotypes were 88% (21/154), 78% (17/61), and 59% (7/10), respectively (log-rank $P_{diff} = 0.003$) (Fig. 1a). The G allele of the non-synonymous SNP (nsSNP) in the peroxisome proliferator-activated receptor gamma (*PPARG*, 3p25.2) gene,

Table 2 Associations of 12 single-nucleotide polymorphisms (SNPs) with incident osteoporosis in the 6-year follow-up of 270 elderly Koreans

Gene	Chromosome	SNP	N/R[a]	Function	Model 1[a]			Model 2[b]			P^{c}_{diff}
					HR	95% CI	*P*	HR	95% CI	*P*	
CRP	1q23.2	rs2794520[d]	C//T	3′ – 3 kb	1.74	1.09, 2.78	0.020	1.64	1.01, 2.64	0.044	0.106
IL6R	1q21.3	rs2228145	C/A	Asp358Ala	1.59	1.05, 2.42	0.029	1.55	1.01, 2.38	0.044	0.066
TMEM18	2p25.3	rs2293083	C/G	Intron	1.95	1.09, 3.50	0.025	1.91	1.06, 3.43	0.032	0.066
PPARG	3p25.2	rs1801282[d]	C/G	Pro12Ala	3.26	1.36, 7.82	0.008	3.70	1.43, 9.54	0.007	0.042
		rs13306747	C/G	Pro269Pro	2.04	1.00, 4.13	0.049	2.60	1.26, 5.36	0.009	0.084
GATA2	3q21.3	rs2860228	G/A	Intron	1.73	1.05, 2.85	0.032	1.59	0.95, 2.64	0.075	0.568
EPHA7	6q16.1	rs1408282	G/A	3′ – 97 kb	1.87	1.25, 2.78	0.002	1.90	1.26, 2.86	0.002	0.003
		rs345730[d]	A/G	Pro687Pro	1.72	1.06, 2.80	0.028	1.80	1.09, 2.95	0.021	0.545
TSNARE1	8q24.3	rs6583607	C/T	Intron	1.61	1.05, 2.45	0.028	1.52	0.99, 2.31	0.053	0.093
CYP17A1	10q24.3	rs11191548	C/T	5′ – 249 kb	1.74	1.01, 2.99	0.045	1.70	0.98, 2.94	0.057	0.073
MC4R	18q21.32	rs12970134	G/A	3′ – 154 kb	1.57	1.01, 2.43	0.045	1.56	0.99, 2.44	0.054	0.213
PNPLA3	22q13.31	rs2076212	C/A	Gly115Cys	2.24	1.11, 4.52	0.024	2.13	1.05, 4.31	0.036	0.048

N/R non-risk/risk allele

[a,b] Cox proportional hazard (CPH) models were used to estimate the hazard ratios (HRs), 95% confidence intervals (CIs), and *P* values: Model 1 adjusted for age, sex, hemoglobin, and hematocrit; and Model 2 adjusted for Model 1 plus body mass index, cigarette smoking, and alcohol consumption

[c] P_{diff}, *P* value for the log-rank test

[d] SNPs in complete linkage disequilibrium ($r^2 = 1$): rs2794520-rs7553007; rs1801282-rs4684847-rs2197423-rs1899951; rs345730-rs345713

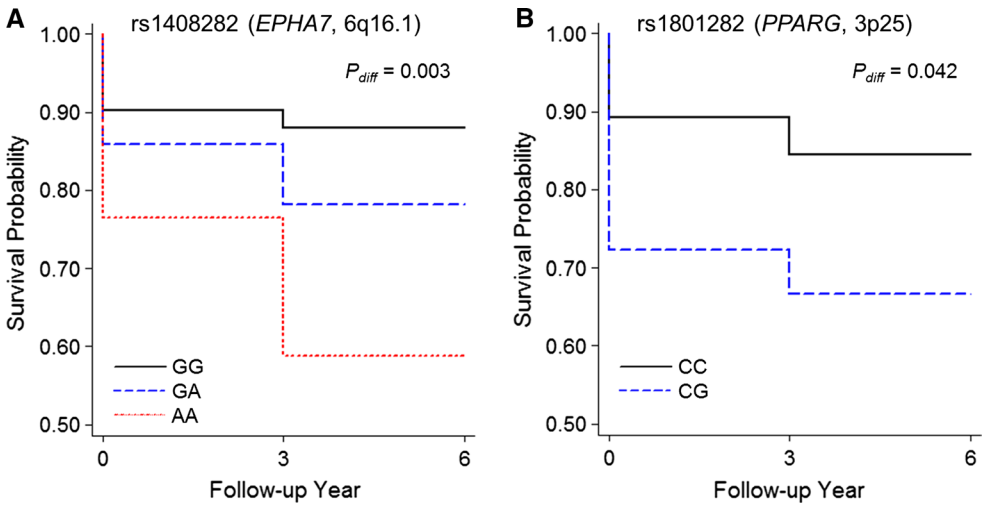

Fig. 1 Kaplan–Meier survival curves and log-rank test *P* values (P_{diff}) for two single-nucleotide polymorphisms (SNPs) that were most significantly associated with incident osteoporosis in the 6-year follow-up data: rs1408282 (*EPHA7*, 6q16.1) (**a**) and rs1801282 (*PPARG*, 3p25) (**b**). Survival rates associated with the three genotypes are indicated by *solid black* (non-risk homozygote), *dashed blue* (heterozygote), and *dotted red* (risk homozygote) *lines*

rs1801282 (Pro12Ala), had the strongest effect size on the risk of developing osteoporosis (HR = 3.26, $P = 0.008$ in Model 1; HR = 3.70, $P = 0.007$ in Model 2); and the CG genotype was associated with worse survival than the CC reference genotype after 6 years of follow-up [i.e., 67% (6/12) vs. 85% (39/213), $P_{diff} = 0.042$] (Fig. 1b). Two additional nsSNPs, rs2076212 (Gly115Cys) and rs2228145 (Asp358Ala), located in the patatin-like phospholipase domain containing three (*PNPLA3*) and interleukin 6 receptor (*IL6R*) genes, respectively, are also promising variants that increase the risk of osteoporosis (Table 2).

Genetic risk prediction for incident osteoporosis

The participants were divided into three (low, middle, and high) osteoporosis risk tertiles based on their PRSs, reflecting the combined effects of 12 independent SNPs (pairwise LD, $r^2 < 0.8$) associated with incident osteoporosis. The results demonstrate a strong association between PRSs and the risk of developing osteoporosis by the 6-year follow-up. Specifically, subjects in the highest PRS group (T3) had a 19.2-fold-greater osteoporosis risk than those in the lowest PRS group (T1) ($P = 0.004$) (Table 3). The

Table 3 A polygenic risk score model for the prediction of osteoporosis risk

Risk model[a]	Event/non-event, *n* (%)[b]	Unadjusted model[c]		Model 1[d]		Model 2[e]	
		HR (95% CI)	*P*	HR (95% CI)	*P*	HR (95% CI)	*P*
Weighted polygenic risk score (PRS, 12 SNPs)							
Tertile 1: 0.47 to 3.00	1 (2)/66 (29)	Reference		Reference		Reference	
Tertile 2: 3.01 to 3.70	9 (20)/54 (24)	10.19 (1.29–80.41)	0.028	13.06 (1.64–104.15)	0.015	12.01 (1.50–96.07)	0.019
Tertile 3: 3.71 to 8.40	35 (78)/104 (47)	19.18 (2.63–139.99)	0.004	27.93 (3.81–204.73)	0.001	26.26 (3.53–195.17)	0.001

[a] Weighted polygenic risk scores (PRS) were stratified into three tertiles (Tertile 1, low risk; Tertile 2, middle risk; Tertile 3, high risk)

[b] The numbers of events in nonosteoporotic subjects during the follow-up period. One subject in the non-event group was excluded for missing genotype data (45 events and 224 non-events)

[c] Hazard ratios (HRs), 95% confidence intervals (CIs), and *P* values were estimated from Cox proportional hazard (CPH) regression analyses

[d,e] HRs, 95% CIs, and *P* values were estimated from Cox proportional hazard (CPH) regression analyses after adjusting for age, sex, hemoglobin, and hematocrit (Model 1); for Model 1 plus three conventional risk factors such as body mass index, cigarette smoking, and alcohol consumption (Model 2), respectively

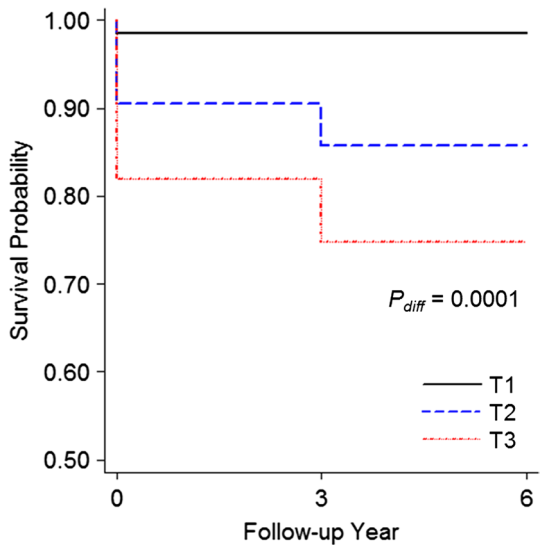

Fig. 2 Kaplan–Meier survival curve and log-rank test *P* value (P_{diff}) for the weighted polygenic risk score (PRS) model comprising 12 SNPs significantly associated with incident osteoporosis in the 6-year follow-up data ($P < 0.05$). Subjects were categorized into three risk groups stratified by their PRSs, and the survival rates of each group are indicated by *solid black* (T1, low risk), *dashed blue* (T2, middle risk), and *dotted red* (T3, high risk) *lines*

association between PRS and osteoporosis risk became more evident after adjustment for four variables, age, sex, Hb, and HCT, selected from 21 potential nongenetic predictors of osteoporosis (Model 1: HR = 27.93, $P = 0.001$). However, this association was somewhat weakened after further adjustment for three other variables, BMI, cigarette smoking, and alcohol consumption, that were previously reported but not selected for the best fitting model in the present study (Model 2: HR = 26.26, $P = 0.001$). Individuals in the middle- and high-risk groups showed much lower survival rates (86% and 75%, respectively) than those in the low-risk group (98%) ($P_{diff} = 0.0001$) (Fig. 2). The prediction accuracy derived from the continuous PRS model (Harrell's $C = 0.73$, 95% CI, 0.64–0.83, $P = 4.9 \times 10^{-37}$) deceased as the PRS model was divided into tertiles (Harrell's $C = 0.67$, 95% CI, 0.60–0.74, $P = 4.1 \times 10^{-52}$).

In silico prediction of gene function and network analyses

We further evaluated biological functions and drew a gene interaction network for 10 candidate genes significantly associated with developing osteoporosis. Among 23 statistically significant GO terms ($P < 0.05$), 4 exceeded the false discovery rate (FDR) of less than 10% (Table 4). The positive regulation of biological process (GO:0048518) was the most enriched in six genes, GATA binding protein 2 (*GATA2*), *EPHA7*, C-reactive protein (*CRP*), *PPARG* , melanocortin 4 receptor (*MC4R*), and *IL6*, among the genes associated with incident osteoporosis ($P = 0.002$, FDR = 3.4%). Three other terms; response to peptide hormone stimulus (GO:0043434), regulation of localization (GO:0032879), and regulation of developmental process (GO:0050793), were also enriched in the subset of the six genes with FDRs less than 10%. In particular, the *PPARG* gene affects a wide range of biological and molecular functions, but the gene product does not contribute to the cellular component (Table 4). A network analysis further supported the importance of *PPARG* as a hub in the interaction network composed of eight candidate genes for osteoporosis, which also included genes involved in inflammation (*CRP* and *IL6R*) and metabolism (*GATA2*, *CYP17A1*, *TMEM18*, *MC4R*, and *PNPLA3*). Human EphA7 protein has a wide range of functions; however, it was not shown to interact with other gene products (Fig. 3).

Table 4 Results of gene ontology enrichment analysis

GO term (*Homo sapiens*)[a]	Gene, *n*	*P*[b]	FDR, %	Gene set
Biological process				
GO:0048518: positive regulation of biological process	6	0.002	3.4	*GATA2, EPHA7, CRP, PPARG, MC4R, IL6R*
GO:0043434: response to peptide hormone stimulus	3	0.003	4.5	*PPARG, MC4R, IL6R*
GO:0032879: regulation of localization	4	0.004	5.4	*GATA2, CRP, PPARG, IL6R*
GO:0050793: regulation of developmental process	4	0.005	7.1	*GATA2, CRP, PPARG, IL6R*
GO:0051094: positive regulation of developmental process	3	0.010	13.5	*GATA2, PPARG, IL6R*
GO:0009893: positive regulation of metabolic process	4	0.012	16.1	*GATA2, PPARG, MC4R, IL6R*
GO:0051239: regulation of multicellular organismal process	4	0.013	16.8	*GATA2, PPARG, MC4R, IL6R*
GO:0051179: localization	6	0.013	17.7	*GATA2, CRP, PPARG, MC4R, IL6R, TSNARE1*
GO:0009725: response to hormone stimulus	3	0.017	22.0	*PPARG, MC4R, IL6R*
GO:0009719: response to endogenous stimulus	3	0.021	25.9	*PPARG, MC4R, IL6R*
GO:0048583: regulation of response to stimulus	3	0.027	32.2	*CRP, PPARG, IL6R*
GO:0042221: response to chemical stimulus	4	0.029	35.1	*CYP17A1, PPARG, MC4R, IL6R*
GO:0045595: regulation of cell differentiation	3	0.030	35.1	*CRP, PPARG, IL6R*
GO:0048699: generation of neurons	3	0.037	42.3	*GATA2, EPHA7, PPARG*
GO:0016192: vesicle-mediated transport	3	0.040	44.1	*GATA2, CRP, TSNARE1*
GO:0022008: neurogenesis	3	0.043	46.8	*GATA2, EPHA7, PPARG*
GO:0006810: transport	5	0.044	48.2	*GATA2, CRP, PPARG, MC4R, TSNARE1*
GO:0006952: defense response	3	0.045	48.2	*CRP, PPARG, IL6R*
GO:0051234: establishment of localization	5	0.046	49.4	*GATA2, CRP, PPARG, MC4R, TSNARE1*
GO:0048468: cell development	3	0.048	51.2	*GATA2, EPHA7, PPARG*
Cellular component				
GO:0016021: integral to membrane	7	0.042	34.6	*CYP17A1, EPHA7, TMEM18, MC4R, IL6R, PNPLA3, TSNARE1*
GO:0031224: intrinsic to membrane	7	0.049	39.6	*CYP17A1, EPHA7, TMEM18, MC4R, IL6R, PNPLA3, TSNARE1*
Molecular function				
GO:0004888: transmembrane receptor activity	4	0.026	25.5	*EPHA7, PPARG, MC4R, IL6R*

[a] Three gene ontology categories

[b] Fisher's exact *P* values and false discovery rates (FDRs) for each GO term were estimated using the DAVID tool

Discussion

Pleiotropic genetic effects cause a wide range of inter-individual variability in humans and may impact their ability to rapidly adapt to new environments [22]. A systematic review of GWASs revealed abundant evidence of pleiotropy in complex human diseases. Many genetic loci, such as *PTPN22* and *APOE4*, show cross-phenotype associations with aging-related diseases such as T2DM, Alzheimer's disease, and atherosclerotic heart disease [23].

In the current study, among 185 SNPs within or near 63 candidate genes related to metabolic and inflammatory traits, we identified 12 SNPs located in or near 10 genes (*CRP*, *IL6R*, *TMEM18*, *PPARG*, *GATA2*, *EPHA7*, *TSNARE1*, *CYP17A1*, *MC4R*, and *PNPLA3*) with pleiotropic effects on incident osteoporosis using 6-year follow-up data. None of these SNPs achieved statistical significance after Bonferroni corrections for multiple hypothesis testing ($P < 0.00027$). The sample size may not be large enough either to detect true associations that survived stringent correction for multiple comparisons or to avoid reporting false-positive associations. However, the Bonferroni correction is considered overly conservative because many SNPs being tested are in substantial LD [24]. Furthermore, two nsSNPs, rs1801282 (*PPARG*, Pro12Ala) and rs2076212 (*PNPLA3*, Gly115Cys), and an intergenic SNP, rs1408282 (*EPHA7*), were associated with differential survival among the different genotype groups. These pleiotropic genes were all shown to contribute to various biological processes based on GO term and KEGG pathway enrichment analyses. Specifically, *PPARG* is a central hub that connects inflammatory genes and metabolic genes in the pleiotropic gene network.

Among these 10 genes, *CRP*, *IL6R*, *PPARG*, *GATA2*, *CYP17A1*, and *MC4R* have been investigated in previous

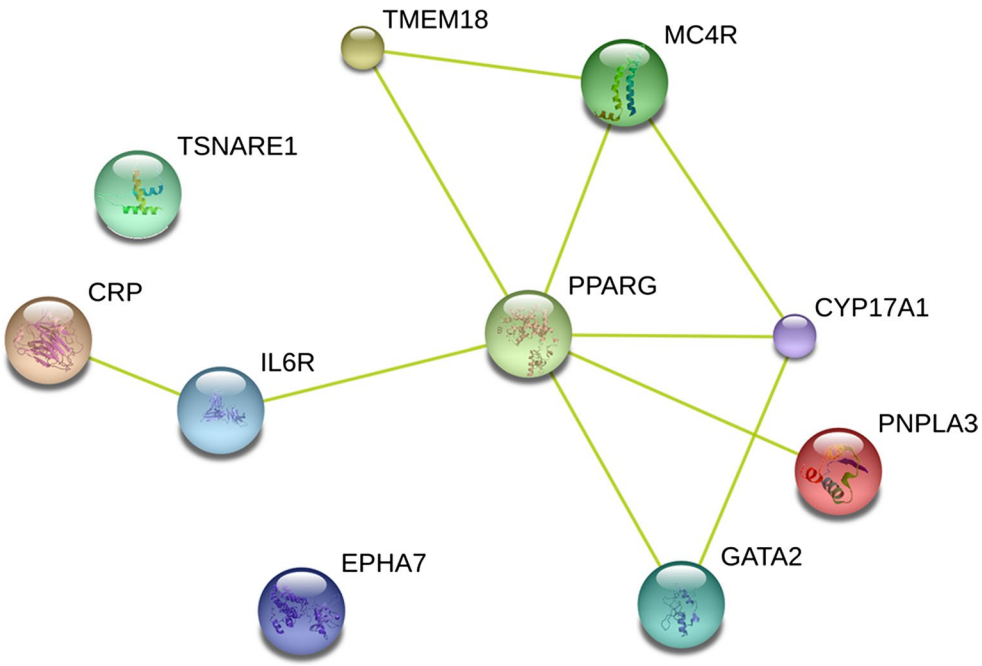

Fig. 3 Gene interaction network for ten candidate genes associated with incident osteoporosis (*Homo sapiens*). The network nodes represent the proteins encoded by the genes

genetic association studies identified with the term 'osteoporosis' in a HuGE Navigator Phenopedia search. Two genes, *PPARG* and *GATA2*, encode transcription factors known to play pivotal roles in both adipogenesis and osteogenesis [25]. The *PPARG* gene has pleiotropic functions in the regulation of lipid metabolism in the liver and skeletal muscle, and is associated with various diseases, such as metabolic syndrome, inflammation, and cancer [26]. A recent study showed that the ERK/MAP kinase phosphorylation of *RUNX2* and *PPARG* transcription factors regulates both osteoblast and adipocyte differentiation in mesenchymal cells [27]. The CCA-to-GCA (Pro12Ala) missense mutation (rs1801282) at codon 12 of the *PPARG* gene, which showed the strongest effect size in the present study (HR = 3.26, $P = 0.008$), has also been associated with low BMD in Chinese postmenopausal women and low serum osteoprotegerin levels in Korean women [28, 29]. Several studies have reported the protective effect of the Ala12 allele against obesity, diabetes, atherosclerosis, inflammation, and malignancy through inhibition of PPARG transcriptional activity [30].

β-Catenin activation followed by *GATA2* expression is required for early osteoclastogenesis [31]. Cytochrome P450 family 17 subfamily A member 1 (*CYP17A1*) is also associated with osteoporosis in postmenopausal women via regulating rate-limiting steps in steroidogenic pathways [32]. The *MC4R* gene is a causal gene for autosomal-dominant obesity, and an MC4R deficiency correlates with increased BMD and decreased bone resorption [33]. The inflammatory markers *CRP*, *IL6*, and *IL6R* are associated with bone resorption in postmenopausal women [34]. Additionally, genetic pleiotropy among CRP, lipids, and CVD provides insights into the shared biology of chronic inflammation and lipid metabolism [35, 36].

The *EPHA7* gene, which showed the most significant association with osteoporosis risk in this follow-up study, is a tumor suppressor causally implicated in cancer development [37]. Bone cell interactions between ephrins and Eph receptors, particularly ephrinB2 and ephrinA2, are involved in bone formation; however, no studies to date have reported a relationship between osteoporosis and the *EPHA7* gene [38]. *PNPLA3* facilitates energy metabolism, and the rs738409 (I148 M) variant has been identified as a pleiotropic marker for metabolic traits, nonalcoholic fatty liver disease, and atherosclerosis; however, no previous studies have reported the effects of this gene on osteoporosis [39]. In the current study, the SNP rs738409 showed no evidence of association ($P = 0.49$, data not shown), but another SNP, rs2076212 (Gly115Cys), was associated with incident osteoporosis (HR = 2.24, $P = 0.024$). The transmembrane protein 18 (*TMEM18*) and t-SNARE domain containing 1 (*TSNARE1*) genes are susceptibility loci for obesity and schizophrenia, respectively, but their association with osteoporosis has not been previously reported [40, 41].

The PRS based on 12 SNPs identified significantly greater risks than any SNP considered alone. However, the most susceptible genotype groups for the *EPHA7* (rs1408282) and *PPARG* (rs1801282) gene variants had higher osteoporotic event rates than those in the highest PRS tertile (41%, 33%, and 25% for AA, CG, and T3 groups, respectively). The continuous PRS model demonstrated better predictive accuracy than the categorical PRS model, as expected (*C*-statistic, 0.73 vs. 0.67). The evidence for this association became

stronger after adjustment for confounding factors. Individuals in the highest PRS tertile were 27.9 fold more likely to develop osteoporosis than those in the lowest tertile after adjusting for age, sex, Hb, and HCT ($P = 0.004$). However, estimating the effect of PRS in the same subjects can lead to overestimation of the true risk of osteoporosis. Therefore, further external validation studies using data from independent cohorts are warranted to evaluate the true effect size of PRS.

Selection bias can occur in a prospective cohort study as a result of differential loss to follow-up. Our calculated likelihood that a subject develops osteoporosis over 6 years might be an underestimate of the true probability if the subjects in better health are more likely to participate over longer follow-up times. In contrast, it would be an overestimate if the subjects lost to follow-up are less likely to develop osteoporosis than those who remained in the study [42]. Heavy tobacco and alcohol use has been shown to increase the risk of osteoporosis; however, a few studies have reported that moderate alcohol consumption and cigarette smoking are associated with reduced osteoporosis risk [43, 44]. Although we could not evaluate the relationship between PRS and BMD because of the lack of follow-up BMD data in the current study, previous studies showed a negative linear relationship between the genetic risk scores composed of BMD-associated SNPs and BMD [45, 46].

Our findings highlight the pleiotropic modulations of metabolism- and inflammation-related genes in the development of osteoporosis, and the results demonstrate that increased predictive power can be achieved by combining pleiotropic variants when developing risk prediction models for osteoporosis. Identifying pleiotropic variants underlying aging-related traits followed by replication in large human cohorts and functional validation in animal models will provide valuable insights into the potential clinical applications of these variants in risk stratification for osteoporosis prevention by reducing environmental risks, modifying lifestyles, and finding new therapeutic targets to treat osteoporosis. Understanding the functional mechanisms of the gene network underlying these comorbidities may lay the foundation for preventing osteoporosis.

Acknowledgments This research was supported by the Basic Science Research Program of the National. Research Foundation of Korea (NRF) funded by the Ministry of Education, Science and Technology (2014R1A1A3053168), the Bio & Medical Technology Development Program of the National Research Foundation of Korea (NRF) funded by the Ministry of Science, ICT & Future Planning (NRF-2012M3A9D1054450), and Hallym University Research Fund 2017 (HRF-201704-015). We are greatly thankful to all participants and staff in the Hallym Aging Study for their contribution to the project.

Compliance with ethical standards

Conflict of interest All authors have no conflicts of interest.

References

1. NIH Consensus Development Panel on Osteoporosis Prevention, Diagnosis, and Therapy (2001) Osteoporosis prevention, diagnosis, and therapy. JAMA 285:785–795
2. Sambrook P, Cooper C (2006) Osteoporosis. Lancet 367:2010–2018
3. Park EJ, Joo IW, Jang MJ, Kim YT, Oh K, Oh HJ (2014) Prevalence of osteoporosis in the Korean population based on Korea National Health and Nutrition Examination Survey (KNHANES), 2008–2011. Yonsei Med J 55:1049–1057
4. Duncan EL, Brown MA (2010) Clinical review 2: genetic determinants of bone density and fracture risk: state of the art specific regulation of bone mass. J Clin Endocrinol Metab 95:2576–2587
5. Thulkar J, Singh S, Sharma S, Thulkar T (2016) Preventable risk factors for osteoporosis in postmenopausal women: systematic review and meta-analysis. J Midlife Health 7:108–113
6. Clark GR, Duncan EL (2015) The genetics of osteoporosis. Br Med Bull 113:73–81
7. Karasik D, Rivadeneira F, Johnson M (2016) The genetics of bone mass and susceptibility to bone diseases. Nat Rev Rheumatol 12:323–334
8. Eichler EE, Flint J, Gibson G, Kong A, Leal SM, Moore JH, Nadeau JH (2010) Missing heritability and strategies for finding the underlying causes of complex disease. Nat Rev Genet 11:446–450
9. Styrkarsdottir U, Thorleifsson G, Sulem P, Gudbjartsson DF, Sigurdsson A et al (2013) Nonsense mutation in the LGR4 gene is associated with several human diseases and other traits. Nature (Lond) 497:517–520
10. Styrkarsdottir U, Thorleifsson G, Eiriksdottir B, Gudjonsson SA, Ingvarsson T, Center JR, Nguyen TV, Eisman JA, Christiansen C, Thorsteinsdottir U, Sigurdsson G, Stefansson K (2016) Two rare mutations in the COL1A2 gene associate with low bone mineral density and fractures in Iceland. J Bone Miner Res 31:173–179
11. Zheng HF, Forgetta V, Hsu YH, Estrada K, Rosello-Diez A et al (2015) Whole-genome sequencing identifies EN1 as a determinant of bone density and fracture. Nature (Lond) 526:112–117
12. Sennerby U, Melhus H, Gedeborg R, Byberg L, Garmo H, Ahlbom A, Pedersen NL, Michaëlsson K (2009) Cardiovascular diseases and risk of hip fracture. JAMA 302:1666–1673
13. Ackert-Bicknell CL (2012) HDL cholesterol and bone mineral density: is there a genetic link? Bone (NY) 50:525–533
14. Billings LK, Hsu YH, Ackerman RJ, Dupuis J, Voight BF et al (2012) Impact of common variation in bone-related genes on type 2 diabetes and related traits. Diabetes 61:2176–2186
15. Zhao F, Gao LH, Li SS, Wei ZY, Fu WZ, He JW, Liu YJ, Hu YQ, Dong J, Zhang ZL (2016) Association between SNPs and haplotypes in the METTL21C gene and peak bone mineral density and body composition in Chinese male nuclear families. J Bone Miner Metab 35:437–447
16. Karasik D, Kiel DP (2010) Evidence for pleiotropic factors in genetics of the musculoskeletal system. Bone (NY) 46:1226–1237
17. Li C, Yang C, Gelernter J, Zhao H (2014) Improving genetic risk prediction by leveraging pleiotropy. Hum Genet 133:639–650
18. Manolio TA, Bailey-Wilson JE, Collins FS (2006) Genes, environment and the value of prospective cohort studies. Nat Rev Genet 7:812–820
19. Jang SN, Choi YH, Choi MG, Kang SH, Jeong JY, Choi YJ, Kim DH (2006) Prevalence and associated factors of osteoporosis among postmenopausal women in Chuncheon: Hallym Aging Study (HAS). J Prev Med Public Health 39:389–396
20. Cho YG, Song HJ, Lee SK, Jang SN, Jeong JY, Choi YH, Hong KS, Choi MG, Kang SH, Kang JH, Kim DH, Caterson I (2009)

The relationship between body fat mass and erectile dysfunction in Korean men: Hallym Aging Study. Int J Impot Res 21:179–186
21. Hong EP, Kim DH, Suh JG, Park JW (2013) Analyses of longitudinal effects of gene-environment interactions on plasma C-reactive protein levels: the Hallym Aging Study. Genes Genom 35:131–139
22. Solovieff N, Cotsapas C, Lee PH, Purcell SM, Smoller JW (2013) Pleiotropy in complex traits: challenges and strategies. Nat Rev Genet 14:483–495
23. Sivakumaran S, Agakov F, Theodoratou E, Prendergast JG, Zgaga L, Manolio T, Rudan I, McKeigue P, Wilson JF, Campbell H (2011) Abundant pleiotropy in human complex diseases and traits. Am J Hum Genet 89:607–618
24. Jorgensen TJ, Ruczinski I, Kessing B, Smith MW, Shugart YY, Alberg AJ (2009) Hypothesis-driven candidate gene association studies: practical design and analytical considerations. Am J Epidemiol 170:986–993
25. Szczerbal I, Foster HA, Bridger JM (2009) The spatial repositioning of adipogenesis genes is correlated with their expression status in a porcine mesenchymal stem cell adipogenesis model system. Chromosoma (Berl) 118:647–663
26. Debril MB, Renaud JP, Fajas L, Auwerx J (2001) The pleiotropic functions of peroxisome proliferator-activated receptor gamma. J Mol Med (Berl) 79:30–47
27. Ge C, Cawthorn WP, Li Y, Zhao G, Macdougald OA, Franceschi RT (2016) Reciprocal control of osteogenic and adipogenic differentiation by ERK/MAP kinase phosphorylation of Runx2 and PPARγ transcription factors. J Cell Physiol 231:587–596
28. Yue H, He JW, Zhang H, Hu WW, Hu YQ, Li M, Liu YJ, Wu SH, Zhang ZL (2010) No association between polymorphisms of peroxisome [corrected] proliferator-activated receptor-gamma gene and peak bone mineral density variation in Chinese nuclear families. Osteoporos Int 21:873–882
29. Rhee EJ, Oh KW, Yun EJ, Jung CH, Park CY, Lee WY, Oh ES, Baek KH, Kang MI, Park SW, Kim SW (2007) The association of Pro12Ala polymorphism of peroxisome proliferator-activated receptor-gamma gene with serum osteoprotegerin levels in healthy Korean women. Exp Mol Med 39:696–704
30. Namvaran F, Rahimi-Moghaddam P, Azarpira N, Dabbaghmanesh MH, Bakhshayeshkaram M, Namvaran MM (2013) Changes in bone biological markers after treatment of Iranian diabetic patients with pioglitazone: no relation to polymorphism of PPAR-γ (Pro12Ala). J Res Med Sci 18:277–282
31. Wei W, Zeve D, Suh JM, Wang X, Du Y, Zerwekh JE, Dechow PC, Graff JM, Wan Y (2011) Biphasic and dosage-dependent regulation of osteoclastogenesis by β-catenin. Mol Cell Biol 31:4706–4719
32. Yilmaz MB, Pazarbasi A, Guzel AI, Kocaturk-Sel S, Kasap H, Kasap M, Urunsak IF, Basaran S, Alptekin D, Demirhan O (2011) Association of serum sex steroid levels and bone mineral density with CYP17 and CYP19 gene polymorphisms in postmenopausal women in Turkey. Genet Mol Res 10:1999–2008
33. Garg G, Kumar J, McGuigan FE, Ridderstråle M, Gerdhem P, Luthman H, Åkesson K (2014) Variation in the MC4R gene is associated with bone phenotypes in elderly Swedish women. PLoS ONE 9:e88565
34. Ferrari SL, Ahn-Luong L, Garnero P, Humphries SE, Greenspan SL (2003) Two promoter polymorphisms regulating interleukin-6 gene expression are associated with circulating levels of C-reactive protein and markers of bone resorption in postmenopausal women. J Clin Endocrinol Metab 88:255–259
35. Hong EP, Kim DH, Suh JG, Park JW (2014) Genetic risk assessment for cardiovascular disease with seven genes associated with plasma C-reactive protein concentrations in Asian populations. Hypertens Res 37:692–698
36. Ligthart S, Vaez A, Hsu YH, Inflammation Working Group of the CHARGE Consortium, PMI-WG-XCP, LifeLines Cohort Study, Stolk R, Uitterlinden AG, Hofman A, Alizadeh BZ, Franco OH, Dehghan A (2016) Bivariate genome-wide association study identifies novel pleiotropic loci for lipids and inflammation. BMC Genomics 17:443
37. Xi HQ, Wu XS, Wei B, Chen L (2012) Eph receptors and ephrins as targets for cancer therapy. J Cell Mol Med 16:2894–2909
38. Matsuo K, Otaki N (2012) Bone cell interactions through Eph/ephrin: bone modeling, remodeling and associated diseases. Cell Adh Migr 6:148–156
39. Posadas-Sánchez R, López-Uribe ÁR, Posadas-Romero C, Pérez-Hernández N, Rodríguez-Pérez JM, Ocampo-Arcos WA, Fragoso JM, Cardoso-Saldaña G, Vargas-Alarcón G (2016) Association of the I148 M/PNPLA3 (rs738409) polymorphism with premature coronary artery disease, fatty liver, and insulin resistance in type 2 diabetic patients and healthy controls. The GEA study. Immunobiology. https://doi.org/10.1016/j.imbio.2016.08.008
40. Liu CT, Young KL, Brody JA, Olden M, Wojczynski MK et al (2014) Sequence variation in TMEM18 in association with body mass index: cohorts for heart and aging research in genomic epidemiology (charge) consortium targeted sequencing study. Circ Cardiovasc Genet 7:344–349
41. Sleiman P, Wang D, Glessner J, Hadley D, Gur RE, Cohen N, Li Q, Hakonarson H, Janssen-CHOP Neuropsychiatric Genomics Working Group (2013) GWAS meta analysis identifies TSNARE1 as a novel schizophrenia/bipolar susceptibility locus. Sci Rep 3:3075
42. Grimes DA, Schulz KF (2002) Cohort studies: marching towards outcomes. Lancet 359:341–345
43. Daniel M, Martin AD, Drinkwater DT (1992) Cigarette smoking, steroid hormones, and bone mineral density in young women. Calcif Tissue Int 50:300–305
44. Felson DT, Zhang Y, Hannan MT, Kannel WB, Kiel DP (1995) Alcohol intake and bone mineral density in elderly men and women. The Framingham Study. Am J Epidemiol 142:485–492
45. Estrada K, Styrkarsdottir U, Evangelou E, Hsu YH, Duncan EL et al (2012) Genome-wide meta-analysis identifies 56 bone mineral density loci and reveals 14 loci associated with risk of fracture. Nat Genet 44:491–501
46. Warrington Nicole M, Kemp John P, Tilling Kate, Tobias Jonathan H, Evans David M (2015) Genetic variants in adult bone mineral density and fracture risk genes are associated with the rate of bone mineral density acquisition in adolescence. Hum Mol Genet 24:4158–4166