

# Genome-wide association meta-analysis identifies seven loci for osteocalcin levels

Katerina Trajanoska\*, Alexander Teumer\*, David Karasik, Andre G Uitterlinden, Fernando Rivadeneira, Douglas P Kiel, Yi-Hsiang Hsu, on behalf of CHARGE and GEFOS consortia

*\*denotes equal contribution in preparation*

*In preparation*

## ABSTRACT

Osteocalcin, the most abundant noncollagenous bone matrix protein involved in bone mineralization and bone turnover, is encoded by the *BGLAP* gene; however, there is no data on the influence of common genetic variation on osteocalcin levels. We performed a genome-wide association study for osteocalcin levels including 20,922 European individuals. Genome-wide associations were tested using natural log transformed osteocalcin levels adjusting for age, age<sup>2</sup>, sex, BMI and ancestral genetic background estimated by PCA. Inverse variance meta-analysis was performed, and significance was set at  $P \leq 5 \times 10^{-8}$ . LD-score regression was used to estimate heritability and genetic correlation (shared heritability) with other traits. SNP-based heritability was estimated as 0.09 (SE: 0.02). Variants at seven loci reached genome-wide significance (GWS): 1q22 (*SMG5*,  $p=3.66 \times 10^{-25}$ ), 8q24.12 (*TNFRSF11B*,  $p=3.72 \times 10^{-22}$ ), 11q13.1 (*SSSCA1-AS1*,  $1.75 \times 10^{-9}$ ), 11q24.3 (*FLI1*,  $p=5.68 \times 10^{-9}$ ), 13q14.11 (*AKAP11*,  $p=8.2 \times 10^{-18}$ ), 16p13.3 (*TSC2*,  $p=7.24 \times 10^{-9}$ ) and 19q13.41 (*FPR3*,  $p=1.01 \times 10^{-8}$ ); after validation in five cohorts ( $n=9,467$ ). Osteocalcin had modest negative genetic correlations (shared heritability) with BMD (femoral neck BMD  $rg:-0.27$ ,  $se:0.10$ ; lumbar spine BMD  $rg:-0.35$ ,  $se:0.12$ ), type 2 diabetes ( $rg:-0.26$ ,  $se:0.11$ ) and several metabolic traits and diseases such as waist-to-hip ratio ( $rg:-0.27$ ,  $se:0.09$ ) and triglycerides ( $rg:-0.27$ ,  $se:0.08$ ). MR analysis did not show any causal effect of osteocalcin on several glycemic traits. The results of this study highlight the polygenicity of serum osteocalcin levels and offers insights into its regulation in humans. Although osteocalcin was genetically correlated with several metabolic traits its causal role was not confirmed.

## INTRODUCTION

Osteocalcin (gamma-carboxyglutamic acid protein, BGLAP), a bone-derived hormone, comprises about 15% of the noncollagenous protein component of the extracellular bone matrix and is the most abundant noncollagenous protein of the human body.<sup>1</sup> Serum osteocalcin is a sensitive and specific marker of osteoblastic activity and reflects the rate of bone formation.<sup>2</sup> Currently, it is used as a non-invasive bone formation biomarker for the clinical assessment of calcium-phosphorus metabolic disturbances of bone in diseases or drug trials.<sup>2</sup>

The osteocalcin protein contains gamma carboxyglutamate (Gla) residues which are carboxylated by the enzyme gamma glutamyl carboxylase, a process that is highly dependent to vitamin K. The carboxylated osteocalcin is secreted from the osteoblasts and incorporated in the bone matrix by binding to hydroxyapatite in a calcium-dependent fashion. Some of its calcium-binding properties led to the suggestion that osteocalcin is involved in the bone matrix mineralization.<sup>3</sup> The extracellular bone matrix provides elasticity and flexibility to bone and also determines its structural organization. Uncarboxylated or undercarboxylated osteocalcin residues cannot be incorporated in the matrix and are release into the circulation. Intact serum osteocalcin reflects the portion of newly synthesized osteocalcin that does not bind to the mineral phase of the bone but is released directly into the circulation.

In addition to its involvement in bone formation, osteocalcin has been suggested to play an important role in endocrine regulation of energy metabolism,<sup>5</sup> male fertility,<sup>6</sup> exercise capacity,<sup>7</sup> brain development,<sup>8</sup> and age-related decline in cognition.<sup>9</sup> Studies suggested that osteocalcin regulates glucose metabolism, insulin secretion and adiponectin expression via directly interacting with  $\beta$ -cells in the pancreas and adipocytes in the adipose tissues as a feedback loop of Leptin's regulation of osteoblasts.<sup>10,11</sup> Osteocalcin-deficient mice are obese with decreased  $\beta$ -cell proliferation and greater insulin resistance.<sup>5</sup> Osteocalcin-treated deficient mice had a significantly decrease in blood glucose and an increased insulin secretion. Moreover, wild type mice treated with non-carboxylated osteocalcin improved insulin sensitivity (partially through the enhancement of  $\beta$ -cell proliferation); increased  $\beta$ -cell insulin secretion, energy expenditure and adiponectin expression; and were protected against the development of obesity and type 2 diabetes (T2DM) induced by high-fat diet or hyperphagia.<sup>12</sup>

Although, as described above, strong evidences demonstrated osteocalcin plays an important role in endocrine regulation of energy metabolism; epidemiological studies, on the other hand, have not been conclusive. Low concentrations of baseline serum osteocalcin were strongly associated with an increased risk of incident diabetes during a 5-year follow-up.<sup>13</sup> In incident diabetic subjects, concentrations of serum osteocalcin were inversely associated with insulin resistance (HOMA-IR). In non-diabetic

men exposure to higher plasma osteocalcin during a 3-year follow-up was associated with a significantly lower rise in fasting plasma glucose.<sup>14</sup> However, no correlation between baseline serum osteocalcin and the incidence of diabetes was observed after a 3-year follow-up in a high-risk cohort with T2D.<sup>15</sup> In addition, bisphosphonate treatment for osteoporosis reduced both carboxylated and non-carboxylated forms of osteocalcin, but no association was found between these changes and glucose parameters after 16 weeks of treatment in subjects with impaired fasting glucose.<sup>16</sup> The variation of serum osteocalcin due to environmental and/or genetic factors may confound the findings from observation studies. For example, an increase of 16%-70% in serum osteocalcin in winter compared with that of summer was found in some studies.<sup>17,18</sup> Serum osteocalcin appears to be slightly higher in males at all ages.<sup>19,20</sup> Whereas an age-related decline of serum osteocalcin 19 positively correlated with estradiol concentration, was observed in women.<sup>21</sup>

The estimated heritability of serum osteocalcin was 62% in a Mexican American population, suggesting that the variation of serum osteocalcin may also be due to genetic factors.<sup>22</sup> A common SNP, rs1800247, located at the 5' UTR of the osteocalcin encoded gene, bone gamma-carboxyglutamate protein, BGLAP, was reported to be associated with lower serum osteocalcin,<sup>23</sup> but didn't replicate in other studies.<sup>24,25</sup> Given that the BGLAP gene doesn't seem to explain the variability of serum osteocalcin, seeking to identify polygenetic factors that contribute to serum osteocalcin variation and to causally infer relations between serum osteocalcin and its effect on metabolic disorders, we conducted a GWAS meta-analysis on 13 studies of European-descent adults and replicated top genome-wide associated findings in additional five independent studies. The replicated SNPs were used as instrumental variables for Mendelian randomization (MR) analyses to examine the causal effect of osteocalcin on metabolic and diabetic relevant disorders.

## SUBJECTS AND METHODS

### Study Subjects, Genotyping and Imputation

The Discovery stage of this GWAS meta-analysis utilized data from 11 cohort studies, including AGEs, Framingham, GOOD, HealthABC (European Ancestry), Mr.OS Sweden, Mr.OS US, PIVUS, SHIP, SHIP-Trend, Rotterdam study ( I, II and III) and TwinsUK cohorts. A total of 20,922 European-descent adults were included in the final analyses. The independent replication of top associated findings, including genome-wide associated SNPs (association p-values  $\leq 5 \times 10^{-8}$ ) and genome-wide suggestive associated SNPs ( $5 \times 10^{-8} \leq$  association p-values  $\leq 5 \times 10^{-6}$ ), were performed in five additional studies, including DECODE, BPROOF, GINLISA, Health2006 and OPRA cohorts. A total of 9,467

adult European-descent adults were included in the replication analyses. Distribution of age, sex, BMI and T2DM prevalence in the participating discovery and replication cohorts is available in the Supplementary Materials (**Supplementary table 1**). All studies were approved by institutional ethics review committees at the relevant organizations and all participants provided written informed consent.

All the cohorts were genotyped using commercially available SNP arrays from Affymetrix (Affymetrix Inc., Santa Clara, CA, USA) or Illumina (Illumina Inc., San Diego, CA, USA). The detail of genotype arrays, calling algorithms, quality controls for genotyping, inclusion criteria of SNPs and subjects is available in the Supplementary Materials (Table S2). Imputation was performed within each study with IMPUTE,<sup>26</sup> IMPUTE2,<sup>27</sup> MACH<sup>28</sup> or Minimac<sup>29</sup> software (**Supplementary table 2**) using genotypes from the HapMap Phase II release 22, NCBI build 37 (CEU samples) or HRC1.1<sup>30</sup> as reference panels. Before performing an analysis, poorly imputed and low frequency or rare polymorphisms were excluded. Specifically, the quality control filters applied for exclusions of SNPs were: imputation quality score <0.3 for MACH/Minimac and <0.4 for IMPUTE/IMPUTE2, average minor allele frequency (MAF) of <1% across studies, and SNPs missing from ≥50% of the cohorts. After quality control, ~2.5 million SNPs were available from each cohort for the discovery meta-analysis.

## Serum Osteocalcin Measurement

Serum total osteocalcin was measured by assays with different specificities including chemiluminescence assay, electrochemiluminescence immunoassay, ELISA, EIA and RIA. The intact osteocalcin is 49 amino acid long. The N-terminal part of the osteocalcin includes molecule fragment from amino acid#1 to amino acid#19. The C-terminal part of the osteocalcin includes molecule fragment from amino acid#44 to amino acid#49. The middle “mid” part of the osteocalcin includes molecule fragment from amino acid#20 to amino acid#43. The majority of these assays capture the intact osteocalcin and/or large N-terminal-mid molecule (fragment amino acid #1 to amino acid#43). For example, RIA was used to measure serum osteocalcin in Framingham Study participants. This RIA assay captures both intact osteocalcin (49 amino acids) and the large N-terminal-mid molecule fragment (amino acid#1 to amino acid#43). Given these differences the serum osteocalcin levels may not be directly comparable across studies. As showed in **Supplementary table 1**, the average serum osteocalcin level in men is 26.3, 16.5 and 4.4, respectively, in participants from GOOD Study (using ELISA), MROS (using EIA) and Framingham Study (using RIA). The distribution of serum osteocalcin across studies as well as the assays they used are available in the Supplementary Materials (**Supplementary table 1**). The non-carboxylated form of osteocalcin can be further separated from the carboxylated form. As showed in Table S1, the % of the non-carboxylated osteocalcin among total serum osteocalcin is 18%,

34% and 19%, respectively, in Framingham, HealthABC and MROS US studies. Since the other studies didn't measure non-carboxylated osteocalcin, we only perform GWAS analyses on the total osteocalcin level.

## SNP-Osteocalcin association Analyses

In the Discovery and replication phases, each cohort conducted analyses according to a standard pre-specified analysis plan under an additive (i.e. per allele count) genetic model. Both sex-specific and sex-combined analyses were performed. The phenotype was log<sub>2</sub> transformed serum total osteocalcin level. The covariate adjustment in the model included sex (only for sex-combined models), age, age<sup>2</sup>, BMI and ancestral genetic background estimated by PCA. For family-based studies with pedigree structures, such as the Framingham Study, a mixed-effect model to take into account the within familiar correlation was applied. The mixed-effect model was implemented in the R Kinship2 package ([https://cran.rproject.org/web/packages/available\\_packages\\_by\\_name.html](https://cran.rproject.org/web/packages/available_packages_by_name.html)). Study-specific summary statistics (beta-coefficients, standard errors, p-values and sample size) were used for the meta-analysis with the genome-wide SNPs.

## Meta-Analysis

Meta-analysis of the GWAS discovery results was conducted and tested independently by two researchers. Due to the relatively limited statistical power to detect sex-specific associations, we only present results from sex-combined meta-analysis. Given the relative incomparable serum total osteocalcin measurements across studies, we applied a fixed-effect, sample size weighted meta-analysis implemented in METAL software.  $\lambda$  genomic control was estimated and applied to meta-analysis to correct for potential inflation of the test statistics. The QQ plot was generated by R package (EasyQQ). The regional plots were generated by LocusZoom,<sup>31</sup> using chromosome position coordinates as provided in GrCh37/hg19.  $I^2$  was estimated per SNP to quantify heterogeneity across studies. We consider genome-wide level of statistical significance at p-value  $\leq 5 \times 10^{-8}$  and the suggestive level of significance at p-value  $\leq 5 \times 10^{-6}$ .

## Conditional Analyses for Independent Signals

To identify independent association signals in the regions containing SNPs that were genome-wide significant or genome-wide suggestively significant, we performed chromosomal-wide association analyses conditioning on the most significantly associated SNP(s) from locus (loci) on the same chromosome. This was done by using variance-component models implemented in the GCTA (Genome-wide Complex Trait Analysis) tool package<sup>32</sup> with summary statistics from discovery meta-analysis along

with LD information from HRC1.1 using the BPROOF study.<sup>33</sup> Conditionally independent variants were then annotated to the closest gene using the Ensembl annotation system implemented by SNP-NEXUS (<https://snp-nexus.org/index.html>).

## Replication Analysis

In each genome-wide associated locus, suggestive associated locus or genome-wide associated independent signals in the HapMap effort we selected the lead SNP with the lowest association p-value for de-novo replication in three cohorts. In addition, SNP with  $p < 5 \times 10^{-6}$  from the combined discovery (HapMap and HRC1.1) were followed for replication in three cohorts with available genome-wide data. The SNP-osteocalcin association analysis has been described in the "Osteocalcin Association Analyses" section. A fixed-effect meta-analysis was performed to combine results from these four replication studies. The meta-analysis approach has been described in the "Meta-Analysis" section.

## Meta-Analysis of Results from Discovery and Replication Studies

We further combine summary statistics from the discovery studies and replication studies for those SNPs with replication results. The meta-analysis was done by using the fixed-effect sample size weighted meta-analysis implemented in METAL software. These association p-values are considered as "combined association p-values". A successful replication was considered if the combined association p-values were genome-wide significant ( $p \leq 5 \times 10^{-8}$ ) and lower than the discovery.

## Functional Annotations

We used the FUMA online platform<sup>34</sup> to obtain gene-based, gene-set and tissue-specific annotations. In addition, for all variants, we predicted their function by PolyPhen2.<sup>35</sup> For all variants, we annotated potential regulatory functions of our replicated GWAS SNPs and loci based on experimental epigenetic evidence including DNase hypersensitive sites, histone modifications, and transcription factor binding sites in human cell lines and tissues from the ENCODE<sup>36,37</sup> and the Epigenetic Roadmap projects.<sup>38</sup> We first constructed haplotype blocks for GWAS lead SNPs and SNPs in high LD ( $r^2 > 0.8$ ) with GWAS lead SNPs. We then identified regulatory elements (enhancers and promoters estimated by chromatin state) in the haplotype blocks across 98 healthy human tissues/normal cell lines available in the ENCODE Project and the Epigenetic Roadmap Project using HaploReg4 web browser.<sup>39</sup> GWAS lead SNPs or the SNPs in high LD with GWAS lead SNPs were annotated as to whether they were located within these regulatory elements, including, promoters, flanking active promoters, active promoters, weak promoters, strong and weak enhancers, weak enhancers in transcribed regions,

weak enhancers, active enhancers, active enhancers with weak H3K4me3 and strong H3K27ac and poised enhancers estimated by the ChromHMM approach.<sup>40</sup>

## SNP Heritability of Osteocalcin and Genetic Correlations with Energy Metabolic Traits

LD score regression was used to estimate the SNP heritability ( $h^2$ ) of serum osteocalcin and to estimate the genetic correlation ( $r_g$ ) between serum osteocalcin and variety of bone and energy metabolic traits and diseases with publicly available summary GWAS data using LD Hub.<sup>41</sup> The tool is a centralized database of summary-level GWAS results for hundreds of diseases/traits, as well as a web interface that automates the LD score regression analysis pipeline.<sup>42</sup>

## Mendelian Randomization Analysis

To determine the causal relation of serum total osteocalcin with the risk of T2D, metabolic relevant disorders, and bone health we selected only GWAS SNPs associated with serum total osteocalcin as instrumental variables from the combined meta-analysis. To estimate the per SNP causal relation between serum total osteocalcin and disease risks, we used the 2-sample MR analytical approach. To increase statistical power, we combined each SNP result into a multi-instrument summary-level MR variable.<sup>43</sup> One of the fundamental assumptions of Mendelian randomization analyses is that there should be no pleiotropy in genetic variants; thus, the genetic variants associated with serum total osteocalcin are exclusively associated with disease risks through only serum total osteocalcin, but no other confounders. To avoid bias of Mendelian randomization analyses from pleiotropic genetic effects, we applied the MR Egger<sup>44</sup> and the Mendelian Randomization Pleiotropy RESidual Sum and Outlier (MR-PRESSO).<sup>45</sup> We also performed a bidirectional MR to test for the possibility that BMD may causally affect the levels of total serum osteocalcin. Given several SNPs were associated with both BMD and osteocalcin instead of removing the SNPs we applied Steiger filtering<sup>46</sup> to estimate the true direction of the SNP-trait association.

## RESULTS

### Genome-Wide Associated Loci and Replication

Sex-combined GWAS meta-analyses were performed on 20,922 European adults. We identified six GWAS loci, including chr1q22 SMG5-BGLAP; chr8q24 TNFRSF11B, chr11q13.1 SSSCA1-AS1, chr11q24.3 FLI1-KCNJ1, chr13q14.11 AKAP11, chr16p13.3 TSC2 (Table 1; Supplementary Figure 1-Manhattan plot). The QQ plots and  $\lambda$  genome control values of GWAS meta-analyses didn't demonstrate any potential systemic

Table 1 | Conditionally independent SNPs associated with serum total osteocalcin levels ( $p < 5 \times 10^{-6}$ ).

Locus	SNP	Closest Gene	Annotation	EA	NEA	EAF	Discovery ( $N_{\text{max}}=20,922$ )			Replication ( $N_{\text{max}}=9,467$ )			Combined ( $N_{\text{max}}=30,389$ )		
							Effect	se	P	Effect	se	P	Effect	se	P
1q22	rs2246476	SMG5-BGLAP	intronic	T	C	0.780	0.037	0.004	$2.64 \times 10^{-19}$	0.057	0.016	$3.2 \times 10^{-4}$	0.038	0.004	$3.66 \times 10^{-25}$
1q22	rs2104404	TMEM79	non-coding intronic	A	C	0.201	-0.029	0.004	$6.94 \times 10^{-11}$	-0.020	0.019	0.30	-0.030	0.004	$1.83 \times 10^{-12}$
8q24.12	rs10955908	TNFRSF11B	intergenic	A	C	0.467	-0.028	0.003	$4.53 \times 10^{-16}$	-0.038	0.013	$2.7 \times 10^{-3}$	-0.030	0.003	$3.73 \times 10^{-22}$
11q13.1	rs1346	SSSCA1-AS1	non-coding	A	T	0.812	0.025	0.005	$2.16 \times 10^{-8}$	0.018	0.020	0.36	-0.021	0.004	$1.75 \times 10^{-9}$
11q24.3	rs1787971	FLJ1.1C/NJ1	intergenic	C	G	0.670	-0.020	0.004	$7.73 \times 10^{-8}$	-0.045	0.016	$4.3 \times 10^{-5}$	0.026	0.004	$5.68 \times 10^{-9}$
13q14.11	rs9594738	TNFSF11	intergenic	T	C	0.498	0.030	0.003	$5.04 \times 10^{-19}$	0.005	0.015	0.79	0.029	0.003	$8.02 \times 10^{-19}$
16p13.3	rs13337626	TSC2	synonymus	T	C	0.916	0.039	0.007	$4.82 \times 10^{-8}$	0.036	0.042	0.40	0.041	0.007	$7.24 \times 10^{-9}$
19q13.41	rs8111274	FPR3	intronic	T	G	0.291	-0.021	0.004	$8.25 \times 10^{-8}$	-0.018	0.016	0.28	-0.021	0.004	$1.01 \times 10^{-8}$
1p36.33	rs12758705	C1orf222	intronic	A	G	0.269	0.022	0.004	$3.06 \times 10^{-8}$	-0.012	0.017	0.48	0.021	0.004	$1.43 \times 10^{-7}$
1p36.33	rs12081179	TMEM79	non-coding intronic	A	G	0.794	-0.022	0.005	$1.47 \times 10^{-6}$	0.007	0.018	0.71	-0.021	0.004	$2.66 \times 10^{-6}$
1p21.1	rs4908190	OLFM3	non-coding intronic	A	T	0.448	-0.017	0.004	$2.26 \times 10^{-6}$	-0.020	0.015	0.19	-0.017	0.003	$4.92 \times 10^{-7}$
1q41	rs1538742	RP11-95P13.2	LincRNA	A	C	0.589	-0.016	0.003	$4.56 \times 10^{-6}$	-0.045	0.016	$4.8 \times 10^{-3}$	-0.017	0.003	$4.11 \times 10^{-7}$
4p14	rs1386625	RP11-83C7.1	LincRNA	A	G	0.089	0.029	0.006	$1.42 \times 10^{-6}$	-0.007	0.040	0.85	0.029	0.006	$1.44 \times 10^{-6}$
6q22.31	rs1688630	CEP85L	intronic	T	C	0.018	-0.072	0.014	$3.63 \times 10^{-7}$	0.075	0.053	0.16	-0.061	0.014	$9.20 \times 10^{-6}$
6q25.3	rs2181190	SYNJ2	intronic	T	C	0.784	-0.020	0.004	$2.90 \times 10^{-6}$	0.010	0.019	0.61	-0.018	0.004	$2.04 \times 10^{-5}$
7p21.1	rs2214919	AC005062.2	non-coding intronic	T	C	0.810	-0.023	0.004	$7.65 \times 10^{-8}$	0.036	0.028	0.20	-0.022	0.004	$3.55 \times 10^{-7}$
10q21.2	rs11597844	ZNF365	intronic	T	C	0.063	-0.040	0.009	$2.17 \times 10^{-6}$	0.008	0.049	0.88	-0.039	0.008	$3.69 \times 10^{-6}$
11p15.1	rs4325274	SOX6	intronic	C	G	0.269	-0.018	0.004	$3.84 \times 10^{-6}$	-0.001	0.017	0.94	-0.016	0.004	$1.63 \times 10^{-5}$
18q21.33	rs884205	TNFRSF11A	3utr	A	C	0.268	0.019	0.004	$1.77 \times 10^{-6}$	0.012	0.028	0.96	0.018	0.004	$4.29 \times 10^{-6}$
19q13.41	rs7254413	FPR3	intronic	A	G	0.431	-0.017	0.004	$1.90 \times 10^{-6}$	-0.029	0.015	0.05	-0.018	0.003	$1.91 \times 10^{-7}$
20q12	rs6129693	RNA5SP484	rRNA	A	C	0.426	-0.017	0.003	$8.68 \times 10^{-7}$	-0.005	0.023	0.85	-0.017	0.003	$8.31 \times 10^{-7}$

EA=effect allele; NEA=non-effect allele; EAF=effect allele frequency; se= standard error; the effect and p-value are for the effect allele

biases (**Supplementary Figure 2**). The most significantly associated SNP, rs2246476, is in the intron region of the SMG5 gene with p-value equal to  $2.64 \times 10^{-19}$ . The minor allele frequency is 22% in the studied cohorts. We did not observe any differences between the inverse-variance weighted and z-score weighted meta-analysis (Table S3). Therefore, in all post GWAS analysis we used the summary statistics from the IVW meta-analysis.

### The GWAS Secondary Signals

To identify independent GWAS signals, we performed a conditional analysis by conditioning on GWAS SNPs. Summary statistics from model 1 and HRC1.1 CEU LD structure were used for the conditional analyses. On Chr1 SMG5 locus (**Supplementary Figure 1**), SNP rs2104404 (**Table 1**), located in the *TMEM79* gene, remained genome-wide significantly associated with serum total osteocalcin after conditioning on the most significant SNP (rs2246476). SNP rs2104404 is not in LD with rs2246476 ( $R^2=0.03$ ,  $D'=0.56$ ).

### Combined Meta-analysis

The top associated SNPs from the HapMap effort and SNPs with  $p < 5 \times 10^{-6}$  from the total discovery effort were selected for replication. Replication genotyping (including de novo genotyping and genome-wide genotyping) were done in 9,467 Caucasian samples. All GWS loci were successfully replicated in the combined meta-analysis. In addition, one suggestive locus on chr19q13.41 reached GWS in the combined meta-analysis along with 14 suggestive loci (**Table 1**). Overall, seven GWAS loci were identified in the combined meta-analysis. **Supplementary Figure 2** and **Supplementary Figure 3** are the regional plots and forest plots.

### Annotation and Functional Role of Associated Variants

We annotated SNPs with discovery p-values  $\leq 5 \times 10^{-6}$  from those seven replicated GWAS loci (**Supplementary Table 4**). Except for SNP rs2073618, the other SNPs are intronic, 5'upstream, 3'downstream or intergenic. SNP rs2073618 is a missense variant modifying the third amino acid of OPG protein (encoded by *TNFRSF11B* gene) from Asn to Lys (N3K). SNP rs2073618 was genome-widely associated with serum total osteocalcin and in LD ( $r^2=0.57$ ,  $D'=0.89$ ) with the top SNP at chr8 (rs10955908). However, the functional impact of this missense SNP is predicted to be "benign" based on the ACMG Guideline.<sup>47</sup>

To annotate potential functions of these non-coding variants and to predict potential functions of these non-coding variants, we overlapped locations of these SNPs with locations of gene regulatory elements, especially on promoters, promoter flanking regions with transcription factor (TF) bindings, enhancers, CTCF binding

sites and TF binding sites (TFBS). The CTCF binding sites represent the activity of insulators, sequences that block the interaction between enhancers and promoters. The average length is 4.3Kb, 1.8Kb, 0.5kb, 0.6kb and 6kb, respectively for promoters, proximal enhancers, distal enhancers, CTCF bindings, and TFBS in human genome. 48 We overlapped SNP locations with these gene regulatory elements observed from human primary osteoblasts. The lead SNPs on chr1 *SMG5-BGLAP*, chr8 -*TNFRSF11B* and chr11 *FLI1* loci are in LD with SNPs located in gene regulatory elements (**Table 2**). rs1346 is located at the promoter flanking TF binding regions of the *SSSCA1-AS1* gene. Moreover, rs2104404, rs10955908, rs611307 and rs11840862 are in LD with SNPs located in the promoter flanking TB bonding region of their corresponding genes. With the exception of rs1246, these regulatory elements were only observed in cells from other lineages but not osteoblasts. (The “Multiple cells” column in **Table 2**). The majority of the associated SNPs were located in the enhancer regions of the mapped genes.

To estimate the confidence of a non-coding variant to be functional, we utilized the RegulomeDB score. It scored variants from 1 (high evidence) to 6 (no or low evidence) based on evidence from eQTL, TF binding, matched TF motif, matched DNase Footprint and DNase peak TF (Table 2). We identified one SNPs (rs6469789) on chr8 *TNFRSF11B* locus scored as “1d” (evidences from eQTL + TF binding + any motif + DNase peak). In addition, 7 SNPs on chr1 *SMG5-BGLAP* locus scored as “1f” (evidences from eQTL + TF binding + matched TF motif). Among them, only rs2246476 was a lead SNP. The SNPs scored as “1d” or “1f” are likely to affect binding and linked to expression of a gene target.

### Gene-based association Analyses

The gene-based association test implemented in MAGMA 49 using the FUMA platform was performed using summary statistics from discovery GWAS meta-analyses. All available SNPs were mapped to 18,793 protein coding genes. Genome-wide significance was found for *TNFRSF11B*, *SMG5*, *CCT3*, *TMRM79*, *PMF1-BGLAP* and *PAQR6* genes with p-values  $\leq 2.6 \times 10^{-6}$  ( $=0.05/18793$ ) after Bonferroni correction (**Supplementary Figure 5**). *TSC2* and *FPR3* were also strongly associated.

### Chromatin 3D Interactions

Hi-C looping structure (chromatin 3D interactions) was used to find interactions between SNPs located regions and potential promoter regions of protein coding genes. We only considered looping structures that overlapped with estimated enhancers and promoters, so that chromatin interactions are always between enhancer regions and promoter regions. We combined chromatin interactions as well as gene regulatory element information (enhancers and promoters) from multiple tissues together; thus,

Table 2 | Annotation of lead SNPs and SNPs in LD in osteoblast and other type of cells.

CHR	Lead SNP	SNP in LD	(r <sup>2</sup> )	Position	RegulomeDB	Feature Type	Osteoblasts	Other cell tissues	annotation	Genes
1	rs2246476		1.00	156271910	<b>1f</b>	Enhancer		x	intronic	SMG5
1		rs2253358	0.81	156258269	<b>1f</b>	Enhancer		x	intronic	SMG5
1		rs2985711	0.80	156312433	7	CTCF Binding Site	x	x	intronic	CCT3
1		rs2495269	0.80	156312858	6	CTCF Binding Site	x	x	intronic	CCT3
1	rs2277872		0.98	156276937	6	Promoter flanking with TF		x	intronic	SMG5
1	rs2104404	rs2104404	0.98	156287122	5	CTCF Binding Site		x	intronic	TMEM79
1		rs1543294	0.84	156240066	<b>1f</b>	Enhancer	x	x	3'-UTR	PMF1-BGLAP
1		rs12090516	0.84	156240766	4	Enhancer		x	intronic	RP11-54H19.8
1		rs2287025	0.84	156260350	7	Enhancer		x	intronic	SMG5
1		rs2287024	0.84	156260415	<b>1f</b>	Enhancer		x	intronic	SMG5
1		rs759328	0.84	156261475	<b>1f</b>	Enhancer		x	intronic	SMG5
1		rs4661036	0.88	156263688	6	Enhancer	x	x	intronic	SMG5
1		rs35478936	0.85	156335371	<b>1f</b>	Enhancer		x	intronic	CCT3
1		rs4661037	0.85	156268688	<b>1f</b>	Enhancer	x	x	intronic	SMG5
1		rs16837272	0.81	156271044	7	Enhancer		x	intronic	SMG5
1		rs4661038	1.00	156288712	5	CTCF Binding Site	x	x	intronic	TMEM79
8	rs10955908			118892318	5			x	intronic	31kb 3' of TNFRSF11B
8		rs6469788	0.80	118940511	7	Promoter flanking with TF		x	intronic	TNFRSF11B
8		rs6469789	0.81	118948422	<b>1d</b>	Promoter flanking with TF		x	intronic	TNFRSF11B
8		rs4407910	1.00	118904878	7	Enhancer	x	x	intronic	19kb 3' of TNFRSF11B
8		rs13439134	0.99	118906629	5	Enhancer		x	intronic	17kb 3' of TNFRSF11B
8		rs13277230	0.99	118910754	5	Enhancer	x	x	intronic	13kb 3' of TNFRSF11B
8		rs4355801	0.99	118911634	5	Enhancer	x	x	intronic	12kb 3' of TNFRSF11B
8		rs4319131	0.80	118935412	5	Enhancer	x	x	intronic	TNFRSF11B
8		rs7463176	0.81	118945386	7	Enhancer		x	intronic	TNFRSF11B
8		rs11573829	0.81	118947384	5	Enhancer	x	x	intronic	TNFRSF11B

Table 2 | Annotation of lead SNPs and SNPs in LD in osteoblast and other type of cells. (continued)

CHR	Lead SNP	SNP in LD	(r <sup>2</sup> )	Position	RegulomeDB	Feature Type	Osteoblasts	Other cell tissues	annotation	Genes
8		rs4242592	0.80	118956736	4	Enhancer		x		4.5kb 5' of TNFRSF11B
8		rs7006553	0.80	118958540	7	Enhancer		x		6.3kb 5' of TNFRSF11B
8		rs2062375	0.81	118965553	6	Enhancer		x		11kb 5' of U6
8		rs10101385	0.99	118911611	6	CTCF Binding Site		x		12kb 3' of TNFRSF11B
8		rs13250753	1.00	118894453	7	CTCF Binding Site		x		29kb 3' of TNFRSF11B
8		rs6415470	0.81	118942872	7	CTCF Binding Site		x	intronic	TNFRSF11B
8		rs10505348	0.80	118960457	7	CTCF Binding Site		x		8.3kb 5' of TNFRSF11B
11	rs1346	rs1346		65569780	4	Promoter flanking with TF	x	x		SSSCA1-AS1
11	rs611307	rs611307		128799925	5	Enhancer		x	intronic	FLI1
11		rs1787971	0.97	128813668	7	Enhancer		x		401bp 3' of FLI1
11		rs608764	0.95	128790125	5	Promoter flanking with TF	x	x	intronic	FLI1
11		rs549809	1.00	128815422	4	Promoter flanking with TF	e	x		2.2kb 3' of FLI1
11		rs550621	0.80	128815467	<b>3a</b>	Promoter flanking with TF	e	x	intronic	2.2kb 3' of FLI1
11		rs680952	0.96	128794013	6	Enhancer		x	intronic	FLI1
11		rs648193	0.96	128797907	5	Enhancer		x	intronic	FLI1
11		rs611307	0.97	128799925	5	Enhancer		x	intronic	FLI1
11		rs579355	0.97	128806265	7	Enhancer		x	intronic	FLI1
11		rs582825	1.00	128809518	4	Enhancer		x	intronic	FLI1
11		rs619255	1.00	128809555	5	Enhancer		x	intronic	FLI1
11		rs682695	1.00	128812565	<b>3a</b>	Enhancer		x	3'-UTR	FLI1
11		rs522331	1.00	128814724	6	Enhancer		x		1.5kb 3' of FLI1
11		rs681296	1.00	128816368	5	Enhancer	x	x		3.1kb 3' of FLI1
11		rs496501	1.00	128817050	5	Enhancer		x		3.8kb 3' of FLI1
11		rs500081	0.80	128817410	7	CTCF Binding Site		x		4.1kb 3' of FLI1
11		rs654788	1.00	128817755	5	CTCF Binding Site		x		4.5kb 3' of FLI1

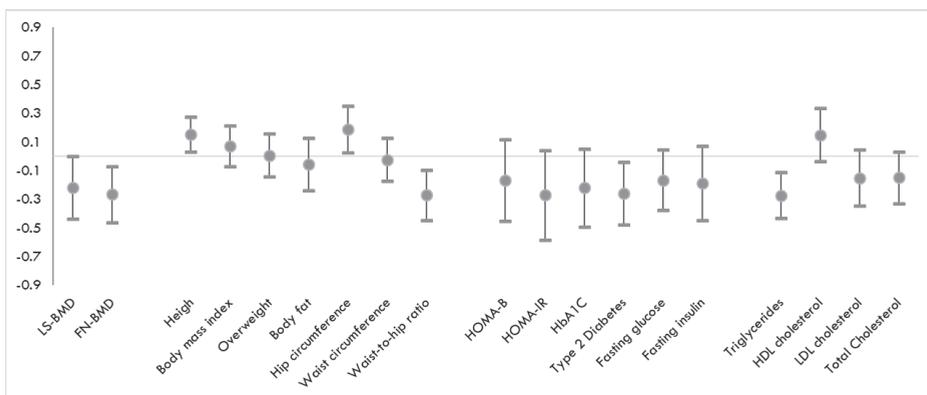
Table 2 | Annotation of lead SNPs and SNPs in LD in osteoblast and other type of cells. (continued)

CHR	Lead SNP	SNP in LD	(r <sup>2</sup> )	Position	RegulomeDB	Feature Type	Osteoblasts	Other cell tissues	annotation	Genes
11		rs654344	1.00	128817851	5	CTCF Binding Site		x		4.6kb 3' of FLI1
11		rs617954	1.00	128817901	7	CTCF Binding Site		x		4.6kb 3' of FLI1
11		rs504760	0.80	128817940	6	CTCF Binding Site		x		4.7kb 3' of FLI1
12	rs11840862	rs11840862		42382327	5	Enhancer		x		59kb 3' of AKAP11
13		rs11840862	0.40	42378009	4	Enhancer	x	x		54kb 3' of AKAP11
13		rs9533090	0.40	42377313	4	Promoter flanking with TF		x		54kb 3' of AKAP11
16	rs13337626	rs13337626	1.00	2075833	5	Enhancer			synonymous	TSC2
16		rs1800715	0.95	2075936	5	Enhancer		x	intronic	TSC2
19	rs8111274	rs8111274		51801340	5	Enhancer			intronic	FPR3

results are not tissue-type or cell-type specific. As showed on the circus plot (**Supplementary Figure 6A**) of chr1 *SMG5-BGLAP* locus, the chromatin interactions (colored orange) were found between top associated SNPs and several genes in the regions. eQTLs between top associated SNPs and genes were highlighted in green. Genes highlighted in red, including *SMG5*, *BGLAP* among others, showed both evidences from chromatin interactions and eQTLs. On chr 8 *TNFRSF11B*, *COLE10*, *MAL2*, *NOV* and *SMAD12* were all marked by significant chromatin interactions and eQTLs (**Supplementary Figure 6B**). Next, on chr11q24.3 both *KCNJ1* and *FLJ1* can be considered possible candidate genes, whereas on chr11q13.1 the list of genes marked by both chromatin interaction and eQTLs is large (**Supplementary Figure 6c** and **Supplementary Figure 6D**). The chr 13 locus was highlighted by *AKAP11*(**Supplementary Figure 6E**). Finally, on chr16 and chr19 (**Supplementary Figure 6F** and **Supplementary Figure 6g**, respectively), *TSC2* and *FPPR3* are likely to be the candidate genes with support from chromatin interactions and eQTLs.

## Heritability and Genetic Correlations

The estimated SNP heritability ( $h^2$ ) of serum total osteocalcin is 9% (95%CI: 4-12%). We found negative genetic correlations between serum total osteocalcin and triglycerides, waist-to-hip ratio, type 2 diabetes and femoral neck BMD ( $p < 0.05$ ). Positive genetic correlation was also found for height (genetic correlation=0.15) (**Figure 1**). Since BMI was adjusted in our osteocalcin GWAS analyses, the genetic correlations with quantitative BMI relevant phenotypes may not be estimated accurately.



**Figure 1 |** Genetic correlation between osteocalcin and variety of related traits. LS-BMD=Lumbar spine bone mineral density; FN-BMD= femoral neck bone mineral density. HOMA-B= Homeostatic model assessment of  $\beta$ -cell function index; HOMA-IR= Homeostatic model assessment of insulin resistance index; HDL= high-density lipoproteins; LDL= low-density lipoproteins.

## Mendelian Randomization Analyses

SNPs from seven replicated GWAS loci were selected as instrumental variables for total serum osteocalcin. We did not detect any pleiotropic biases in our MR analysis (Egger-intercept). We did not find any causal effect of genetically determined osteocalcin on type 2 diabetes and other glycemic traits with the exception of HOMA-IR (Table 3). We did not look into obesity relevant phenotypes (BMI, waist-to-hip ration, obesity) since BMI was adjusted in our osteocalcin GWAS analyses; therefore, the analyses may not be accurately.

As a positive control, we also performed Mendelian randomization analyses on bone relevant phenotypes. We found a causal effect of serum total osteocalcin on lower femoral-neck BMD (beta=-1.32; p-value=1.4x10<sup>-6</sup>) (Table 4).

**Table 3 |** Causal effect of serum total osteocalcin on different bone and metabolic traits.

Outcome	IVW		Weighted median		MR-PRESSO		Egger intercept	
	effect (95%CI)	P	Effect (95%CI)	P	global P	Effect (95%CI)	P	P
Femoral neck BMD	-0.685 (-1.239, -0.130)	0.016	-0.707 (-1.111, -0.304)	0.001	<0.001	-0.715 (-1.173, -0.257)	0.028	0.51
Lumbar spine BMD	-1.056 (-1.948, -0.163)	0.020	-0.590 (-1.106, -0.073)	0.026	<0.001	-0.616 (-1.087, -0.145)	0.083	0.94
Type 2 Diabetes	-0.395 (-1.046, 0.256)	0.234	-0.216 (-0.651, 0.234)	0.331	0.001	-0.370 (-0.729, 0.010)	0.100	0.19
Fasting glucose	-0.048 (-0.140, 0.044)	0.303	-0.017 (-0.136, 0.102)	0.781	0.583	NA	NA	0.67
Fasting insulin	-0.042 (-0.136, 0.052)	0.384	-0.062 (-0.190, 0.065)	0.775	0.421	NA	NA	0.51
Homa-B	-0.083 (-0.203, 0.038)	0.180	-0.181 (-0.232, 0.040)	0.012	0.159	NA	NA	0.02
Homa-IR	-0.128 (-0.250, -0.006)	0.041	-0.183 (-0.341, -0.026)	0.022	0.818	NA	NA	0.22

global p= test for horizontal pleiotropy; MR-Presso beta (95%CI) = corrected effect estimate for horizontal pleiotropy via outlier removal; NA= not applicable. If the global test is not significant (P<0.05) corrected effect is not calculated as there are no outliers

**Table 4** | Causal effect of BMD on osteocalcin.

Exposure	IVW		Weighted median		MR-PRESSO	Egger intercept		
	beta (95%CI)	P	beta (95%CI)	P	global P	beta (95%CI)	P	
Femoral neck BMD	-0.065 (-0.111, -0.019)	0.005	-0.035 (-0.069, -0.000)	0.048	<0.001	-0.042 (-0.069, -0.015)	0.004	0.85
Lumbar spine BMD	-0.056 (-0.097, -0.015)	0.008	-0.028 (-0.062, 0.005)	0.09	<0.001	-0.018 (-0.041, -0.005)	0.14	0.01

global p= test for horizontal pleiotropy; MR-Presso beta (95%CI) = corrected effect estimates for horizontal pleiotropy via outlier removal; NA= not applicable. If the global test is not significant (P<0.05) corrected effect is not calculated as there are no outlier

## DISCUSSION

In the current GWAS study, we successfully replicated seven GWAS loci associated with total serum osteocalcin. Several of the SNPs were located within or near genes implicated in relevant bone pathways such as *TNFRSF11B* and *AKAP11*. Total serum osteocalcin had modest heritability (~9%) and weak evidence of shared genetic etiology with femoral neck BMD, type 2 diabetes, triglycerides, waist-to-hip ratio and height. Nonetheless, none of these relationships were causal with the exception of BMD.

The strongest associations arose from the 1q22 locus which harbors the well-known osteocalcin gene, *BGLAP*, but also several other genes such as *SMG5*, *TMEM79* and *CCT3*. The strongest association (rs2246476) mapped to the *SMG5* gene. This SNP or its proxies ( $r^2 > 0.8$ ) have not been reported to be associated with any other traits before. Although *SMG5* have been associated with variety of traits such as mortality, HbA1c, red cell distribution, and adiposity the lead SNPs in these associations were not in LD with our lead SNP ( $r^2 < 0.2$ ). This SNP is highly likely to affect the binding and expression of *SMG5* (regulomeDB score of 1f [eQTL + TF binding/DNase peak]). In addition, the variant is located within the enhancer region of *SMG5* but only in the liver. Additional secondary signal in this locus (rs2277872) overlap with the TF Promoter flanking region of *SMG5* in blood cells and the enhancer region in osteoblasts. The biological function of *SMG5* is not well understood. What is known so far is that it plays important role in nonsense-mediated mRNA decay; process relevant to reduce errors during gene expression.<sup>50</sup> The gene is expressed across various tissues such as testis, blood, and muscle among others. In the proximity of *SMG5* is the *BGLAP* or the so called "osteocalcin gene". *BGLAP* encodes 4 exons from which the first three produce immature proteins, whereas exon 4 produces mature one.<sup>51</sup> In osteoblasts,

the expression of osteocalcin is incited by the Runx2 (Cbfa1) transcription factor; without it the osteoblasts will don't differentiate and produce bone<sup>52</sup> and osteocalcin will not be produced as well. On the other hand, BGLAP expression is inhibited by the transforming growth factor- $\beta$  (TGF- $\beta$ ) in an autocrine fashion. The *BGLAP* promotor is sensitive to vitamin D, alcohol intake and corticosteroids, thus, they may influence the levels of total serum osteocalcin. The lead SNPs in this locus are also in LD with SNPs mapping near *TMEM79* and *CCT3*; thus, any of these genes can be a potential candidate gene. Nevertheless, *SMG5* and *BGLAP* were highlighted by both chromatin interactions and eQTLs; making them stronger candidate genes in this locus.

Several SNPs mapped to genes involved in the nuclear factor- $\kappa$ B ligand (RANKL/RANK /osteoprotegerin (OPG)) signaling pathway such as *TNFRSF11B* and *AKAP11*. *TNFRSF11B* encodes the OPG protein which is primarily secreted from the osteoblasts and it is essential for osteoclast development and function. OPG binds to RANKL, and prevents it from activating its cognate receptor RANK, which is vital for osteoclast differentiation, activation and survival.<sup>53</sup> Next, the *AKAP11* gene (is located ~185 kb upstream of RANKL) encodes a group of A-kinase anchor proteins (AKAP), involved in intracellular signaling events by linking cAMP to other pathways.<sup>54,55</sup> In mice, AKAP11 is continuously expressed during osteoblast maturation.<sup>56</sup> Akap11 knockout cells, generated using CRISPR/Cas9 in mouse pre-osteoblastic MC3T3-E1 cells, presented significant reduction of calcium mineral deposition and decreased alkaline phosphatase (ALP) activity since differentiation phase (Day7). Moreover, the expression of the major structural protein genes such as OC, IBSP and SPP1, was reduced since mineralization phase (Day16).<sup>56</sup> *TNFRSF11B* and *AKAP11* were mapped by both chromatin interaction and eQTLs. Signals also arose from the 11q24.3 locus that have been recently associated with heel BMD<sup>57</sup> which mapped near *FLI1* that have also been linked to the NF- $\kappa$ B Signaling; however, the exact mechanism of action of this gene on bone are still unknown. Overall, these genes play a key role in bone metabolism, thus, can affect the serum osteocalcin levels as well.

Next, we identified two loci that have not been associated with any bone traits harboring the *TSC2* and *FPR3* genes. Nevertheless, these genes have shown links to relevant bone pathways. For instance, *TSC2* been related with the mTOR signaling pathway<sup>58</sup> implicated in osteoblast differentiation and function stimulated by several bone anabolic ligands such as Wnt, igf and bmp.<sup>59,60,61</sup> The role of mTOR signaling in osteoclastogenesis is still controversial as inactivation of mTORC1 by deletion of Raptor can enhance, whereas hyperactivation of mTORC1 by deleting Tsc1 in osteoclast precursors can impair osteoclastogenesis.<sup>62</sup> *FPR3* on the other hand, has been proposed to promote osteoblast differentiation via the N-Formyl Peptide Receptor 1-mediated Signaling Pathway in Human Mesenchymal Stem Cells from Bone Marrow GPCRs (G protein coupled receptor) in bone.<sup>63</sup> G protein signaling also has a key role in

bone development and remodeling by influencing the GPCR-G, Wnt, calcium and PTH signaling pathways. In addition, the G-protein-coupled receptor (GPCR) 48 have been related with significant downregulation of the expression levels of osteocalcin and collagen.<sup>64</sup> Interestingly, GPCR represents the largest family of proteins that is currently targeted by approved drugs.<sup>65</sup> The exact role of these genes in the underlying pathways is still unknown and warrants further investigation.

In the past decade there has been increasing recognition of the intimate relationship between bone and energy metabolism. Initially it was shown that energy metabolism influence bone remodeling through the effects of leptin, serotonin and the sympathetic tone [reviewed elsewhere<sup>66</sup>]. Therefore, it was postulated that bone may also provide a feedback loop and affect energy metabolism. An early experimental animal study showed that osteocalcin-deficient mice were glucose intolerant, insulin resistant and fat started.<sup>5</sup> Similarly, observational studies have shown inverse association between osteocalcin, adiposity and diabetes.<sup>67,68,69,70</sup> We did find a negative genetic correlation between osteocalcin and type 2 diabetes. However, our MR analysis, did not show any causal effect of genetically determined total serum osteocalcin on type 2 diabetes; indicating that osteocalcin and type 2 diabetes may share common pathways and the observed associations to be simple due to pleiotropy. In addition, osteocalcin and BMD showed a bidirectional association driven by the SNPs mapping to the RANK/RANK/OPG pathway; establishing the role of osteocalcin as bone turnover marker where as evidence of a causal effect cannot be established. Overall, the MR results need to be interpreted with caution as the variance explained of osteocalcin was low; thus, providing low power to detect any true causal associations.

This study has several limitations that need to be highlighted. First, the majority of the studies were imputed to older imputation panels (HapMap) and only one utilized newer and denser panel (HRC1.1); thus, we might have missed some variants associated with total serum osteocalcin not present in the majority of the studies. Second, the osteocalcin assays were different between studies which may have led to lower power to detect true-positive associations. Finally, we were not able to examine the carboxylated and non-carboxylated forms of osteocalcin separately as the majority of the cohorts estimated only the total serum osteocalcin.

Total serum osteocalcin is a complex, heritable trait implicated in relevant bone mechanisms. We identified seven loci robustly associated with total serum osteocalcin. MR analysis showed no causal link with metabolic and glycemic traits, but did provide evidence for bidirectional association with BMD, highlighting the essential role of osteocalcin in bone biology. The exact role of osteocalcin in energy metabolism in humans is yet to be fully elucidated till then our results do not support any causal association.

## REFERENCES

1. Hauschka P V, Lian JB, Cole DE, Gundberg CM. Osteocalcin and matrix Gla protein: Vitamin K-dependent proteins in bone. *Physiol Rev.* 1989;69(3):990-1047. doi:10.1152/physrev.1989.69.3.990
2. CALVO MS, EYRE DR, GUNDBERG CM. Molecular Basis and Clinical Application of Biological Markers of Bone Turnover\*. *Endocr Rev.* 1996;17(4):333-368. doi:10.1210/edrv-17-4-333
3. Lian JB, Gundberg CM. Osteocalcin. Biochemical considerations and clinical applications. *Clin Orthop Relat Res.* 1988;(226):267-291. doi:10.1097/00003086-198801000-00036
4. Gundberg CM, Weinstein RS. Multiple immunoreactive forms of osteocalcin in uremic serum. *J Clin Invest.* 1986;77(6):1762-1767. doi:10.1172/JCI112499
5. Moser SC, van der Eerden BCJ. Osteocalcin—A Versatile Bone-Derived Hormone. *Front Endocrinol (Lausanne).* 2019;9:794. doi:10.3389/fendo.2018.00794
6. Oury F, Sumara G, Sumara O, et al. Endocrine Regulation of Male Fertility by the Skeleton. *Cell.* 2011;144(5):796-809. doi:10.1016/j.cell.2011.02.004
7. Mera P, Laue K, Ferron M, et al. Osteocalcin Signaling in Myofibers Is Necessary and Sufficient for Optimum Adaptation to Exercise. *Cell Metab.* 2016;23(6):1078-1092. doi:10.1016/j.cmet.2016.05.004
8. Oury F, Khrimian L, Denny CA, et al. XMaternal and offspring pools of osteocalcin influence brain development and functions. *Cell.* 2013;155(1):228. doi:10.1016/j.cell.2013.08.042
9. Khrimian L, Obri A, Ramos-Brossier M, et al. Gpr158 mediates osteocalcin's regulation of cognition. *J Exp Med.* 2017;214(10):2859-2873. doi:10.1084/jem.20171320
10. Hinoi E, Gao N, Jung DY, et al. The sympathetic tone mediates leptin's inhibition of insulin secretion by modulating osteocalcin bioactivity. *J Cell Biol.* 2008;183(7):1235-1242. doi:10.1083/jcb.200809113
11. Hinoi E, Gao N, Jung DY, et al. An osteoblast-dependent mechanism contributes to the leptin regulation of insulin secretion. *Ann N Y Acad Sci.* 2009;1173(SUPPL. 1):E20-30. doi:10.1111/j.1749-6632.2009.05061.x
12. Ferron M, Hinoi E, Karsenty G, Ducy P. Osteocalcin differentially regulates  $\beta$  cell and adipocyte gene expression and affects the development of metabolic diseases in wild-type mice. *Proc Natl Acad Sci.* 2008;105(13):5266-5270. doi:10.1073/pnas.0711119105
13. Díaz-López A, Bulló M, Juanola-Falgarona M, et al. Reduced serum concentrations of carboxylated and undercarboxylated osteocalcin are associated with risk of developing type 2 diabetes mellitus in a high cardiovascular risk population: A nested case-control study. *J Clin Endocrinol Metab.* 2013;98(11):4524-4531. doi:10.1210/jc.2013-2472
14. Kindblom JM, Ohlsson C, Ljunggren Ö, et al. Plasma Osteocalcin Is Inversely Related to Fat Mass and Plasma Glucose in Elderly Swedish Men. *J Bone Miner Res.* 2009;24(5):785-791. doi:10.1359/jbmr.081234
15. Liatis S, Sfikakis PP, Tsiakou A, et al. Baseline osteocalcin levels and incident diabetes in a 3-year prospective study of high-risk individuals. *Diabetes Metab.* 2014;40(3):198-203. doi:10.1016/j.diabet.2014.01.001
16. Hong S, Koo J, Hwang JK, et al. Changes in Serum Osteocalcin are Not Associated with Changes in Glucose or Insulin for Osteoporotic Patients Treated with Bisphosphonate. *J Bone Metab.* 2013;20(1):37. doi:10.11005/jbm.2013.20.1.37
17. Storm D, Porter REES, Musgrave K, et al. Calcium Supplementation Prevents Seasonal Bone Loss and Changes in Biochemical Markers of Bone Turnover in Elderly New England Wom-

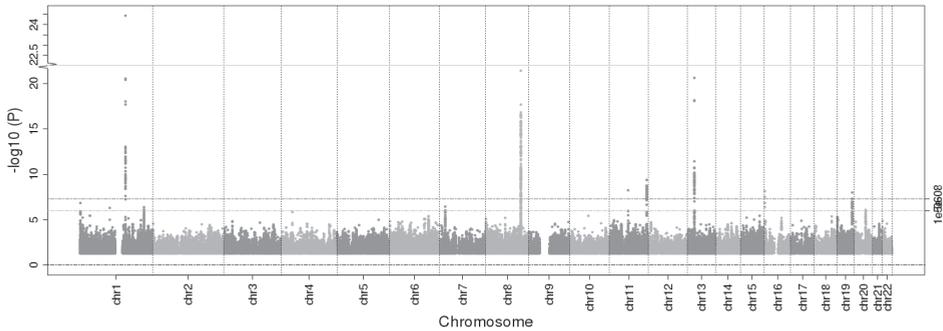
- en: A Randomized Placebo-Controlled Trial 1. *J Clin Endocrinol Metab.* 1998;83(11):3817-3825. doi:10.1210/jcem.83.11.5289
18. Woitge HW, Knothe A, Witte K, et al. Circannual rhythms and interactions of vitamin D metabolites, parathyroid hormone, and biochemical markers of skeletal homeostasis: A prospective study. *J Bone Miner Res.* 2000;15(12):2443-2450. doi:10.1359/jbmr.2000.15.12.2443
  19. Vanderschueren D, Gevers G, Raymaekers G, Devos P, Dequeker J. Sex- and age-related changes in bone and serum osteocalcin. *Calcif Tissue Int.* 1990;46(3):179-182. doi:10.1007/bf02555041
  20. Johansen JS, Giwercman A, Hartwell D, et al. Serum Bone Gla-Protein as a Marker of Bone Growth in Children and Adolescents: Correlation With Age, Height, Serum Insulin-Like Growth Factor I, and Serum Testosterone. *J Clin Endocrinol Metab.* 1988;67(2):273-278. doi:10.1210/jcem-67-2-273
  21. Yasui T, Uemura H, Tomita J, et al. Association of serum undercarboxylated osteocalcin with serum estradiol in pre-, peri- and early post-menopausal women. *J Endocrinol Invest.* 2006;29(10):913-918. doi:10.1007/BF03349196
  22. Mitchell BD, Cole SA, Bauer RL, et al. Genes Influencing Variation in Serum Osteocalcin Concentrations Are Linked to Markers on Chromosomes 16q and 20q 1. *J Clin Endocrinol Metab.* 2000;85(4):1362-1366. doi:10.1210/jcem.85.4.6571
  23. Dohi Y, Iki M, Ohgushi H, et al. A novel polymorphism in the promoter region for the human osteocalcin gene: The possibility of a correlation with bone mineral density in post-menopausal Japanese women. *J Bone Miner Res.* 1998;13(10):1633-1639. doi:10.1359/jbmr.1998.13.10.1633
  24. Gustavsson A, Nordström P, Lorentzon R, Lerner UH, Lorentzon M. Osteocalcin Gene Polymorphism is Related to Bone Density in Healthy Adolescent Females. *Osteoporos Int.* 2000;11(10):847-851. doi:10.1007/s001980070043
  25. Jiang DK, Xu FH, Liu MY, et al. No evidence of association of the osteocalcin gene HindIII polymorphism with bone mineral density in Chinese women. *J Musculoskelet Neuronal Interact.* 7(2):149-154. <http://www.ncbi.nlm.nih.gov/pubmed/17627084>. Accessed April 1, 2020.
  26. Marchini J, Howie B, Myers S, McVean G, Donnelly P. A new multipoint method for genome-wide association studies by imputation of genotypes. *Nat Genet.* 2007;39(7):906-913. doi:10.1038/ng2088
  27. Howie BN, Donnelly P, Marchini J. A Flexible and Accurate Genotype Imputation Method for the Next Generation of Genome-Wide Association Studies. Schork NJ, ed. *PLoS Genet.* 2009;5(6):e1000529. doi:10.1371/journal.pgen.1000529
  28. Li Y, Willer CJ, Ding J, Scheet P, Abecasis GR. MaCH: Using sequence and genotype data to estimate haplotypes and unobserved genotypes. *Genet Epidemiol.* 2010;34(8):816-834. doi:10.1002/gepi.20533
  29. Howie B, Fuchsberger C, Stephens M, Marchini J, Abecasis GR. Fast and accurate genotype imputation in genome-wide association studies through pre-phasing. *Nat Genet.* 2012;44(8):955-959. doi:10.1038/ng.2354
  30. Consortium the HR, McCarthy S, Das S, et al. A reference panel of 64,976 haplotypes for genotype imputation. *Nat Genet.* 2016;48(10):1279-1283. doi:10.1038/ng.3643

31. Pruim RJ, Welch RP, Sanna S, et al. LocusZoom: regional visualization of genome-wide association scan results. *Bioinformatics*. 2010;26(18):2336-2337. doi:10.1093/bioinformatics/btq419
32. Yang J, Lee SH, Goddard ME, Visscher PM. GCTA: a tool for genome-wide complex trait analysis. *Am J Hum Genet*. 2011;88(1):76-82. doi:10.1016/j.ajhg.2010.11.011
33. van Wijngaarden JP, Dhonukshe-Rutten RA, van Schoor NM, et al. Rationale and design of the B-PROOF study, a randomized controlled trial on the effect of supplemental intake of vitamin B12 and folic acid on fracture incidence. *BMC Geriatr*. 2011;11(1):80. doi:10.1186/1471-2318-11-80
34. Zheng J, Erzurumluoglu AM, Elsworth BL, et al. LD Hub: a centralized database and web interface to perform LD score regression that maximizes the potential of summary level GWAS data for SNP heritability and genetic correlation analysis. *Bioinformatics*. 2017;33(2):272-279. doi:10.1093/bioinformatics/btw613
35. Bulik-Sullivan BK, Loh P-R, Finucane HK, et al. LD Score regression distinguishes confounding from polygenicity in genome-wide association studies. *Nat Genet*. 2015;47(3):291-295. doi:10.1038/ng.3211
36. Burgess S, Bowden J, Fall T, Ingelsson E, Thompson SG. Sensitivity analyses for robust causal inference from mendelian randomization analyses with multiple genetic variants. *Epidemiology*. 2017;28(1):30-42. doi:10.1097/EDE.0000000000000559
37. Bowden J, Davey Smith G, Burgess S. Mendelian randomization with invalid instruments: effect estimation and bias detection through Egger regression. *Int J Epidemiol*. 2015;(June):512-525. doi:10.1093/ije/dyv080
38. Verbanck M, Chen CY, Neale B, Do R. Detection of widespread horizontal pleiotropy in causal relationships inferred from Mendelian randomization between complex traits and diseases. *Nat Genet*. 2018;50(5):693-698. doi:10.1038/s41588-018-0099-7
39. Hemani G, Tilling K, Davey Smith G. Orienting the causal relationship between imprecisely measured traits using GWAS summary data. Li J, ed. *PLOS Genet*. 2017;13(11):e1007081. doi:10.1371/journal.pgen.1007081
40. Watanabe K, Taskesen E, van Bochoven A, Posthuma D. Functional mapping and annotation of genetic associations with FUMA. *Nat Commun*. 2017;8(1):1826. doi:10.1038/s41467-017-01261-5
41. Adzhubei I, Jordan DM, Sunyaev SR. Predicting functional effect of human missense mutations using PolyPhen-2. *Curr Protoc Hum Genet*. 2013;0 7(SUPPL.76):Unit7.20. doi:10.1002/0471142905.hg0720s76
42. Neph S, Vierstra J, Stergachis AB, et al. An expansive human regulatory lexicon encoded in transcription factor footprints. *Nature*. 2012;489(7414):83-90. doi:10.1038/nature11212
43. Gerstein MB, Kundaje A, Hariharan M, et al. Architecture of the human regulatory network derived from ENCODE data. *Nature*. 2012;489(7414):91-100. doi:10.1038/nature11245
44. Roadmap Epigenomics Consortium, Kundaje A, Meuleman W, et al. Integrative analysis of 111 reference human epigenomes. *Nature*. 2015;518(7539):317-329. doi:10.1038/nature14248
45. Ward LD, Kellis M. HaploReg: a resource for exploring chromatin states, conservation, and regulatory motif alterations within sets of genetically linked variants. *Nucleic Acids Res*. 2012;40(Database issue):D930-4. doi:10.1093/nar/gkr917
46. Ernst J, Kellis M. ChromHMM: Automating chromatin-state discovery and characterization. *Nat Methods*. 2012;9(3):215-216. doi:10.1038/nmeth.1906

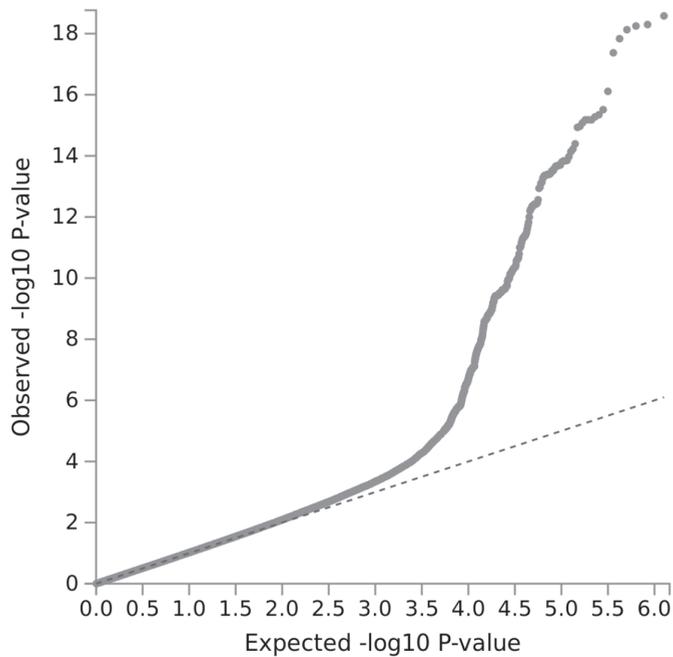
47. Richards S, Aziz N, Bale S, et al. Standards and guidelines for the interpretation of sequence variants: A joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. *Genet Med*. 2015;17(5):405-424. doi:10.1038/gim.2015.30
48. Zerbino DR, Wilder SP, Johnson N, Juettemann T, Flicek PR. The Ensembl Regulatory Build. *Genome Biol*. 2015;16(1):56. doi:10.1186/s13059-015-0621-5
49. de Leeuw CA, Mooij JM, Heskes T, Posthuma D. MAGMA: Generalized Gene-Set Analysis of GWAS Data. Tang H, ed. *PLOS Comput Biol*. 2015;11(4):e1004219. doi:10.1371/journal.pcbi.1004219
50. Nickless A, Bailis JM, You Z. Control of gene expression through the nonsense-mediated RNA decay pathway. *Cell Biosci*. 2017;7(1). doi:10.1186/s13578-017-0153-7
51. McKee MD, Cole WG. Bone Matrix and Mineralization. In: *Pediatric Bone*. Elsevier Inc.; 2012:9-37. doi:10.1016/B978-0-12-382040-2.10002-4
52. Komori T, Yagi H, Nomura S, et al. Targeted disruption of Cbfa1 results in a complete lack of bone formation owing to maturational arrest of osteoblasts. *Cell*. 1997;89(5):755-764. doi:10.1016/S0092-8674(00)80258-5
53. Leibbrandt A, Penninger JM. RANK–RANKL Signaling. In: *Encyclopedia of Cancer*. Springer Berlin Heidelberg; 2008:2537-2540. doi:10.1007/978-3-540-47648-1\_4945
54. Carnegie GK, Means CK, Scott JD. A-kinase anchoring proteins: From protein complexes to physiology and disease. *IUBMB Life*. 2009;61(4):394-406. doi:10.1002/iub.168
55. AKAP11 A-kinase anchoring protein 11 [Homo sapiens (human)] - Gene - NCBI. <https://www.ncbi.nlm.nih.gov/gene/11215>. Accessed April 1, 2020.
56. Cheng VKF, Li GHY, Tan KCB, Lu W, Cheung C. AKAP11 is a positive regulator of osteoblast extracellular matrix formation and mineralization. *46th Annu Meet Eur Calcif Tissue Soc (ECTS 2019)*.
57. Li Y, Corradetti MN, Inoki K, Guan KL. TSC2: Filling the GAP in the mTOR signaling pathway. *Trends Biochem Sci*. 2004;29(1):32-38. doi:10.1016/j.tibs.2003.11.007
58. Inoki K, Ouyang H, Zhu T, et al. TSC2 Integrates Wnt and Energy Signals via a Coordinated Phosphorylation by AMPK and GSK3 to Regulate Cell Growth. *Cell*. 2006;126(5):955-968. doi:10.1016/j.cell.2006.06.055
59. Xian L, Wu X, Pang L, et al. Matrix IGF-1 maintains bone mass by activation of mTOR in mesenchymal stem cells. *Nat Med*. 2012;18(7):1095-1101. doi:10.1038/nm.2793
60. Karner CM, Lee S-Y, Long F. Bmp Induces Osteoblast Differentiation through both Smad4 and mTORC1 Signaling. *Mol Cell Biol*. 2017;37(4). doi:10.1128/mcb.00253-16
61. Zhang Y, Xu S, Li K, et al. mTORC1 Inhibits NF- $\kappa$ B/NFATc1 Signaling and Prevents Osteoclast Precursor Differentiation, In Vitro and In Mice. *J Bone Miner Res*. 2017;32(9):1829-1840. doi:10.1002/jbmr.3172
62. Shin MK, Jang YH, Yoo HJ, et al. N-formyl-Methionyl-Leucyl-Phenylalanine (fMLP) promotes osteoblast differentiation via the N-formyl peptide receptor 1-mediated signaling pathway in human mesenchymal stem cells from bone marrow. *J Biol Chem*. 2011;286(19):17133-17143. doi:10.1074/jbc.M110.197772
63. Luo J, Zhou W, Zhou X, et al. Regulation of bone formation and remodeling by G-protein-coupled receptor 48. *Development*. 2009;136(16):2747-2756. doi:10.1242/dev.033571
64. Sriram K, Insel PA. G protein-coupled receptors as targets for approved drugs: How many targets and how many drugs? In: *Molecular Pharmacology*. Vol 93. American Society for Pharmacology and Experimental Therapy; 2018:251-258. doi:10.1124/mol.117.111062

65. Confavreux CB. Bone: from a reservoir of minerals to a regulator of energy metabolism. *Kidney Int Suppl.* 2011;79(121):S14-S19. doi:10.1038/ki.2011.25
66. Takashi Y, Ishizu M, Mori H, et al. Circulating osteocalcin as a bone-derived hormone is inversely correlated with body fat in patients with type 1 diabetes. Mogi M, ed. *PLoS One.* 2019;14(5):e0216416. doi:10.1371/journal.pone.0216416
67. Polgreen LE, Jacobs DR, Nathan BM, et al. Association of osteocalcin with obesity, insulin resistance, and cardiovascular risk factors in young adults. *Obesity (Silver Spring).* 2012;20(11):2194-2201. doi:10.1038/oby.2012.108
68. Kanazawa I, Yamaguchi T, Yamauchi M, et al. Serum undercarboxylated osteocalcin was inversely associated with plasma glucose level and fat mass in type 2 diabetes mellitus. *Osteoporos Int.* 2011;22(1):187-194. doi:10.1007/s00198-010-1184-7
69. Ngarmukos C, Chailurkit L, Chanprasertyothin S, Hengprasith B, Sritara P, Ongphiphadhanakul B. A reduced serum level of total osteocalcin in men predicts the development of diabetes in a long-term follow-up cohort. *Clin Endocrinol (Oxf).* 2012;77(1):42-46. doi:10.1111/j.1365-2265.2011.04215.x

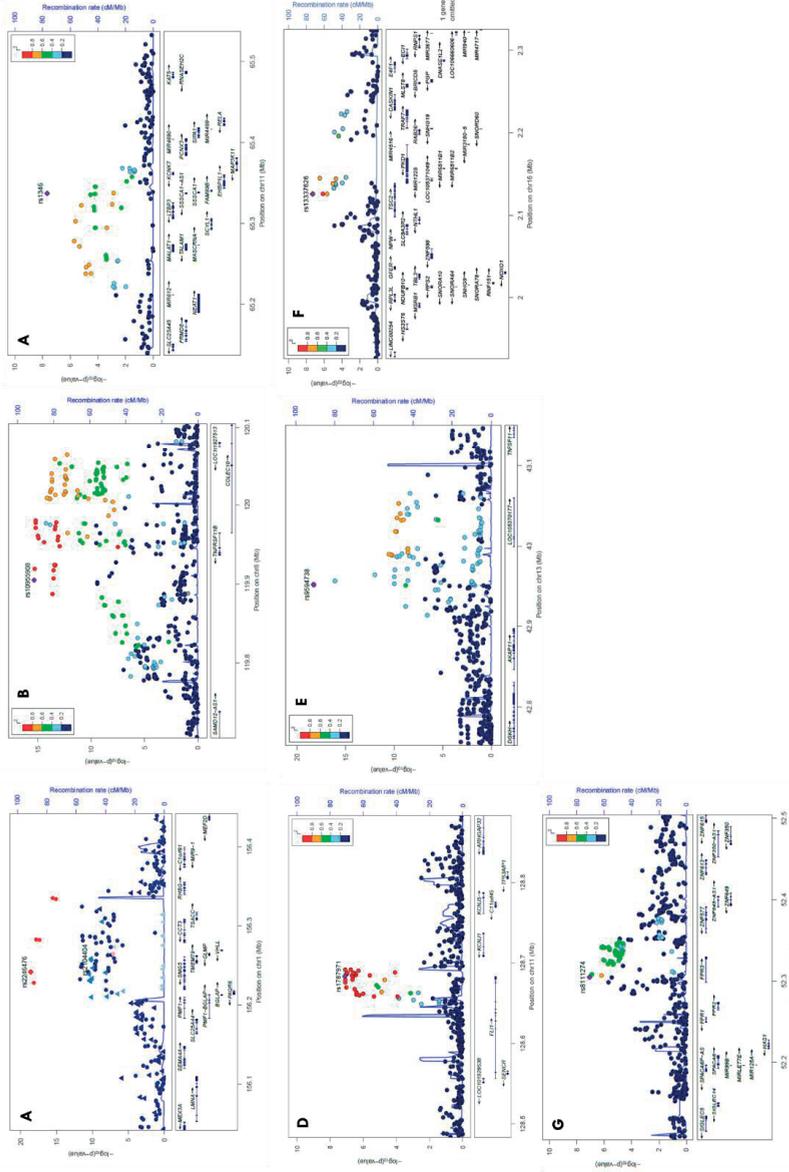
## SUPPLEMENTARY MATERIAL



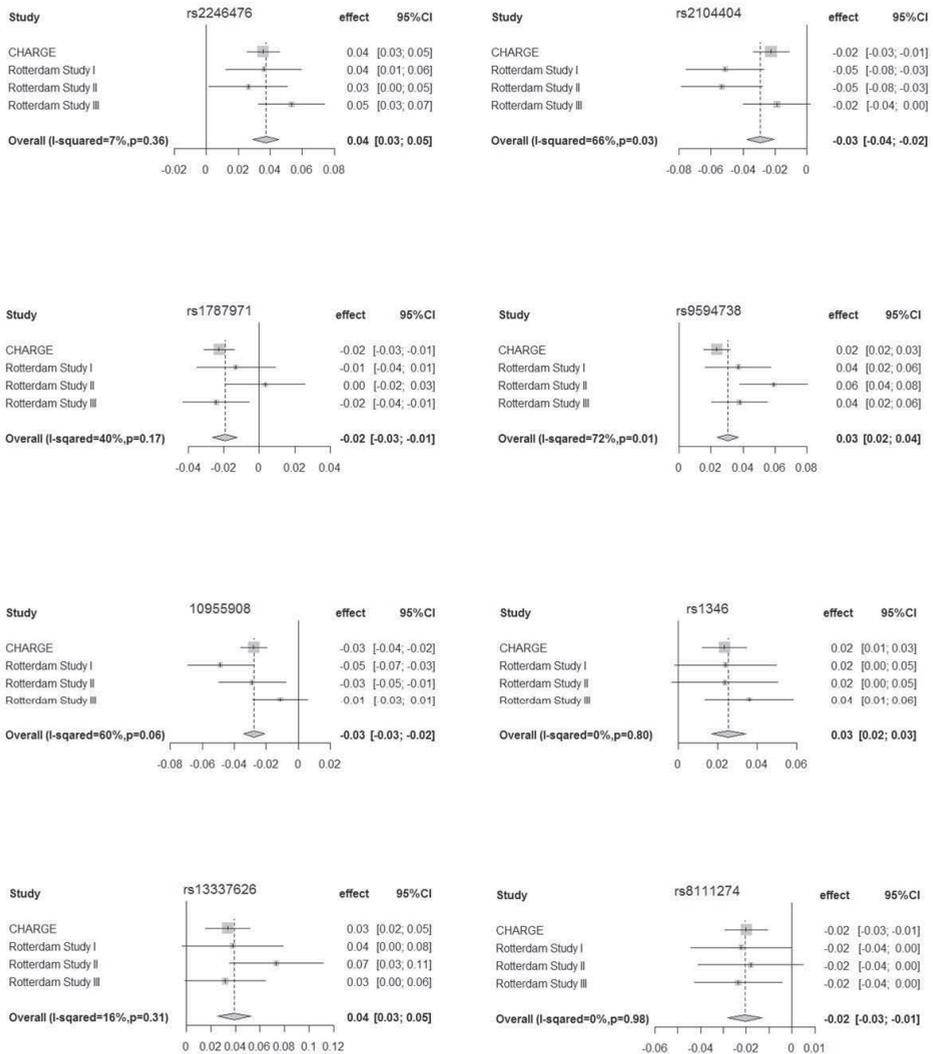
**Supplementary Figure 1** | Manhattan Plot of Association Statistics ( $-\log_{10}(P)$ ) for serum total osteocalcin. Each dot represents a SNP and the x axis indicates its chromosomal position (built 37 NCBI). Dashed red horizontal line marks the GWS threshold ( $P \leq 5 \times 10^{-08}$ ).



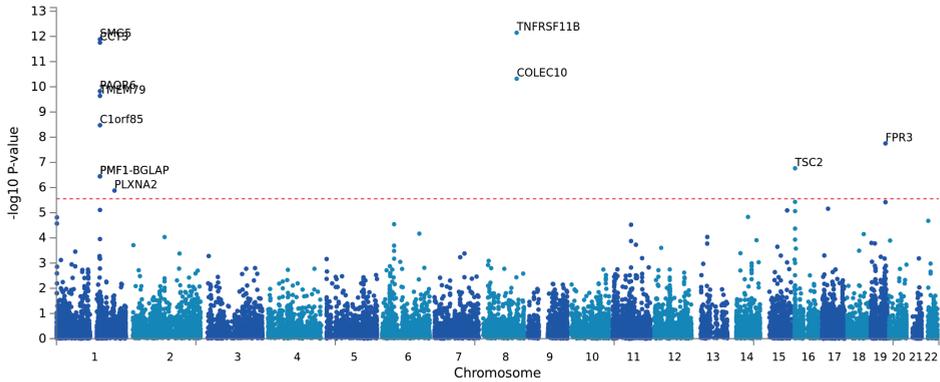
**Supplementary Figure 2** | Quantile-quantile (QQ) plot of observed versus expected P values of the GWAS results. The dashed red line in the QQ plot indicates the distribution of SNPs under the null hypothesis.



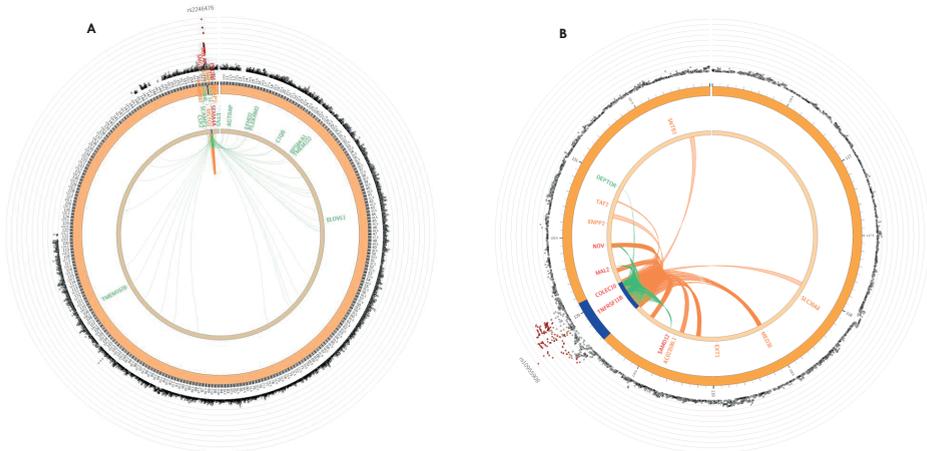
**Supplementary figure 3 |** Regional association plots displaying 1q22 (A), 8q24.12 (B), 11q13.1(C), 11q24.3 (D), 13q14.11 (E), 16p13.3(F) and 19q13.41 (G) loci. In each plot, the  $-\log_{10}$  of p values are on the left y-axis; the SNP genomic position (hg19) on the x-axis; the estimated recombination rate from 1000 genomes (March 2015 EUR) are on the right y-axis and plotted in blue. SNPs are coloured red to reflect linkage disequilibrium (LD) with the most significant SNP in purple (pairwise  $r^2$  from 1000 genomes March 2015 EUR).



**Supplementary figure 4** | Forest plots displaying the effect estimates of the GWS SNPs for each individual study and the pool estimated. The effect estimate represents 1-unit change in the log transformed serum total osteocalcin per copy of the modelled allele.

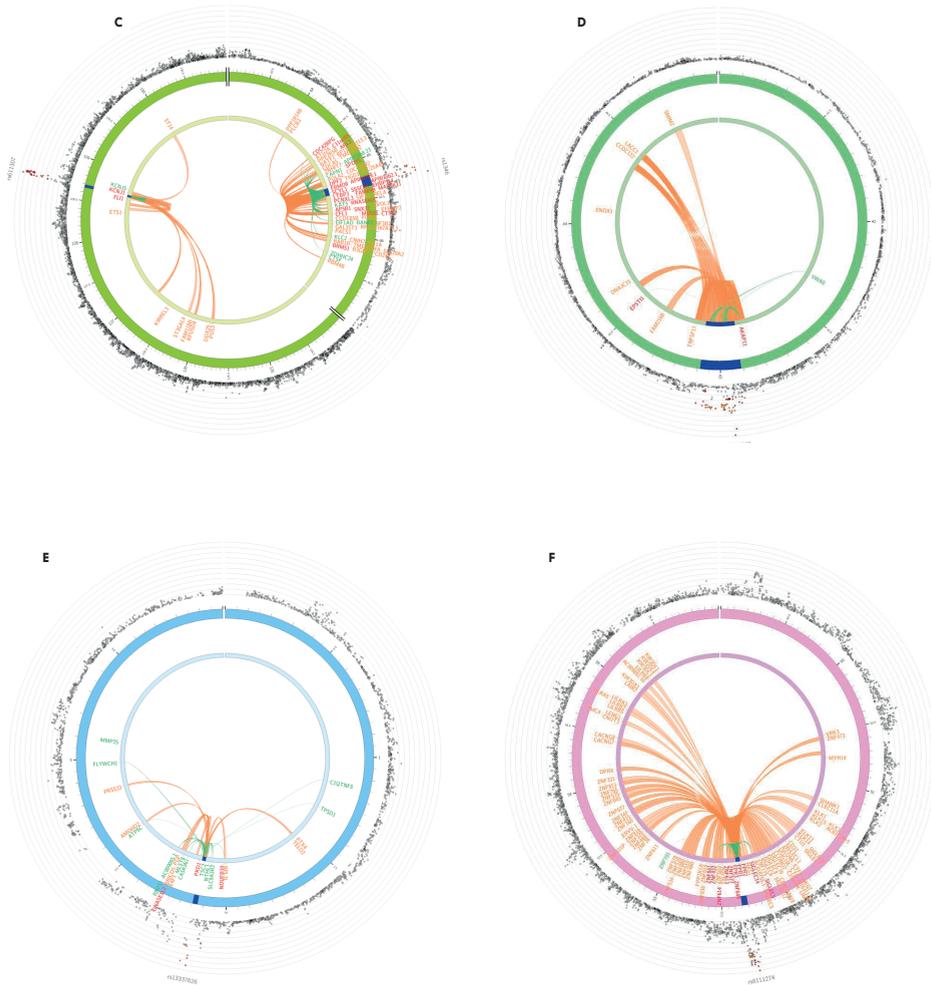


**Supplementary figure 5 | Gene-Based Association Analyses Using Summary Statistics from Discovery Meta-Analysis**



**Supplementary Figure 6 | The Circos Plot with Chromatin Intraction (Orange lines) and eQTLs (Green Lines) on displaying 1q22 (A), 8q24.12 (B), 11q13.1(C), 11q24.3 (D), 13q14.11 (E), 16p13.3(F) and 19q13.41 (G) loci.**

The most outer layer represents the Manhattan plot. SNPs in genomic risk loci are colour-coded as a function of their maximum  $r^2$  to the lead SNPs in the locus, as follows: red ( $r^2 > 0.8$ ), orange ( $r^2 > 0.6$ ), green ( $r^2 > 0.4$ ), blue ( $r^2 > 0.2$ ) and grey ( $r^2 \leq 0.2$ ). The rs ID of the top SNPs in the risk locus is displayed in the most outer layer. The second layer contains the genomic risk locus which is highlighted in blue. Here genes are mapped by chromatin interactions (**orange**) or eQTLs (**green**). When the gene is mapped by both, it is colored in red.



**Supplementary Figure 6** | The Circos Plot with Chromatin Intraction (Orange lines) and eQTLs (Green Lines) on displaying 1q22 (A), 8q24.12 (B), 11q13.1(C), 11q24.3 (D), 13q14.11 (E), 16p13.3(F) and 19q13.41 (G) loci.

The most outer layer represents the Manhattan plot. SNPs in genomic risk loci are colour-coded as a function of their maximum  $r^2$  to the lead SNPs in the locus, as follows: red ( $r^2 > 0.8$ ), orange ( $r^2 > 0.6$ ), green ( $r^2 > 0.4$ ), blue ( $r^2 > 0.2$ ) and grey ( $r^2 \leq 0.2$ ). The rs ID of the top SNPs in the risk locus is displayed in the most outer layer. The second layer contains the genomic risk locus which is highlighted in blue. Here genes are mapped by chromatin interactions (orange) or eQTLs (green). When the gene is mapped by both, it is colored in red. (continued)

**Supplementary Table S2 | Genotype, Imputation and Sample QC information for Participating Cohorts**

Study	Genotyping		SNP Inclusion criteria			
	Platform(s) / Chip(s)	Calling Algorithm	MAF	Call Rate	P-test HWE	# SNPs
Discovery Cohorts						
AGEs	Illumina Hu370CNV	BeadStudio	≥1%	≥97%	>10 <sup>-6</sup>	308340
FHS	Affymatrix 550K	BRLMM	≥1%	≥97%	>10 <sup>-6</sup>	378163
GOOD	Illumina Infinium HumanHap 610K	BeadStudio	≥1%	≥98%	>10 <sup>-6</sup>	521160
HealthABC	Illumina Infinium 1M	BeadStudio	≥1%	≥97%	>10 <sup>-6</sup>	African American = 1007669. European Ancestry 914075
Mr.OS Sweden	Illumina HumanOmni1_Quad_v1-0 B	GenomeStudio	≥1%	≥98%	>10 <sup>-4</sup> for autosomes. >10 <sup>-3</sup> for X chromosome	739477
Mr.OS US	Illumina HumanOmni1_Quad_v1-0 B	GenomeStudio	≥ 1%	≥ 97%	≥ 10 <sup>-4</sup>	725550
PIVUS	Human Omni Express	GenCall	>0	≥95%	>10 <sup>-6</sup>	645318
SHIP	Affymetrix 6.0	Birdseed2	NA	NA	NA	869224
SHIP-Trend	Illumina Omni 2.5	GenCall	>0	>0.9	>0.0001	1782967
TwinsUK	HumanHap300 HumanHap610Q. 1M-Duo and 1.2MDuo 1M).	Illuminus	MAF>1%	SNP call rate >97% (SNPs with MAF≥5%) or >99% (for 1% ≤MAF < 5%)	>10 <sup>-6</sup>	317k=303.940; 610k=554.234; 1M=878.319

SNP Imputation			Samples		Association	
Method	MAF	Quality metric	Total # SNPs	Call rate* selection Discovery Cohorts	Sample QC / Other exclusions	Software
MACH 1.0.16	>0%	$r^2\text{-hat} \geq 0.30$	2533153	$\geq 97\%$	1) Mismatch with previous genotypes from other experiments. 2) missing height and weight	PorbABEL 0.0.5c
MACH 1.0.16	$\geq 1\%$	$(O/E)\sigma^2$ ratio $\geq 0.30$	2471285	$\geq 97\%$	1) Autosomal heterozygosity ( $<0.26$ or $>0.30$ ); (2) Ethnic outliers (Eigenstrata)	Kinship. R-package
MACH 1.0.16	none	none	2543887	$\geq 97.5\%$	1) heterozygosity $> 33\%$ ; 2) ethnic outliers; 3) related individuals and duplicates.	MACH2qtl on GRIMP
MACH 1.0.16	none	none	African American = 1922501. European Ancestry 2543887	$\geq 97\%$	No IBD $> 0.125$ . no ethnic outliers from Eigenstrat. no gender discrepancies from self-report versus genotype data	MACH@ QTL
minimac	none	none		$\geq 97\%$	exclusion based on IBD clustering. checked for duplicates. first and second-degree relatives	PLINK
minimac	none	MACH r2	3020488	$\geq 97\%$	1) related participants and duplicates 2) High BAF in $>5$ chromosomes	R
IMPUTE 2.0	none	info $>0$	2692093	$>95\%$	Heterozygosity ( $\pm 3SD$ . Gender discordance)	PLINK
IMPUTE	NA	excl. duplicate IDs	2748910	$>92\%$	duplicate samples (by estimated IBD). individuals with reported/genotyped gender mismatch	Quicktest v0.95
IMPUTE	NA	excl. duplicate IDs	3437411	$>94\%$	duplicate samples (by estimated IBD). individuals with reported/genotyped gender mismatch	Quicktest v0.95
IMPUTEv2	$\geq 1\%$	info $\geq 0.30$	3044097	$\geq 98\%$	Exclusion criteria were:(ii) heterozygosity across all SNPs $\geq 2$ s.d. from the sample mean (iii) evidence of non-European ancestry as assessed by PCA comparison with HapMap3 populations; (iv) observed pairwise IBD probabilities suggestive of sample identity errors; (v). We corrected misclassified monozygotic and dizygotic twins based on IBD probabilities	GenABEL

**Supplementary Table S3** | Comparing the IVW and ZW meta-analysis for the lead GWS SNPs from each locus

Locus	SNP	EA	NEA	EAF	Effect	se	P	Z-score	P
1q22	rs2246476	T	C	0.780	0.037	0.004	$2.64 \times 10^{-19}$	8.97	$2.60 \times 10^{-19}$
1q22	rs2104404	A	C	0.201	-0.029	0.004	$6.94 \times 10^{-11}$	-6.47	$1.01 \times 10^{-10}$
8q24.12	rs10955908	A	C	0.467	-0.028	0.003	$4.53 \times 10^{-16}$	-8.21	$2.18 \times 10^{-16}$
11q13.1	rs1346	A	T	0.812	0.025	0.005	$2.16 \times 10^{-8}$	5.61	$1.95 \times 10^{-8}$
11q24.3	rs1787971	C	G	0.670	-0.020	0.004	$7.73 \times 10^{-8}$	-5.34	$9.18 \times 10^{-8}$
13q14.11	rs9594738	T	C	0.498	0.030	0.003	$5.04 \times 10^{-19}$	8.93	$4.36 \times 10^{-19}$
16p13.3	rs13337626	T	C	0.916	0.039	0.007	$4.82 \times 10^{-8}$	5.38	$7.40 \times 10^{-8}$
19q13.41	rs8111274	T	G	0.291	-0.021	0.004	$8.25 \times 10^{-8}$	-5.37	$7.70 \times 10^{-8}$