

Twenty bone-mineral-density loci identified by large-scale meta-analysis of genome-wide association studies

Fernando Rivadeneira^{1,2,19,20}, Unnur Styrkársdóttir^{3,20}, Karol Estrada^{1,19,20}, Bjarni V Halldórsson^{3,4,20}, Yi-Hsiang Hsu^{5,20}, J Brent Richards^{6–8,20}, M Carola Zillikens^{1,19,20}, Fotini K Kavvoura^{9,20}, Najaf Amin², Yurii S Aulchenko^{2,19}, L Adrienne Cupples¹⁰, Panagiotis Deloukas¹¹, Serkalem Demissie¹⁰, Elin Grundberg^{7,12}, Albert Hofman^{2,19}, Augustine Kong³, David Karasik⁵, Joyce B van Meurs^{1,2,19}, Ben Oostra¹³, Tomi Pastinen^{7,12}, Huibert A P Pols^{1,2}, Gunnar Sigurdsson^{14,15}, Nicole Soranzo^{6,11}, Gudmar Thorleifsson³, Unnur Thorsteinsdóttir^{3,14}, Frances M K Williams⁸, Scott G Wilson^{8,16}, Yanhua Zhou¹⁰, Stuart H Ralston¹⁷, Cornelia M van Duijn^{2,19,20}, Timothy Spector^{8,20}, Douglas P Kiel^{5,20}, Kari Stefansson^{3,14,20}, John P A Ioannidis^{9,18,20} & André G Uitterlinden^{1,2,19,20} for the Genetic Factors for Osteoporosis (GEFOS) Consortium

Bone mineral density (BMD) is a heritable complex trait used in the clinical diagnosis of osteoporosis and the assessment of fracture risk. We performed meta-analysis of five genome-wide association studies of femoral neck and lumbar spine BMD in 19,195 subjects of Northern European descent. We identified 20 BMD loci that reached genome-wide significance (GWS; $P < 5 \times 10^{-8}$), of which 13 map to regions not previously associated with this trait: 1p31.3 (*GPR177*), 2p21 (*SPTBN1*), 3p22 (*CTNNA1*), 4q21.1 (*MEPE*), 5q14 (*MEF2C*), 7p14 (*STARD3NL*), 7q21.3 (*FLJ42280*), 11p11.2 (*LRP4*, *ARHGAP1*, *F2*), 11p14.1 (*DCDC5*), 11p15 (*SOX6*), 16q24 (*FOXL1*), 17q21 (*HDAC5*) and 17q12 (*CRHR1*). The meta-analysis also confirmed at GWS level seven known BMD loci on 1p36 (*ZBTB40*), 6q25 (*ESR1*), 8q24 (*TNFRSF11B*), 11q13.4 (*LRP5*), 12q13 (*SP7*), 13q14 (*TNFRSF11A*) and 18q21 (*TNFRSF11A*). The many SNPs associated with BMD map to genes in signaling pathways with relevance to bone metabolism and highlight the complex genetic architecture that underlies osteoporosis and variation in BMD.

Osteoporosis is characterized by reduced bone mass and microarchitectural deterioration of bone tissue, leading to loss of bone strength and increased risk of fracture. Osteoporosis increases in incidence with age and is an important health threat to hundreds of millions of elderly individuals worldwide. Linkage analysis in monogenic bone disorders such as osteoporosis-pseudoglioma syndrome, sclerosteosis, high-bone-mass syndrome and Paget's disease have yielded important advances in recent years and highlighted the importance of the Wnt signaling¹ and RANK-RANKL-Opg pathways in the regulation of bone mass and bone turnover². However, linkage studies for the common form of osteoporosis have had limited success³. As with other complex diseases, most of the candidate gene association studies

in osteoporosis have produced conflicting results with limited replication⁴, mostly owing to small sample sizes⁵. However, large studies of major candidate gene polymorphisms within the sufficiently powered setting of the GENOMOS consortium have been successful in obtaining consistent evidence of association between some genetic variants, BMD and fracture risk^{6–10}. Concurrently, genome-wide association studies (GWAS), which perform a hypothesis-free search for genetic determinants¹¹, have already identified ten loci that are associated at a GWS level with BMD^{12,13}. Four of these loci encode members of the Wnt and RANK-RANKL-Opg signaling pathways, highlighting their role in monogenic forms of osteoporosis as well as in the common form.

¹Departments of Internal Medicine and ²Epidemiology, Erasmus Medical Center, Rotterdam, The Netherlands. ³deCODE Genetics, Reykjavik, Iceland. ⁴Reykjavik University, Reykjavik, Iceland. ⁵Hebrew SeniorLife, Harvard Medical School, Boston, Massachusetts, USA. ⁶Department of Medicine and ⁷Department of Human Genetics, McGill University, Montréal, Québec, Canada. ⁸Department of Twin Research and Genetic Epidemiology, Kings College London, London, UK. ⁹Department of Hygiene and Epidemiology, University of Ioannina School of Medicine, Ioannina, Greece. ¹⁰Department of Biostatistics, School of Public Health, Boston University, Boston, Massachusetts, USA. ¹¹Wellcome Trust Sanger Institute, Hinxton, Cambridge, UK. ¹²McGill University and Genome Quebec Innovation Centre, Montréal, Québec, Canada. ¹³Department of Clinical Genetics, Erasmus MC, Rotterdam, The Netherlands. ¹⁴Faculty of Medicine, University of Iceland, Reykjavik, Iceland. ¹⁵Department of Endocrinology and Metabolism, University Hospital, Reykjavik, Iceland. ¹⁶School of Medicine & Pharmacology, The University of Western Australia and Department of Endocrinology & Diabetes, Sir Charles Gairdner Hospital, Nedlands, Western Australia, Australia. ¹⁷Rheumatic Diseases Unit, Institute of Genetics and Molecular Medicine, Western General Hospital, University of Edinburgh, Edinburgh, UK. ¹⁸Center for Genetic Epidemiology and Modeling, Tufts Medical Center, Tufts University School of Medicine, Boston, Massachusetts, USA. ¹⁹Netherlands Genomics Initiative (NGI)-sponsored Netherlands Consortium for Healthy Aging (NCHA). ²⁰These authors contributed equally to this work. Correspondence should be addressed to A.G.U. (a.g.uitterlinden@erasmusmc.nl) or J.P.A.I. (jioannid@cc.uoi.gr).

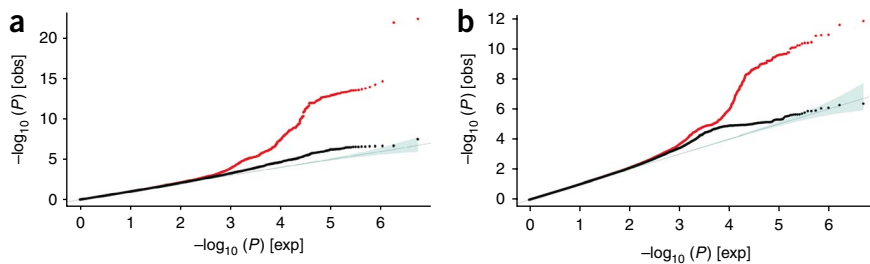


Figure 1 Quantile-quantile (Q-Q) plots. **(a)** Lumbar spine BMD. **(b)** Femoral neck BMD. The plots compare additive model statistics to those expected under the null distribution using fixed-effects for all analyzed HapMap CEU imputed SNPs passing quality control criteria in the studies (red lines) and after exclusion of all genome-wide significant and correlated ($r^2 > 0.1$) SNPs (blue lines).

BMD is used in clinical practice for the diagnosis of osteoporosis and in the assessment of fracture risk¹⁴. BMD is usually measured at the hip (femoral neck) and lumbar spine, which are common sites of fracture. However, BMD measurements at different skeletal sites are predictive of fracture in general because of their high correlation ($r^2 \sim 0.60$)¹⁵. From a genetic perspective, BMD at both spine and hip is a complex but highly heritable trait ($h^2 \sim 0.60$ – 0.80)¹⁶. As shown for human height^{17–19}, dozens and possibly hundreds of loci with small effects can be expected to influence variation in BMD. Detection and reliable documentation of these loci with weak effects require studies with comprehensive coverage of the genome and very large sample sizes. Here we report the findings of a large-scale meta-analysis of 19,195 adults from five GWAS that examined BMD of the lumbar spine (LS-BMD) and the femoral neck (FN-BMD) within the setting of the Genetic Factors for Osteoporosis (GEFOS) Consortium.

RESULTS

Samples, genotyping and analysis of genome-wide scans

The five study populations were from the Rotterdam Study (RS, $n = 4,987$), Erasmus Rucphen Family Study (ERF, $n = 1,228$), Twins UK Study (TUK, $n = 2,734$), deCODE Genetics Study (dCG, $n = 6,743$) and Framingham Osteoporosis Study (FOS, $n = 3,503$). The ages of participants ranged from 18 to 96 years. All studies had a majority of women (range 57–88%) in their samples, with TUK including women only. Additional characteristics of the study populations and subject exclusion criteria are presented in **Supplementary Tables 1 and 2**, respectively.

Identification of BMD loci

Association results (corrected by the genomic control method²⁰) of all HapMap CEU imputed SNPs that passed quality control criteria in each study (**Supplementary Table 3**) were meta-analyzed using METAL²¹. We declared results genome-wide significant at $\alpha = 5 \times 10^{-8}$ after adjusting for all common variant tests in the human genome^{22,23}. We investigated whether there was an excess of significant associations by comparing the test statistics to those expected under the null distribution using interquantile-quantile plots (corrected for overall meta-analysis genomic control $\lambda_{\text{LS-BMD}} = 1.09$ and $\lambda_{\text{FN-BMD}} = 1.08$). We found strong (and not early) deviation of the observed statistics from the null distribution for both BMD traits, corresponding to an excess of significant and probably true associations (**Fig. 1**). Exclusion of all SNPs within 500 kb of the SNPs associated at a GWS level and correction for overall meta-analysis genomic control still left many more SNPs associated with BMD than would be expected by chance alone. This indicates that, among many false positives appearing at less stringent statistical thresholds, additional truly associated BMD variants might exist.

The meta-analysis identified 467 SNPs from 20 genomic loci exceeding the GWS threshold of association with the BMD traits (**Fig. 2**). Of these, 15 loci associated with LS-BMD (**Supplementary Tables 4a and 5a**) and 10 with FN-BMD (**Supplementary Tables 4b and 5b**); 5 of these loci were associated with both skeletal sites. The effect sizes and significance of the top SNPs from the 20 regions that contained markers associated with LS-BMD and FN-BMD at GWS are shown in **Table 1**. Applying correction for overall meta-analysis genomic control resulted in 7 of the 20 loci not reaching GWS (with $7.1 \times 10^{-7} >$

$P > 5.0 \times 10^{-8}$). For most markers, heterogeneity was not very large or statistically significant.

Despite the correlation between LS-BMD and FN-BMD measurements, we found site-specific effects. Seven of the 20 loci showed evidence for skeletal site specificity ($P < 0.05$), and among these 3 showed strong evidence for site specificity ($P < 1 \times 10^{-6}$). This site specificity is to be expected given the differences in heritability and the fact that the genetic correlation (or fraction of ‘shared’ heritability) between the measurements is considerably less than 1 (**Supplementary Table 6**).

Genes in associated regions and their functions

Of the 20 BMD loci identified in this scan at a GWS level, 7 have been reported previously as GWS^{12,13,24}, whereas the remaining 13 have not. Of these 13 loci not previously associated with BMD at a GWS level, 4 were suggestively associated in previous reports^{12,13} and 9 are loci not previously reported to be associated with BMD. In **Supplementary Table 7**, we present a summary of relevant gene annotations including

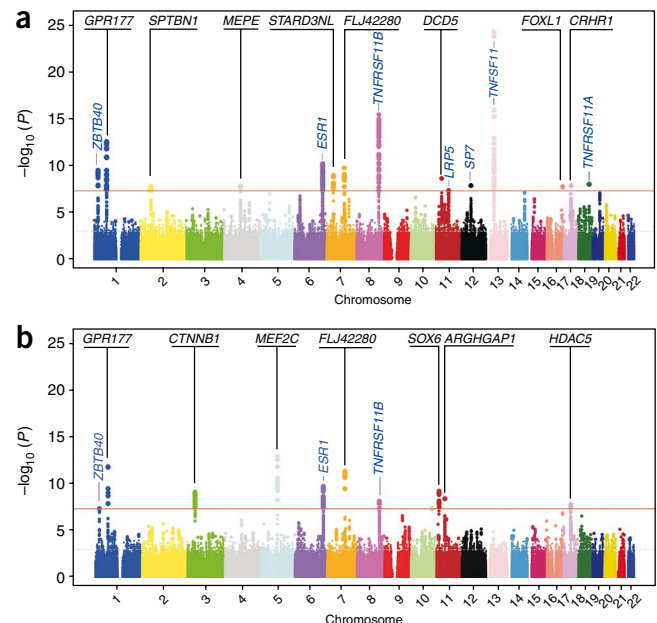


Figure 2 Manhattan plots. **(a,b)** Plots display newly discovered and previously reported (known) loci associated at genome-wide significant level (GWS) with lumbar spine BMD **(a)** and femoral neck BMD **(b)** for all 2,543,686 HapMap CEU-imputed SNPs analyzed using fixed-effects. The 13 new GWS loci are in underlined type. Previously reported GWS loci are in blue type.

Table 1 Top genome-wide significant markers of the 20 loci associated at GWS with lumbar spine and/or femoral neck BMD

Marker information						Lumbar spine BMD						Femoral neck BMD							
Locus	SNP	A1/A2	Freq. A1	Type ^a	Closest gene	Dist. ^b (kb)	Effect estimate (in s.d.)			Heterogeneity			Effect estimate (in s.d.)			Heterogeneity			Site-spec. H ₀ P value ^e
							Beta (β _{LS})	s.e.m.	P value	OMA-GC P value ^c	Q P value ^d	I ²	Beta	s.e.m.	P value	OMA-GC P value ^c	Q P value ^d	I ²	
New loci associated with BMD at GWS level																			
1p31.3	rs1430742	C/T	0.21	I	<i>GPR177</i>	44.0	0.105	0.014	2.6E-13	2.5E-12	0.26	21	0.100	0.014	1.8E-12	1.2E-11	0.75	0	0.82
	rs2566755	C/T	0.21	G	<i>GPR177</i>	44.3	0.104	0.014	3.3E-13	3.1E-12	0.27	20	0.100	0.014	1.7E-12	1.1E-11	0.76	0	0.83
3p22	rs87938	A/G	0.45	G	<i>CTNNB1</i>	103.3	-0.043	0.012	1.7E-04	3.1E-04	0.24	23	-0.070	0.011	8.1E-10	3.4E-09	0.14	34.5	0.10
5q14	rs1366594	C/A	0.45	G	<i>MEF2C</i>	197.0	0.005	0.012	0.65	0.66	NC	NC	-0.085	0.011	1.3E-13	1.1E-12	0.62	0	3.2E-08
7p14	rs1524058	T/C	0.40	I	<i>STARD3NL</i>	81.7	-0.070	0.012	1.1E-09	5.2E-09	0.18	30	-0.038	0.011	8.9E-04	1.4E-03	0.05	48	0.05
7q21.3	rs4729260	G/C	0.32	I	<i>FLJ42280</i>	14.9	-0.081	0.013	1.7E-10	9.5E-10	0.14	35	-0.085	0.012	9.4E-12	5.4E-11	0.77	0	0.82
	rs7781370	T/C	0.34	I	<i>FLJ42280</i>	0.7	-0.074	0.012	1.1E-09	5.5E-09	0.12	37	-0.083	0.012	4.7E-12	2.9E-11	0.68	0	0.60
11p14.1	rs16921914	A/G	0.27	G	<i>DCDC5</i>	61.6	0.077	0.013	2.3E-09	1.0E-08	0.52	0	0.038	0.013	0.003	0.005	NC	NC	0.03
11p15	rs7117858	G/A	0.20	G	<i>SOX6</i>	297.3	-0.042	0.014	0.004	0.005	NC	NC	0.088	0.014	6.4E-10	2.7E-09	0.70	0	1.5E-10
16q24	rs10048146	G/A	0.19	G	<i>FOXL1</i>	95.4	-0.093	0.016	1.7E-08	6.0E-08	0.21	28	-0.085	0.016	1.7E-07	4.9E-07	0.96	0	0.73
17q12	rs9303521	T/G	0.46	G	<i>CRHR1</i>	56.5	-0.068	0.012	1.4E-08	5.0E-08	0.05	49	-0.055	0.012	3.6E-06	8.3E-06	0.07	45	0.46
Suggestive loci now associated with BMD at GWS level																			
2p21	rs11898505	A/G	0.34	G	<i>SPTBN1</i>	1.1	0.067	0.012	1.6E-08	6.3E-08	0.39	6	0.027	0.012	0.02	0.03	NC	NC	0.02
4q21.1	rs1471403	T/C	0.34	G	<i>MEPE</i>	7.3	0.068	0.012	1.5E-08	5.7E-08	0.18	30	0.059	0.012	7.8E-07	2.0E-06	0.56	0	0.58
11p11.2	rs7932354	T/C	0.29	I	<i>ARHGAP1</i>	0.1	0.056	0.013	1.1E-05	2.4E-05	0.61	0	0.073	0.012	4.0E-09	1.5E-08	0.68	0	0.32
17q21	rs228769	G/C	0.20	I	<i>HDAC5</i>	7.8	0.067	0.014	4.0E-06	1.0E-05	0.87	0	0.081	0.014	1.7E-08	5.8E-08	0.94	0	0.49
Known loci associated with BMD at GWS level																			
1p36	rs7524102	G/A	0.17	G	<i>ZBTB40</i>	79.9	0.094	0.015	3.2E-10	1.7E-09	0.27	19.5	0.079	0.015	8.8E-08	2.6E-07	0.64	0	0.48
	rs6426749	C/G	0.17	I	<i>ZBTB40</i>	66.9	0.107	0.017	7.6E-10	3.8E-09	0.74	0	0.082	0.015	4.8E-08	1.5E-07	0.64	0	0.26
6q25	rs2504063	A/G	0.40	G	<i>ESR1</i>	38.0	-0.078	0.012	6.1E-11	3.7E-10	0.03	52.7	-0.066	0.012	3.0E-08	9.6E-08	0.84	0	0.45
	rs2941740	G/A	0.43	I	<i>C6orf97</i>	67.3	0.070	0.012	2.0E-09	9.3E-09	0.01	59.0	0.073	0.012	2.0E-10	9.1E-10	0.02	57.2	0.84
8q24	rs2062377	T/A	0.44	I	<i>TNFRSF11B</i>	43.0	0.094	0.012	3.5E-16	5.7E-15	0.31	15.4	0.062	0.011	5.4E-08	1.7E-07	0.61	0	0.05
	rs11995824	G/C	0.55	I	<i>TNFRSF11B</i>	48.3	-0.093	0.012	1.1E-15	1.6E-14	0.24	22.3	-0.066	0.011	7.1E-09	2.6E-08	0.48	0	0.10
11q13.4	rs599083	G/T	0.31	G	<i>LRP5</i>	24.4	-0.067	0.012	4.7E-08	1.7E-07	0.50	0	-0.047	0.012	9.7E-05	0.0002	0.76	0	0.25
12q13	rs2016266	G/A	0.32	G	<i>SP7</i>	1.6	0.070	0.012	1.3E-08	5.2E-08	0.93	0	0.046	0.012	1.9E-04	0.0003	0.81	0	0.16
13q14	rs9533090	T/C	0.50	I	<i>AKAP11</i>	54.0	-0.120	0.012	5.4E-25	4.6E-23	0.02	57.1	-0.041	0.011	3.9E-04	0.0006	0.39	5.7	1.1E-06
18q21	rs884205	A/C	0.27	I	<i>TNFRSF11A</i>	1.4	-0.078	0.014	9.4E-09	3.8E-08	0.90	0	-0.039	0.013	0.004	0.005	NC	NC	0.04

Boldface indicates $P < 5 \times 10^{-8}$. NC, not calculated; $P > 0.001$.

^aType, type of SNP: G, genotyped (in at least one study); I, imputed. ^bDistance to coding region. ^cOMA-GC, overall meta-analysis genomic control. ^dQ P value, Q-statistic P value.

^eSite-spec. H₀, site-specificity null hypothesis, i.e., $\beta_{LS} = \beta_{FN}$.

related pathways, monogenic syndromes, knockout mouse models and additional functional details of the genes that are most likely to underlie the associated signals at these 13 loci.

New loci associated with BMD at GWS level

We identified nine previously unreported loci that showed associations with BMD, for which we present Forest plots of effects (Fig. 3) and regional association plots (Supplementary Fig. 1) for the top SNPs.

1p31.3 (*GPR177*). Two common SNPs (minor allele frequency (MAF) = 0.21) in complete pairwise linkage disequilibrium (LD) (rs1430742 and rs2566755) were associated at a GWS level with both FN-BMD and LS-BMD. Both top SNPs are located within an intron of the *GPR177* (G protein-coupled receptor 177, also called WNTLESS homolog) gene. *GPR177* is part of the highly evolutionarily conserved Wnt signaling pathway²⁵, which is involved in bone cell differentiation and development. The gene is a positive regulator of the I-κB kinase-NF-κB cascade, part of the RANK system. Cross-talk between the NF-κB and mitogen-activated protein kinase (MAPK) pathways has been indicated by the identification of genes that are expressed in both pathways, including several G protein-coupled receptors²⁶.

3p22 (*CTNNB1*). rs87939 (MAF = 0.45), located 103 kb upstream of the *CTNNB1* gene on chromosome 3, was associated at a GWS level with FN-BMD. *CTNNB1* encodes β-catenin, which is integral

to the Wnt signaling pathway and as such is an excellent candidate for BMD regulation, considering that Wnt signaling controls the process of bone resorption by negatively regulating *TNFRSF11B* expression in osteoblasts²⁷.

5q14 (*MEF2C*). The rs1366594 top SNP (MAF = 0.45), with another 60 GWS markers at this locus, showed skeletal site specificity, being associated with FN-BMD only. The SNP is 197 kb upstream of the *MEF2C* (MADS box transcription enhancer factor 2, polypeptide C) gene with no other known annotation within the large LD block. *MEF2C* is a transcription factor that is highly expressed in muscle and allows transcriptional cross-talk between the Ca²⁺/calmodulin-dependent kinase (CaMK) and mitogen-activated protein kinase (MAPK) signaling pathways by signal-dependent dissociation from histone deacetylases (HDACs). *MEF2C* interacts with HDAC4 and HDAC5 (see 17q21 region below), resulting in repression of the transcriptional activity of *MEF2C*. In mice, selective degradation of HDACs by the proteasome allows *MEF2C* to activate the slow myofiber gene program, resulting in enhanced endurance during physical exercise²⁸.

7p14-p13 (*STARD3NL*). rs1524058 (MAF = 0.40), located 81 kb upstream of the *STARD3NL* (*STARD3* N-terminal like) gene (which encodes a cholesterol endosomal transporter), was associated at a GWS level with LS-BMD but less strongly with FN-BMD. A recent study of Asian individuals²⁹ identified a SNP (rs1721400) in this 7p14

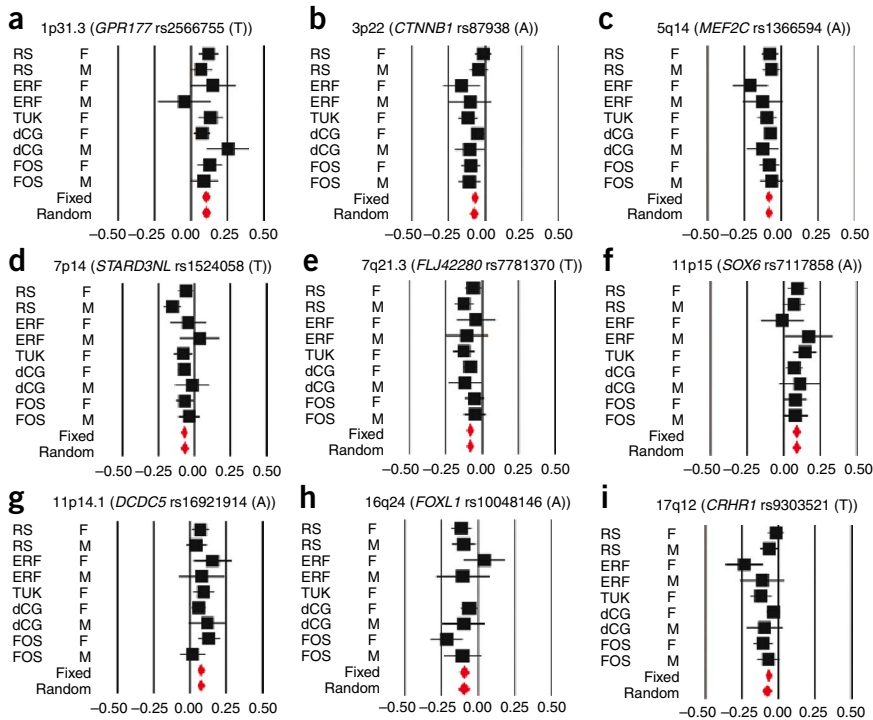


Figure 3 Forest plots for the top SNPs for each of the nine newly discovered loci. (a) 1p31.3. (b) 3p22. (c) 5q14. (d) 7p14. (e) 7q21.3. (f) 11p15. (g) 11p14.1. (h) 16q24. (i) 17q12. Black squares represent effect estimate and 95% CI for each study; red diamonds are summary effect estimates. Measurements units are in BMD s.d.

region that was associated with BMD measurements of the radius, tibia and heel (lowest $P = 1.4 \times 10^{-7}$). However, the rs1721400 SNP was not associated with either LS-BMD or FN-BMD in our study ($P > 0.05$). rs1721400 (mapping close to *SFRP4*; secreted frizzled-related protein 4) is in very low LD with our top SNP rs1524058 in HapMap individuals of European descent ($r^2 = 0.015$, $D' = 0.194$), whereas LD between markers is considerably higher in Chinese or Japanese individuals ($r^2 = 0.230$, $D' = 0.931$). Thus, an underlying signal common to both Europeans and Asians might still be captured by these SNPs, considering the differences in LD across populations.

7q21.3 (FLJ42280). Several SNPs are GWS in this region, with the two top SNPs being in moderate pairwise LD ($r^2 = 0.36$; $D' = 0.84$) and associated at GWS with both LS-BMD and FN-BMD. The SNPs are located within (rs4729260) and just upstream (rs7781370) of *FLJ42280*, of unknown function. There are several genes within the ~480-kb LD region, of which *SHFM1* (split hand/foot malformation (ectrodactyly) type 1) is the closest gene (87–185 kb away).

11p15 (SOX6). rs7117858 (MAF = 0.20) was associated at GWS level with FN-BMD only, displaying strong evidence for skeletal site specificity. The SNP is located 297 kb upstream of the *SOX6* (SRY (sex determining region Y)-box 6) gene, which encodes a SOX family transcription factor defined by a conserved high-mobility group DNA-binding domain. The gene is expressed in various tissues, most abundantly in skeletal muscle. *Sox6* knockout mice show early lethality due to cardiac insufficiency and present with mild skeletal abnormalities that affect the size and mineralization of endochondral elements³⁰. Other SOX-family proteins regulate RUNX2-mediated differentiation of mesenchymal cells during endochondral ossification (skeletogenesis)³¹.

11p14.1 (DCDC5; DCDC1). rs16921914 (MAF = 0.27) is the only marker in the 11p14.1 region that is associated with LS-BMD at a GWS level. All other associated SNPs are in moderate LD with rs16921914 ($r^2 < 0.70$) and show less strong associations ($P > 1 \times 10^{-7}$). The SNP is located 62 kb downstream of *DCDC1* (doublecortin domain containing 1) and 73 kb upstream of *DCDC5*. Doublecortin domains are highly conserved elements that serve as protein-interaction platforms³². Mutations in members of this protein superfamily are linked to several neurogenetic diseases, and to our knowledge these proteins are not expressed in bone.

16q24.3 (FOXC2; FOXL1). rs10048146 (MAF = 0.19) was associated with LS-BMD and is located on the subtelomeric region of chromosome 16, about 95 kb downstream from a cluster of small (1-kb) genes of the 'forkhead' (or winged helix) gene family. The genes are mainly expressed in the gastrointestinal mucosa (*FOXL1*) or are involved in adipocyte metabolism and early stage chondrogenic differentiation (*FOXC2*). *FOXC2* stimulates osteoblast differentiation of mesenchymal cells through activation of canonical Wnt- β -catenin signals³³, whereas *FOXC2* expression is regulated by bone morphogenetic

proteins³⁴. Skeletal defects of the spine have been reported in *Foxc2* mouse knockout models³⁵, and deletions and inactivating mutations affecting the FOX gene cluster can cause severe malformations of the VACTER type in humans, which include vertebral malformations³⁶.

17q12-q22 (CRHR1). rs9303521 (MAF = 0.46) was associated with LS-BMD and is located 56 kb from the *CRHR1* (corticotrophin-releasing factor receptor) gene. Among other genes in this LD region, *MAP3K14* is a candidate that could be involved in bone-active pathways, particularly through the activation of NF- κ B.

Suggestive loci now associated with BMD at GWS level

Four BMD loci for which this is the first demonstration of a GWS level of association were suggestively associated with BMD in previous reports^{12,13}. We present Forest plots of effects (Fig. 4) and regional association plots (Supplementary Fig. 2) for these SNPs.

2p21 (SPTBN1). rs11898505 (MAF = 0.34), located in an intron of *SPTBN1* (spectrin, beta, non-erythrocytic 1), which encodes a major cytoskeletal scaffolding protein, was associated with LS-BMD. The same SNP was previously shown to be associated with BMD and fractures, even though it did not reach GWS level for BMD in the same study¹³. In mice, disruption of β -spectrin isoforms (Elf) leads to the disruption of TGF- β signaling by Smad proteins³⁷.

4q21.1 (MEPE). The 4q21.1 region contains a cluster of structurally and phylogenetically related genes that encode matricellular phosphoglycoproteins with functions in bone formation and growth³⁸. The top associated rs1471403 SNP (MAF = 0.34) is located 7 kb 3' of *MEPE* (matrix, extracellular, phosphoglycoprotein; also known as osteoblast/osteocyte factor 45), 42 kb 3' of *IBSP* (integrin-binding

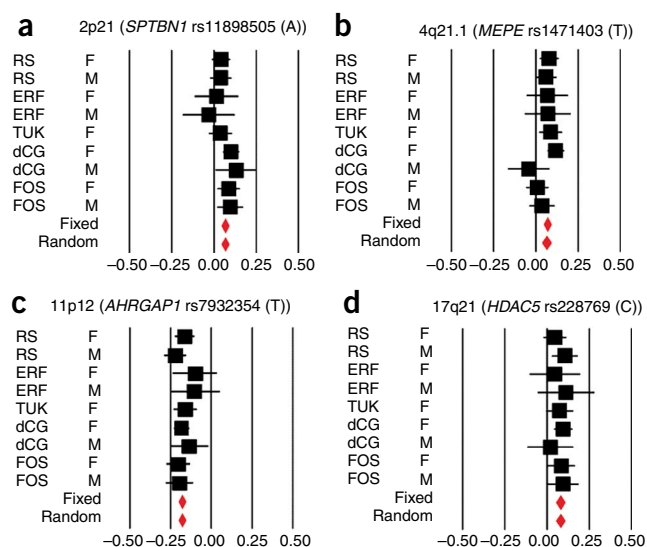


Figure 4 Forest plots for the top SNPs for each of the four loci attaining GWS for the first time in this study. (a) 2p21. (b) 4q21.1. (c) 11p11.2. (d) 17q21. Black squares represent effect estimate and 95% CI for each study; red diamonds are summary effect estimates. Measurements units are in BMD s.d.

sialoprotein) and 122 kb 5' of *SPP1* (secreted phosphoprotein 1, also known as osteopontin). *IBSP* and *SPP1* are highly expressed in osteoblasts, osteoclasts and hypertrophic chondrocytes. *MEPE* is predominantly expressed by osteocytes in human bone, and inhibits bone formation. Mouse knockout models of these genes show diverse skeletal phenotypes. *Mepe* (*Of45*) knockout mice show increased bone mass and inhibition of age-related bone loss³⁹; *Ibsp* knockout mice show high trabecular bone density with low bone turnover but respond to bone loss caused by disuse⁴⁰; and *Spp1* knockout mice have high trabecular bone mass and are resistant to bone loss⁴¹. Previously, a nonsynonymous SNP in *IBSP* (rs1054627, G195E) in moderate LD with rs1471403 ($r^2 = 0.2$, $D' = 0.8$) was reported as being suggestively associated with hip BMD¹².

11p11.2 (*ARHGAP1*; *LRP4*). At 11p11.2, a large LD block extends through a region that encompasses several genes, including *C11orf49*, *LRP4*, *CKAP5*, *F2* and *ARHGAP1*. Two fully correlated SNPs ($r^2 = 1$; $MAF = 0.29$) — rs7932354, which lies in the promoter region of *ARHGAP1* (Rho GTPase activating protein 1), and rs2070852, located in intron 5 of *F2* (coagulation factor II) — were associated with FN-BMD at GWS level. Other correlated SNPs in the region ($r^2 = 0.2$ – 0.8) were previously suggestively associated with hip BMD and attributed to *LRP4* (low-density lipoprotein receptor-related protein 4)¹³. *ARHGAP1*, a ubiquitous GTPase-activating protein that represses RhoA, is another good candidate. RhoA is a small G protein of the Rho family that regulates cell morphology by actin-cytoskeleton reorganization, which is thought to be a potential commitment switch in the differentiation of mesenchymal stem cells to osteoblasts⁴². In addition, *Arhgap1* knockout mouse models show a marked skeletal phenotype, including a threefold reduction in BMD, decreased cortical thickness and bone fragility in older animals⁴³.

17q21 (*HDAC5*; *C17orf53*). The 17q21 region contains more than 30 genes in 1 Mb surrounding the top rs228769 SNP ($MAF = 0.20$), which is located 8 kb upstream of *HDAC5* (histone deacetylase 5) and 26 kb upstream of *C17orf53*. A nonsynonymous SNP in *C17orf53*,

rs227584 (T126P), which is moderately correlated with rs228769 ($r^2 = 0.64$), was found to be associated with hip BMD, albeit not at GWS level¹². These SNPs probably represent the same signal. *HDAC5* encodes a class II histone deacetylase (homologous to yeast Hda1) that is ubiquitously expressed and responsible for the deacetylation of lysine residues on the N-terminal part of the core histones (H2A, H2B, H3 and H4). Histone acetylation and deacetylation are important in transcriptional regulation, cell cycle progression and developmental events, particularly for myocyte differentiation⁴⁴. In undifferentiated myoblasts, *HDAC5* is present in the nucleus where it binds to the myocyte enhancer *MEF2C* (see above) to repress transcription and inhibit muscle maturation. In bone, recruitment of class II histone deacetylases such as *HDAC5* is needed for TGF- β mediated osteoblast differentiation⁴⁵, which occurs through inhibition of Runx2 function by Smad3 (ref. 46).

Known loci associated with BMD at GWS level

We have replicated at GWS level seven loci that were associated with BMD in previous GWAS^{12,13,24}: 1p36 (*ZBTB40*), 6q25 (*ESR1*), 8q24 (*TNFRSF11B*, which encodes Opg), 11q13.4 (*LRP5*), 12q13 (*SP7*, which encodes osterix, or *OSX*), 13q14 (*TNFSF11*, which encodes RANKL) and 18q21 (*TNFRSF11A*, which encodes RANK). Regional association plots for all seven regions are included in **Supplementary Figure 3**. In the 6q25 (*ESR1*) region, at least two independent GWS signals are seen on either side of a high-recombination-rate peak (**Supplementary Fig. 3b**). This locus showed a complex pattern of association in a previous study, indicating three independent signals at the locus¹³. In **Supplementary Table 8**, we report the associations observed in this study for all loci reported previously as attaining GWS or suggestive association^{12,13,24}. SNPs from the 6p21.32 (MHC locus), 14q32 (*MARK3*) and 17q21 (*SOST*) regions described previously to be GWS¹³ were still significantly associated with BMD in our study, but not at a GWS level. Note that in the previous study the 14q32 and the 17q21 regions were associated with total hip BMD, which differs from the femoral neck BMD phenotype used in the current study.

Gene expression quantitative trait locus (eQTL) associations

We tested the association of SNPs (or proxies) from the 13 newly GWS associated regions with *cis*-allelic expression of gene transcripts in primary human osteoblasts. All SNPs associated with gene expression at $P < 0.05$ and located in the same LD block as the strongest associated variants ($D' \geq 0.8$) are shown in **Supplementary Table 9**. Associations were seen for transcripts of *GPR177*, *MEF2C* and *FOXC2*. Similarly, for variants in (or in LD with variants in) *MEPE*, the most significant correlation with expression in osteoblasts was seen with *IBSP*, whereas *MEPE* seems not to be highly expressed in osteoblasts. However, the statistical evidence is not fully conclusive as only subtle over-representation of the associated loci was observed (10.5% versus 7% for nonassociated control SNPs, $\chi^2 = 8.9$ and $P = 0.003$). The small over-representation of the associated loci indicates that several of the associated genes might be expressed in cells and tissues other than those from osteoblast lineages.

Combined effect of the 20 GWS BMD loci and fracture risk

We examined the combined effect of the top SNPs arising from the 20 associated BMD loci in 4,983 individuals from the prospective population-based Rotterdam Study. Risk allele counts were derived from the top associated SNPs from the 15 LS-BMD and 10 FN-BMD loci, all of which showed a normal frequency distribution in the study population (**Fig. 5**). The 15 LS-SNPs combined explained

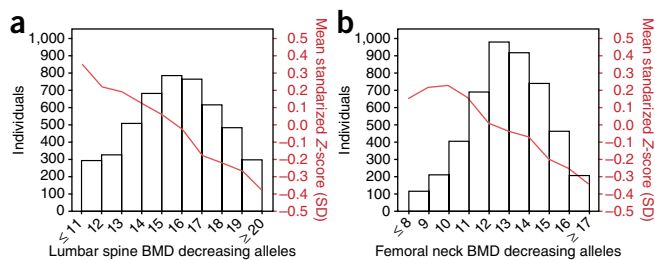


Figure 5 Histogram and line plot modeling in the Rotterdam Study of the combined allelic effect across all genome-wide significant associated loci. **(a)** Lumbar spine BMD. **(b)** Femoral neck BMD. Subjects were classified according to the number of BMD-decreasing (risk) alleles at the lumbar spine (20 SNPs) and the femoral neck (15 SNPs) BMD loci. The mean for each risk allele-count group was determined and the extremes of the distribution counts were pooled into the nearest risk allele-count group of >100 individuals.

~2.9% of the variance in LS-BMD and the 10 FN-SNPs combined explained ~1.9% of the variance in FN-BMD. A highly significant linear decrease in the mean LS-BMD and FN-BMD of individuals was seen with increasing numbers of 'low BMD' risk alleles: those carrying 20 or more alleles that associated with low LS-BMD ('low BMD' alleles, $n = 245$) had 0.74 s.d. (~0.13 g/cm²) lower BMD ($P = 2 \times 10^{-18}$) than those carrying 11 or fewer 'low-BMD' alleles ($n = 291$). A similar (but less pronounced) trend was seen for FN-BMD ($P = 0.0001$). The association between the compound allelic scores and the risk of fracture were assessed in 2,727 radiographically screened individuals (302 vertebral fracture cases) and in 4,865 individuals followed up an average of 8.2 years (672 nonvertebral fractures). The compound FN-BMD allelic score was significantly associated with the risk of incident nonvertebral fracture in the Rotterdam Study dataset (HR = 1.042, 95% CI [1.003, 1.084]; $P = 0.04$), whereas it was borderline significant for association with the risk of vertebral fracture (OR = 1.061, 95% CI [0.997, 1.129]; $P = 0.06$). In contrast, the compound LS-BMD allelic score was significantly associated with the risk of vertebral fracture (OR = 1.061, 95% CI [1.009, 1.116]; $P = 0.02$), whereas it was not significant for association with the risk of incident nonvertebral fracture (HR = 1.025, 95% CI [0.993, 1.058]; $P = 0.13$). Adjustment for BMD showed that at least 54% of the genetic effect on incident nonvertebral fracture could be explained by FN-BMD (HR_{adjusted} = 1.019, 95% CI [0.980, 1.060]; $P = 0.35$), whereas 78% of the genetic effect on vertebral fracture was through LS-BMD (OR_{adjusted} = 1.013, 95% CI [0.962, 1.068]; $P = 0.62$). Despite power limitations, consistent significant associations between the compound BMD allelic scores were observed with the risk of both vertebral and nonvertebral fracture.

DISCUSSION

The GEFOS Consortium has been assembled to identify the genetic determinants of osteoporosis and fracture. This study represents the first step in a collaborative effort and expands our knowledge of the underlying genetics of BMD, a clinical measurement used for the diagnosis of osteoporosis and the assessment of fracture risk. BMD measures of the lumbar spine and femoral neck were analyzed independently because, despite a relatively high phenotypic correlation, the genetic correlation is not perfect. This is illustrated by the site specificity detected in some of the associations, probably reflecting genuine biological mechanisms (differences in cortical versus trabecular bone content), but also intrinsic measurement differences (artifacts influencing BMD values, such as osteophytes of the lumbar spine or aortic calcifications).

Performing meta-analysis of GWAS has limitations. False positive associations might be generated from multiple hypothesis testing and population stratification. Nevertheless, we applied established methods to minimize the impact of multiple testing by applying stringent GWS thresholds to determine significance. Also, our findings are constrained by the power of our current sample, suited to identify effect sizes explaining ~0.2% of the variance of the trait. Further, we are not powered to detect real effects of the same (small) magnitude arising within specific sex and/or age groups. Similarly, owing to power limitations, we cannot address potential gene-gene and gene-environment interactions or the effects of rare alleles that are not captured by the SNP panels screened in GWAS, nor can we determine the effect on fracture risk. Despite the study's being underpowered to assess heterogeneity of effects, some markers showed significant heterogeneity and were not GWS when analyzed under the random effects model. Further evaluation of such markers in larger populations is needed to determine the source(s) of heterogeneity across datasets, which may include inaccuracies in the imputations, subtle differences in phenotype ascertainment, differences in linkage to the causal variant, differences in environmental modifying factors, or genuinely different genetic effects between populations⁴⁷. Achieving such a sufficiently powered setting is the target of the expanding GEFOS Consortium.

All our most highly associated markers showed high-quality imputation scores, high correlation after *de novo* genotyping and/or at least one genotyped proxy in complete LD that was also associated at GWS level (**Supplementary Table 10**). Thus, we can exclude imputation inaccuracies as a major source of heterogeneity and/or false-positive associations. In addition, we excluded individuals with non-European profiles, which confines our current findings to the context of populations of Northern European descent. Similarly, test statistics corrected by the genomic inflation factors that affected each study makes it unlikely that population stratification or relatedness (like those observed in the ERF, TUK and dCG populations) are important in our associations. In addition, we examined the effect of applying a second correction for overall meta-analysis genomic control in our results. Only one locus (*LRP5*) was substantially below the GWS threshold after correction. This is remarkable considering that the association of variants in *LRP5* has been consistently replicated at GWS level in at least two previous efforts^{10,24}.

In summary, we identified and/or confirmed at least 20 loci associated with lumbar spine and femoral neck BMD, highlighting the complex genetic architecture that underlies the variation in BMD. However, these loci explain only a small fraction of the variance in BMD and hence an even smaller fraction of the heritability for fracture risk. Nevertheless, these findings highlight molecules within new and known biological pathways that influence BMD variation, particularly the Wnt and NF- κ B signaling pathways (including about half of the identified loci: *GPR177*, *CTNBN1*, *FOXC2*, *LRP5*, *SPTBN1*, *HDAC5*, *TNFSF11*, *TNFRSF11A* and *TNFRSF11B*). None of the SNPs we have identified can be unequivocally designated as the underlying 'true' variants driving the associations. Additional efforts to identify such variants are warranted to maximize the application of this genetic knowledge to the prediction of risk in individuals and the translation into new pharmacological agents for treating osteoporosis. Increasing the sample size further will aid the identification of additional loci associated not only with BMD but also with risk of fracture, the ultimate consequence of osteoporosis.

METHODS

Methods and any associated references are available in the online version of the paper at <http://www.nature.com/naturegenetics/>.

Accession numbers. OMIM^{48,49}: *GPR177* (611514); *CTNNB1* (116806); *MEF2C* (600662); *STARD3NL* (611759); *FLJ42280* (none); *SHFM1* (183600); *SOX6* (607257); *DCDC5* (612321); *DCDC1* (608062); *FOXC2* (602402); *FOXL1* (603252); *CRHR1* (<http://www.ncbi.nlm.nih.gov/entrez/dispomim.cgi?id=122561>); *SPTBN1* (182790); *MEPE* (605912); *IBSP* (147563); *SPP1* (166490); *ARHGAP1* (602732); *LRP4* (604270); *F2* (176930); *HDAC5* (605315); *C17orf53* (none); *ZBTB40* (612106); *ESR1* (133430); *C6orf97* (none); *TNFRSF11B* (602343); *LRP5* (603506); *SP7* (606633); *TNFSF11* (602642); *AKAP11* (604696); *TNFRSF11A* (603499).

Note: Supplementary information is available on the Nature Genetics website.

ACKNOWLEDGMENTS

We thank all study participants for making this work possible. This research and the Genetic Factors for Osteoporosis (GEFOS) consortium (<http://www.gefos.org>) have been funded by the European Commission (HEALTH-F2-2008-201865-GEFOS). **Rotterdam Study (RS):** This study was funded by the Netherlands Organization of Scientific Research NWO Investments (175.010.2005.011, 911-03-012), the Research Institute for Diseases in the Elderly (014-93-015; RIDE2) and the Netherlands Genomics Initiative (NGI)/Netherlands Organization for Scientific Research (NWO) project 050-060-810. We thank P. Arp, M. Jhamai, M. Moorhouse, M. Verkerk and S. Bervoets for their help in creating the GWAS database. The Rotterdam Study is funded by Erasmus Medical Center and Erasmus University, Rotterdam, Netherlands Organization for the Health Research and Development (ZonMw), the Research Institute for Diseases in the Elderly (RIDE), the Ministry of Education, Culture and Science, the Ministry for Health, Welfare and Sports, the European Commission (DG XII) and the Municipality of Rotterdam. We thank the staff from the Rotterdam Study, particularly L. Buist and J.H. van den Boogert and also the participating general practitioners and pharmacists. **Erasmus Rucphen Family (ERF):** The study was supported by grants from the Netherlands Organization for Scientific Research (NWO), Erasmus MC and the Centre for Medical Systems Biology (CMSB). We thank all general practitioners for their contributions, P. Veraart for help in genealogy, J. Vergeer for supervision of the laboratory work and P. Snijders for help in data collection. **Twins UK (TUK):** The study was funded by the Wellcome Trust, the Arthritis Research Campaign, the Chronic Disease Research Foundation, the Canadian Institutes of Health Research (J.B.R.), the European Society for Clinical and Economic Aspects of Osteoporosis (J.B.R.) and the European Union FP-5 GenomEUtwin Project (QLG2-CT-2002-01254). The study also receives support from a National Institute for Health Research (NIHR) comprehensive Biomedical Research Centre award to Guy's & St. Thomas' NHS Foundation Trust in partnership with King's College London. We thank the staff of the Twins UK study; the DNA Collections and Genotyping Facilities at the Wellcome Trust Sanger Institute for sample preparation; Quality Control of the Twins UK cohort for genotyping (in particular A. Chaney, R. Ravindrarajah, D. Simpkin, C. Hinds and T. Dibling); P. Martin and S. Potter of the DNA and Genotyping Informatics teams for data handling; Le Centre National de Génotypage, France, led by M. Lathrop, for genotyping; Duke University, North Carolina, USA, led by D. Goldstein, for genotyping; and the Finnish Institute of Molecular Medicine, Finnish Genome Center, University of Helsinki, led by A. Palotie. **Icelandic deCODE Study (dCG):** We thank the staff of the deCODE core facilities and recruitment centre for their important contributions to this work. **Framingham Osteoporosis Study (FOS):** The study was funded by grants from the US National Institute for Arthritis, Musculoskeletal and Skin Diseases and National Institute on Aging (R01 AR/AG 41398; DPK and R01 AR 050066; DK). The Framingham Heart Study of the National Heart, Lung, and Blood Institute of the National Institutes of Health and Boston University School of Medicine were supported by the National Heart, Lung, and Blood Institute's Framingham Heart Study (N01-HC-25195) and its contract with Affymetrix, Inc. for genotyping services (N02-HL-6-4278). Analyses reflect intellectual input and resource development from the Framingham Heart Study investigators participating in the SNP Health Association Resource (SHARe) project. A portion of this research was conducted using the Linux Cluster for Genetic Analysis (LinGA-II) funded by the Robert Dawson Evans Endowment of the Department of Medicine at Boston University School of Medicine and Boston Medical Center. **eQTL Hob Study:** The study was supported by Genome Quebec, Genome Canada and the Canadian Institutes of Health Research (CIHR).

T.P. holds a Canada Research Chair. We thank O. Nilsson, H. Mallmin and Ö. Ljunggren at the Departments of Surgical and Medical Sciences, Uppsala University Hospital, Sweden, for large-scale collection of primary bone samples.

AUTHOR CONTRIBUTIONS

F.R., K.E., B.V.H., F.K.K. and J.P.A.I. ran meta-analysis; F.R., K.E., B.V.H., Y.-H.H., J.B.R., M.C.Z., N.A., Y.S.A., L.A.C., S.D., N.S., G.T. and Y.Z. ran statistical analysis in studies; F.R., U.S., P.D., J.B.v.M., U.T. and A.G.U. coordinated GWA genotyping of studies; E.G. and T.P. did expression studies; F.R., U.S., M.C.Z., A.H., B.O., H.A.P.P., G.S., G.T., F.M.K.W., S.G.W., C.M.v.D., T.S., D.P.K. and A.G.U. coordinated/collected phenotypic information; U.S., L.A.C., A.H., A.K., D.K., B.O., H.A.P.P., U.T., C.M.v.D., T.S., D.P.K., K.S. and A.G.U. designed studies; F.R., U.S., J.B.v.M., T.S., U.T., S.H.R., J.P.A.I. and A.G.U. established the consortium and U.T., S.H.R., J.P.A.I. and A.G.U. obtained funding; all authors interpreted results; all authors critically read the manuscript; and F.R. wrote the manuscript draft.

COMPETING INTERESTS STATEMENT

The authors declare competing financial interests: details accompany the full-text HTML version of the paper at <http://www.nature.com/naturegenetics/>.

Published online at <http://www.nature.com/naturegenetics/>.

Reprints and permissions information is available online at <http://npg.nature.com/reprintsandpermissions/>.

- Ellies, D.L. *et al.* Bone density ligand, Sclerostin, directly interacts with LRP5 but not LRP5G171V to modulate Wnt activity. *J. Bone Miner. Res.* **21**, 1738–1749 (2006).
- Theoleyre, S. *et al.* The molecular triad Opg/RANK/RANKL: involvement in the orchestration of pathophysiological bone remodeling. *Cytokine Growth Factor Rev.* **15**, 457–475 (2004).
- Ioannidis, J.P. *et al.* Meta-analysis of genome-wide scans provides evidence for sex- and site-specific regulation of bone mass. *J. Bone Miner. Res.* **22**, 173–183 (2007).
- Ioannidis, J.P. Why most published research findings are false. *PLoS Med.* **2**, e124 (2005).
- Ioannidis, J.P. Genetic associations: false or true? *Trends Mol. Med.* **9**, 135–138 (2003).
- Ioannidis, J.P. *et al.* Differential genetic effects of *ESR1* gene polymorphisms on osteoporosis outcomes. *J. Am. Med. Assoc.* **292**, 2105–2114 (2004).
- Langdahl, B.L. *et al.* Large-scale analysis of association between polymorphisms in the transforming growth factor beta 1 gene (*TGFB1*) and osteoporosis: the GENOMOS study. *Bone* **42**, 969–981 (2008).
- Ralston, S.H. *et al.* Large-scale evidence for the effect of the *COL1A1* Sp1 polymorphism on osteoporosis outcomes: the GENOMOS study. *PLoS Med.* **3**, e90 (2006).
- Uitterlinden, A.G. *et al.* The association between common vitamin D receptor gene variations and osteoporosis: a participant-level meta-analysis. *Ann. Intern. Med.* **145**, 255–264 (2006).
- van Meurs, J.B. *et al.* Large-scale analysis of association between *LRP5* and *LRP6* variants and osteoporosis. *J. Am. Med. Assoc.* **299**, 1277–1290 (2008).
- McCarthy, M.I. *et al.* Genome-wide association studies for complex traits: consensus, uncertainty and challenges. *Nat. Rev. Genet.* **9**, 356–369 (2008).
- Styrkarsdottir, U. *et al.* New sequence variants associated with bone mineral density. *Nat. Genet.* **41**, 15–17 (2009).
- Styrkarsdottir, U. *et al.* Multiple genetic loci for bone mineral density and fractures. *N. Engl. J. Med.* **358**, 2355–2365 (2008).
- Kanis, J.A., Delmas, P., Burckhardt, P., Cooper, C. & Torgerson, D. Guidelines for diagnosis and management of osteoporosis. The European Foundation for Osteoporosis and Bone Disease. *Osteoporosis. Int.* **7**, 390–406 (1997).
- Blake, G.M., Knapp, K.M., Spector, T.D. & Fogelman, I. Predicting the risk of fracture at any site in the skeleton: are all bone mineral density measurement sites equally effective? *Calcif. Tissue Int.* **78**, 9–17 (2006).
- Peacock, M., Turner, C.H., Econs, M.J. & Foroud, T. Genetics of osteoporosis. *Endocr. Rev.* **23**, 303–326 (2002).
- Weedon, M.N. *et al.* Genome-wide association analysis identifies 20 loci that influence adult height. *Nat. Genet.* **40**, 575–583 (2008).
- Gudbjartsson, D.F. *et al.* Many sequence variants affecting diversity of adult human height. *Nat. Genet.* **40**, 609–615 (2008).
- Visscher, P.M. Sizing up human height variation. *Nat. Genet.* **40**, 489–490 (2008).
- Devlin, B. & Roeder, K. Genomic control for association studies. *Biometrics* **55**, 997–1004 (1999).
- Willer, C.J. *et al.* Six new loci associated with body mass index highlight a neuronal influence on body weight regulation. *Nat. Genet.* **41**, 25–34 (2009).
- Pe'er, I., Yelensky, R., Altshuler, D. & Daly, M.J. Estimation of the multiple testing burden for genomewide association studies of nearly all common variants. *Genet. Epidemiol.* **32**, 381–385 (2008).
- Frazer, K.A. *et al.* A second generation human haplotype map of over 3.1 million SNPs. *Nature* **449**, 851–861 (2007).

24. Richards, J.B. *et al.* Bone mineral density, osteoporosis, and osteoporotic fractures: a genome-wide association study. *Lancet* **371**, 1505–1512 (2008).
25. Bänziger, C. *et al.* Wntless, a conserved membrane protein dedicated to the secretion of Wnt proteins from signaling cells. *Cell* **125**, 509–522 (2006).
26. Matsuda, A. *et al.* Large-scale identification and characterization of human genes that activate NF- κ B and MAPK signaling pathways. *Oncogene* **22**, 3307–3318 (2003).
27. Glass, D.A. II *et al.* Canonical Wnt signaling in differentiated osteoblasts controls osteoclast differentiation. *Dev. Cell* **8**, 751–764 (2005).
28. Pothoff, M.J. *et al.* Histone deacetylase degradation and MEF2 activation promote the formation of slow-twitch myofibers. *J. Clin. Invest.* **117**, 2459–2467 (2007).
29. Cho, Y.S. *et al.* A large-scale genome-wide association study of Asian populations uncovers genetic factors influencing eight quantitative traits. *Nat. Genet.* **41**, 527–534 (2009).
30. Smits, P. *et al.* The transcription factors L-Sox5 and Sox6 are essential for cartilage formation. *Dev. Cell* **1**, 277–290 (2001).
31. Zhou, G. *et al.* Dominance of SOX9 function over RUNX2 during skeletogenesis. *Proc. Natl. Acad. Sci. USA* **103**, 19004–19009 (2006).
32. Reiner, O. *et al.* The evolving doublecortin (DCX) superfamily. *BMC Genomics* **7**, 188 (2006).
33. Kim, S.H. *et al.* The forkhead transcription factor Foxc2 stimulates osteoblast differentiation. *Biochem. Biophys. Res. Commun.* **386**, 532–536 (2009).
34. Nifuji, A., Miura, N., Kato, N., Kellermann, O. & Noda, M. Bone morphogenetic protein regulation of forkhead/winged helix transcription factor Foxc2 (Mfh1) in a murine mesodermal cell line C1 and in skeletal precursor cells. *J. Bone Miner. Res.* **16**, 1765–1771 (2001).
35. Winnier, G.E., Hargett, L. & Hogan, B.L. The winged helix transcription factor MFH1 is required for proliferation and patterning of paraxial mesoderm in the mouse embryo. *Genes Dev.* **11**, 926–940 (1997).
36. Stankiewicz, P. *et al.* Genomic and genic deletions of the FOX gene cluster on 16q24.1 and inactivating mutations of *FOXF1* cause alveolar capillary dysplasia and other malformations. *Am. J. Hum. Genet.* **84**, 780–791 (2009).
37. Tang, Y. *et al.* Disruption of transforming growth factor- β signaling in ELF- β -spectrin-deficient mice. *Science* **299**, 574–577 (2003).
38. Alford, A.I. & Hankenson, K.D. Matricellular proteins: extracellular modulators of bone development, remodeling, and regeneration. *Bone* **38**, 749–757 (2006).
39. Gowen, L.C. *et al.* Targeted disruption of the osteoblast/osteocyte factor 45 gene (OF45) results in increased bone formation and bone mass. *J. Biol. Chem.* **278**, 1998–2007 (2003).
40. Malaval, L. *et al.* Bone sialoprotein plays a functional role in bone formation and osteoclastogenesis. *J. Exp. Med.* **205**, 1145–1153 (2008).
41. Yoshitake, H., Rittling, S.R., Denhardt, D.T. & Noda, M. Osteopontin-deficient mice are resistant to ovariectomy-induced bone resorption. *Proc. Natl. Acad. Sci. USA* **96**, 8156–8160 (1999).
42. Meyers, V.E., Zayzafoon, M., Douglas, J.T. & McDonald, J.M. RhoA and cytoskeletal disruption mediate reduced osteoblastogenesis and enhanced adipogenesis of human mesenchymal stem cells in modeled microgravity. *J. Bone Miner. Res.* **20**, 1858–1866 (2005).
43. Wang, L., Yang, L., Debidia, M., Witte, D. & Zheng, Y. Cdc42 GTPase-activating protein deficiency promotes genomic instability and premature aging-like phenotypes. *Proc. Natl. Acad. Sci. USA* **104**, 1248–1253 (2007).
44. McKinsey, T.A., Zhang, C.L., Lu, J. & Olson, E.N. Signal-dependent nuclear export of a histone deacetylase regulates muscle differentiation. *Nature* **408**, 106–111 (2000).
45. Schroeder, T.M. & Westendorf, J.J. Histone deacetylase inhibitors promote osteoblast maturation. *J. Bone Miner. Res.* **20**, 2254–2263 (2005).
46. Kang, J.S., Alliston, T., Delston, R. & Derynck, R. Repression of Runx2 function by TGF- β through recruitment of class II histone deacetylases by Smad3. *EMBO J.* **24**, 2543–2555 (2005).
47. Ioannidis, J.P., Thomas, G. & Daly, M.J. Validating, augmenting and refining genome-wide association signals. *Nat. Rev. Genet.* **10**, 318–329 (2009).
48. Online Mendelian Inheritance in Man (OMIM) database (McKusick-Nathans Institute of Genetic Medicine, Johns Hopkins University, and National Center for Biotechnology Information, National Library of Medicine, Bethesda, Maryland, USA). <<http://www.ncbi.nlm.nih.gov/omim/>>.
49. McKusick, V.A. *Mendelian Inheritance in Man. A Catalog of Human Genes and Genetic Disorders* (Johns Hopkins University Press, Baltimore, Maryland, USA, 1998).

ONLINE METHODS

Study population. The GENetic Factors for Osteoporosis (GEFOS) Consortium is a coalition of teams of investigators working on the genetics of osteoporosis. The current meta-analysis incorporated 19,195 individuals of Northern European ancestry derived from five GWAS on BMD of the lumbar spine (LS-BMD) and the femoral neck (FN-BMD) including the Rotterdam Study (RS, $n = 4,987$), Erasmus Rucphen Family Study (ERF, $n = 1,228$), Twins UK Study (TUK, $n = 2,734$), deCODE Genetics Study (dCG, $n = 6,743$) and Framingham Osteoporosis Study (FOS, $n = 3,503$). All studies were approved by institutional ethics review committees at the relevant organizations and all participants provided written informed consent. The Rotterdam Study is a prospective population-based cohort study of chronic disabling conditions in Dutch elderly individuals aged 55 years and over (<http://www.epib.nl/research/ergo.htm>)^{50,51}. The Erasmus Rucphen Family study is a family-based study of a genetic isolate in the southwest Netherlands to identify genetic risk factors for complex disorders⁵². The Twins UK study is a population-based sample of twins from the UK studying the hereditary basis of a variety of age-related traits and diseases (<http://www.twinsUK.ac.uk>)⁵³. The Icelandic deCODE Genetics study uses a population-based sample to identify the genetic basis of complex diseases¹³. The Framingham Osteoporosis Study is embedded in the Framingham Heart Study (FHS), a community-based, longitudinal, prospective cohort comprising three generations of individuals in multigenerational pedigrees and additional unrelated individuals (<http://www.framinghamheartstudy.org/>). Individuals of 'Generation 1' include those first examined in 1948 (ref. 54), 'Generation 2' includes those examined at the first cycle from 1971 to 1975 (ref. 55) and 'Generation 3' includes those examined at the first cycle beginning in 2002–2005 (ref. 56). For these analyses, 812 members of the Generation 1 cohort (22.6% of the sample) and 2,783 (77.4%) members of the Generation 2 cohort who had BMD measured as part of the FOS were included.

BMD and anthropometric measurements. BMD was measured in all cohorts at the lumbar spine (L1–L4 or L2–L4) and femoral neck using dual-energy X-ray absorptiometry following standard manufacturer protocols (GE-Lunar Corp. or Hologic Inc.). See **Supplementary Table 1** for details. All DXA and anthropometric measurements were performed in the RS at the baseline visit between 1991 and 1992, in ERF between 2002 and 2003, in TUK at the latest follow-up, in dCG at the baseline visit, and in FOS Generation 1 between 1992 and 1997 and Generation 2 between 1996 and 2001.

Phenotype modeling. The overall strategy involved linear regression to adjust BMD measurements for effects of age, weight, sex and study using standardized residuals with mean 0 and s.d. 1 in the genotype-phenotype association testing. Such residuals were obtained by regressing within each study the raw BMD measurements on age and weight (and principal components in FOS to adjust for population substructure using EIGENSTRAT⁵⁷) in sex-specific models. Thus, in studies including both men and women, the data for each gender are included as separate estimates in the meta-analysis.

Genotyping. The five GWAS were genotyped using the Illumina Infinium HumanHap550 Beadchip (RS), the Illumina Infinium HumanHap300 or HumanCNV370 Beadchip (ERF, TUK and dCG) or the Affymetrix Dual NspI/StyI GeneChip 2 × 250K with 50K gene-centered MIP set (FOS), all according to manufacturers' protocols and quality control standards. The exclusion and filtering criteria for individuals and SNPs are described in **Supplementary Tables 2 and 3**.

Genotype imputation. Imputation was used to evaluate associations for the same SNPs across study populations using scans from different genotyping platforms. Genotypes were imputed for all polymorphic SNPs oriented to the positive strand from phased autosomal chromosomes of the HapMap CEU Phase II panel (release 22, build 36)⁵⁸. Hidden Markov model-based

algorithms were used to infer unobserved genotypes probabilistically as implemented in either MACH⁵⁹ or IMPUTE⁶⁰. Imputation quality control metrics from MACH and IMPUTE were used. Detailed descriptions of quality control and imputation procedures are summarized for all studies in **Supplementary Table 3**. We performed technical validation of the imputed genotypes in an independent set of 880 individuals of Icelandic origin using Centaurus (Nanogen)⁶¹ discrimination assays for the top associated hits that reached GWS for the first time in this study (**Supplementary Table 10**).

Genotype-phenotype association testing. Each study performed genome-wide association for BMD using sex-specific, age- and weight-adjusted standardized residuals analyzed under an additive (per allele) genetic model⁶². Analysis of imputed genotype data accounted for uncertainty in each genotype prediction by using either the dosage information from MACH or the genotype probabilities from IMPUTE. Studies used MACH2QTL⁵⁹, which uses genotype dosage value (0–2, continuous) as a predictor in a linear regression framework, SNPTEST⁶⁰, Merlin⁶³ or the linear mixed effects model of the Kinship⁶⁴ and ProABEL⁶⁵ packages in R⁶⁶ to account for relatedness (**Supplementary Table 3**). The genomic control method²⁰ was used to correct the standard error by the square root of the genomic inflation factor (λ): $SE_{corrected} = s.e.m. \sqrt{\lambda}$, which is equivalent to the proposed correction of the χ^2 statistic by λ . Genomic inflation factors for the studies are presented in **Supplementary Table 3**. Overall meta-analysis genomic control inflation factors were calculated as described⁶⁷. Genomic inflation factors scaled to a standard size (1,000 individuals) to calibrate for the effect of sample size on λ ⁶⁷ showed that residual genomic inflation was negligible ($\lambda_{LSBMD1000} = 1.005$ and $\lambda_{FN-BMD1000} = 1.004$).

Meta-analysis. The minor allele from HapMap CEU genotypes was used to define the coded allele in all analyses, regardless of frequency in individual cohorts. All meta-analysis calculations were done using the METAL²¹ software package applying inverse-variance methodology assuming fixed effects with Cochran's Q and I^2 metrics used to quantify between-study heterogeneity. We also calculated the summary results by random effects using STATA software⁶⁸ for those markers associated at GWS level with Cochran's Q P value < 0.05 and/or I^2 estimates $> 50\%$, reflecting large heterogeneity beyond chance. Random effects models also incorporate in the calculations the between-study heterogeneity and estimate the average genetic effect from the population of genetic effects that may differ in different studies. In the absence of between-study heterogeneity, fixed and random effects calculations give identical results. We declared results GWS at $\alpha = 5 \times 10^{-8}$ after adjusting for all common variant tests in the human genome^{22,23}. To test for BMD site specificity, we estimated the effect difference ($\Delta\beta$) as $\beta_{FN} - \beta_{LS}$ (where FN is femoral neck and LS is lumbar spine), the s.e.m. of the mean difference ($\Delta s.e.m.$) as $\sqrt{SE_{FN}^2 + SE_{LS}^2}$ and the Z statistic as $\Delta\beta/\Delta s.e.m.$ from which the P values were computed.

eQTL analysis in human osteoblasts. SNPs from new loci associated with BMD at the genome-wide significance (GWS) level were tested for association with *cis*-allelic expression of neighboring gene transcripts, in primary human osteoblasts derived from 95 unrelated Swedish donors. Detailed cell culture methods have been described⁶⁹. Expression profiling was performed using Illumina HumRef-8 BeadChips according to the manufacturer's protocol. Genotyping for genotype-expression association was performed using the Illumina HumanHap 550k Duo chip. Individuals with low genotyping rate and SNPs showing significant deviation from Hardy-Weinberg equilibrium ($P < 0.05$) were excluded. Similarly, low-frequency ($MAF < 0.05$) SNPs and SNPs with high rates of missing data were excluded. The association of the expression levels was focused on *cis*-acting genetic variants, defined as being within a 250-kb window flanking the gene, using a linear regression model implemented in the PLINK⁷⁰ software package with age and sex as covariates.

SNPs included on the Illumina 550K chip were assessed for expression *cis* associations directly. In addition, all genotyped SNPs included on the Illumina chip that were in strong LD (defined as $D' \geq 0.8$) and mapping ± 50 kb from the GWS hit were included in the association study. To test for significant enrichment of functional SNPs (those associated with gene expression in human osteoblasts at $P < 0.05$) among the candidate SNPs, a χ^2 statistic was obtained to test whether observed associations were different from expected associations in the expression dataset beyond chance. Expected values (7.1%) were based on the proportion of SNPs with $P < 0.05$ seen in the human osteoblast gene expression dataset and in a random selection of 1,200 SNPs associated with both LS-BMD and FN-BMD at $P > 0.90$, $MAF > 0.20$, and present in the Illumina HumanHap550 array (assessed as negative controls for association with human osteoblast expression).

Combined effect of associated loci. Within the setting of the prospective population-based Rotterdam Study, the combined effect of all 20 BMD loci was studied by classifying subjects according to the number of BMD-decreasing (risk) alleles. This was based on 15 lumbar spine and 10 femoral neck BMD loci as follows. Lumbar spine (15 SNPs): rs7524102 (*ZBTB40*), rs1430742 (*GPR177*), rs11898505 (*SPTBN1*), rs1471403 (*MEPE*), rs2504063 (*ESR1*), rs1524058 (*STARD3NL*), rs4729260 (*FLJ42280*), rs2062377 (*TNFRSF11B*), rs16921914 (*DCDC5*), rs599083 (*LRP5*), rs2016266 (*SP7*), rs9533090 (*TNFSF11*), rs10048146 (*FOXC2*), rs9303521 (*CRHR1*) and rs884205 (*TNFRSF11A*). Femoral neck (10 SNPs): rs6426749 (*ZBTB40*), rs2566755 (*GPR177*), rs87938 (*CTNNA1*), rs1366594 (*MEF2C*), rs2941740 (*ESR1*), rs7781370 (*FLJ42280*), rs11995824 (*TNFRSF11B*), rs7117858 (*SOX6*), rs7932354 (*ARHGAP1*) and rs228769 (*HDAC5*). The mean BMD for each risk allele-count group was determined, and at the extremes of the distribution counts were pooled into the nearest risk allele-count group of >100 individuals. The approximate BMD difference in g/cm^2 was obtained by multiplying in each group the mean Z -score LS-BMD by 0.18 g/cm^2 and the mean Z -score FN-BMD by 0.13 g/cm^2 (the s.d. values for BMD in the Rotterdam Study). The allele score was obtained by dividing the number of 'BMD-decreasing alleles' by the total number of alleles. Also within the setting of the prospective population-based Rotterdam Study, we determined the risk for vertebral and nonvertebral fracture for the combined allelic scores constructed for all the top hits associated at a GWS level with BMD. Risk ratio (RR) estimates were obtained from logistic regression (vertebral fractures) and Cox-proportional hazards (incident nonvertebral fractures) models adjusted for sex, age and weight. To determine the fraction of fracture risk explained by BMD, we applied the following formula: $[\ln RR_{unadjusted} - \ln RR_{BMDadjusted}] / \ln RR_{unadjusted}$. Methods describing the fracture datasets have been published^{71,72}. In summary, thoracolumbar radiographs of the spine were obtained in 3,308 (genotyped) individuals who survived on average 6.4 ± 0.4 (s.d.) years and were scored for presence of vertebral fractures ($n = 329$) using the McCloskey/Kanis method⁷³. Record of the incidental nonvertebral fractures ($n = 900$) occurring between the baseline visit from 1990 through 1993 until 1 January 2002, was obtained from

the computerized records of general practitioners and hospital registries for 5,974 genotyped individuals followed on average 8.2 ± 2.7 (s.d.) years after the baseline visit.

50. Hofman, A. *et al.* The Rotterdam Study: 2010 objectives and design update. *Eur. J. Epidemiol.* **24**, 553–572 (2009).
51. Hofman, A., Grobbee, D.E., de Jong, P.T. & van den Ouweland, F.A. Determinants of disease and disability in the elderly: the Rotterdam Elderly Study. *Eur. J. Epidemiol.* **7**, 403–422 (1991).
52. Aulchenko, Y.S. *et al.* Linkage disequilibrium in young genetically isolated Dutch population. *Eur. J. Hum. Genet.* **12**, 527–534 (2004).
53. Arden, N.K., Baker, J., Hogg, C., Baan, K. & Spector, T.D. The heritability of bone mineral density, ultrasound of the calcaneus and hip axis length: a study of postmenopausal twins. *J. Bone Miner. Res.* **11**, 530–534 (1996).
54. Dawber, T.R., Kannel, W.B. & Lyell, L.P. An approach to longitudinal studies in a community: the Framingham Study. *Ann. NY Acad. Sci.* **107**, 539–556 (1963).
55. Kannel, W.B., Feinleib, M., McNamara, P.M., Garrison, R.J. & Castelli, W.P. An investigation of coronary heart disease in families. The Framingham offspring study. *Am. J. Epidemiol.* **110**, 281–290 (1979).
56. Splansky, G.L. *et al.* The Third Generation Cohort of the National Heart, Lung, and Blood Institute's Framingham Heart Study: design, recruitment, and initial examination. *Am. J. Epidemiol.* **165**, 1328–1335 (2007).
57. Price, A.L. *et al.* Principal components analysis corrects for stratification in genome-wide association studies. *Nat. Genet.* **38**, 904–909 (2006).
58. International HapMap Consortium. A second generation human haplotype map of over 3.1 million SNPs. *Nature* **449**, 851–861 (2007).
59. Li, Y. & Abecasis, G.R. Mach 1.0: rapid haplotype reconstruction and missing genotype inference. *Am. J. Hum. Genet.* **S79**, 2290 (2006).
60. 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.* **39**, 906–913 (2007).
61. Kutayvin, I.V. *et al.* A novel endonuclease IV post-PCR genotyping system. *Nucleic Acids Res.* **34**, e128 (2006).
62. Little, J. *et al.* Strengthening the reporting of genetic association studies (STREGA): an extension of the STROBE statement. *Eur. J. Epidemiol.* **24**, 37–55 (2009).
63. Abecasis, G.R., Cherny, S.S., Cookson, W.O. & Cardon, L.R. Merlin—rapid analysis of dense genetic maps using sparse gene flow trees. *Nat. Genet.* **30**, 97–101 (2002).
64. Therneau, T. Kinship: mixed effects Cox models, sparse matrices, and modelling data from large pedigrees. R package version 1.1.0, edn. 19 (R Foundation for Statistical Computing, Vienna, Austria, 2008).
65. Aulchenko, Y.S., Ripke, S., Isaacs, A. & van Duijn, C.M. GenABEL: an R library for genome-wide association analysis. *Bioinformatics* **23**, 1294–1296 (2007).
66. R Developmental Core Team. *A Language and Environment for Statistical Computing* (R Foundation for Statistical Computing, Vienna, Austria, 2007).
67. de Bakker, P.I. *et al.* Practical aspects of imputation-driven meta-analysis of genome-wide association studies. *Hum. Mol. Genet.* **17**, R122–R128 (2008).
68. Stata Statistical Software. Release 10 (StataCorp LP, College Station, Texas, USA, 2007).
69. Grundberg, E. *et al.* Systematic assessment of the human osteoblast transcriptome in resting and induced primary cells. *Physiol. Genomics* **33**, 301–311 (2008).
70. Purcell, S. *et al.* PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am. J. Hum. Genet.* **81**, 559–575 (2007).
71. Rivadeneira, F. *et al.* Estrogen receptor beta (*ESR2*) polymorphisms in interaction with estrogen receptor alpha (*ESR1*) and insulin-like growth factor 1 (*IGF1*) variants influence the risk of fracture in postmenopausal women. *J. Bone Miner. Res.* **21**, 1443–1456 (2006).
72. Schuit, S.C. *et al.* Fracture incidence and association with bone mineral density in elderly men and women: the Rotterdam Study. *Bone* **34**, 195–202 (2004).
73. McCloskey, E.V. *et al.* The assessment of vertebral deformity: a method for use in population studies and clinical trials. *Osteoporos. Int.* **3**, 138–147 (1993).

Erratum: Twenty bone-mineral-density loci identified by large-scale meta-analysis of genome-wide association studies

Fernando Rivadeneira, Unnur Styrkársdóttir, Karol Estrada, Bjarni V Halldórsson, Yi-Hsiang Hsu, J Brent Richards, M Carola Zillikens, Fotini K Kavvoura, Najaf Amin, Yuri S Aulchenko, L Adrienne Cupples, Panagiotis Deloukas, Serkalem Demissie, Elin Grundberg, Albert Hofman, Augustine Kong, David Karasik, Joyce B van Meurs, Ben Oostra, Tomi Pastinen, Huibert A P Pols, Gunnar Sigurdsson, Nicole Soranzo, Gudmar Thorleifsson, Unnur Thorsteinsdóttir, Frances M K Williams, Scott G Wilson, Yanhua Zhou, Stuart H Ralston, Cornelia M van Duijn, Timothy Spector, Douglas P Kiel, Kari Stefansson, John P A Ioannidis & André G Uitterlinden for the Genetic Factors for Osteoporosis (GEFOS) Consortium

Nat. Genet.; doi:10.1038/ng.446; corrected online 11 October 2009

In the version of this article initially published online, the seventh and eighth sentences under the heading “Combined effect of the 20 GWS BMD loci and fracture risk” were incorrect. The correct wording is as follows: “The compound FN-BMD allelic score was significantly associated with the risk of incident nonvertebral fracture in the Rotterdam Study dataset (HR = 1.042, 95% CI [1.003, 1.084]; P = 0.04), whereas it was borderline significant for association with the risk of vertebral fracture (OR = 1.061, 95% CI [0.997, 1.129]; P = 0.06). In contrast, the compound LS-BMD allelic score was significantly associated with the risk of vertebral fracture (OR = 1.061, 95% CI [1.009, 1.116]; P = 0.02), whereas it was not significant for association with the risk of incident nonvertebral fracture (HR = 1.025, 95% CI [0.993, 1.058]; P = 0.13).” The error has been corrected for all versions of the article.