

Genome-wide association study meta-analysis for quantitative ultrasound parameters of bone identifies five novel loci for broadband ultrasound attenuation

Benjamin H. Mullin^{1,2*}, Jing Hua Zhao³, Suzanne J. Brown¹, John R. B. Perry³, Jian'an Luan³, Hou-Feng Zheng⁴, Claudia Langenberg³, Frank Dudbridge⁵, Robert Scott³, Nick J. Wareham³, Tim D. Spector⁶, J Brent Richards^{6,7†}, John P. Walsh^{1,2†}, Scott G. Wilson^{1,2,6†}

¹Dept. of Endocrinology & Diabetes, Sir Charles Gairdner Hospital, Perth 6009, Australia,

²School of Medicine and Pharmacology, University of Western Australia, Perth 6009,

Australia, ³MRC Epidemiology Unit, University of Cambridge, Cambridge CB2 0QQ, UK,

⁴Institute of Aging Research, School of Medicine, Hangzhou Normal University, and the Affiliated Hospital of Hangzhou Normal University, Hangzhou, Zhejiang 310036, China,

⁵Department of Non-communicable Disease Epidemiology, London School of Hygiene and Tropical Medicine, London WC1E 7HT, UK, ⁶Twin & Genetic Epidemiology Research Unit,

St Thomas' Hospital Campus, King's College London, London SE1 7EH, UK, ⁷Departments

of Medicine, Human Genetics, Epidemiology and Biostatistics, Jewish General Hospital, Lady Davis Institute, McGill University, Montreal H3T 1E2, Canada

[†]Equal contribution

*Corresponding author: Benjamin H. Mullin
School of Medicine and Pharmacology
The University of Western Australia
Nedlands, Western Australia 6009
Phone: +61 8 6457 2466
Fax: +61 8 6457 3221
Email: Benjamin.Mullin@uwa.edu.au

Abstract

Osteoporosis is a common and debilitating bone disease that is characterised by low bone mineral density, typically assessed using dual-energy X-ray absorptiometry. Quantitative ultrasound (QUS), commonly utilising the two parameters velocity of sound (VOS) and broadband ultrasound attenuation (BUA), is an alternative technology used to assess bone properties at peripheral skeletal sites. The genetic influence on the bone qualities assessed by QUS remains an under-studied area. We performed a comprehensive GWAS including low-frequency variants ($MAF \geq 0.005$) for BUA and VOS using a discovery population of individuals with whole-genome sequence (WGS) data from the UK10K project ($n=1,268$). These results were then meta-analysed with those from two deeply imputed GWAS replication cohorts ($n=1,610$ and $13,749$). In the gender-combined analysis, we identified eight loci associated with BUA and five with VOS at the genome-wide significance level, including three novel loci for BUA at 8p23.1 (*PPP1R3B*), 11q23.1 (*LOC387810*) and 22q11.21 (*SEPT5*) ($P=2.4 \times 10^{-8}$ – 1.6×10^{-9}). Gene-based association testing in the gender-combined dataset revealed eight loci associated with BUA and seven with VOS at the genome-wide significance level, with one novel locus for BUA at *FAM167A* (8p23.1) ($P=1.4 \times 10^{-6}$). An additional novel locus for BUA was seen in the male-specific analysis at *DEFB103B* (8p23.1) ($P=1.8 \times 10^{-6}$). Fracture analysis revealed significant associations between variation at the *WNT16* and *RSPO3* loci and fracture risk ($P=0.004$ and 4.0×10^{-4} respectively). In conclusion, by performing a large GWAS meta-analysis for QUS parameters of bone using a combination of WGS and deeply imputed genotype data, we have identified five novel genetic loci associated with BUA.

Introduction

Osteoporosis is a common and debilitating bone disease that is characterised by a low bone mineral density (BMD) and deterioration of the bone micro-architecture, which leads to an increased risk of fracture (1). The disease is particularly prevalent in postmenopausal women due to a reduction in oestrogen production, with subsequent effects on bone as well as intestinal and renal calcium handling (2,3). Excess mortality caused by osteoporotic fracture in women has been estimated 1-year post fracture at 9% and 5-years post fracture at 24% (4). In addition to the effects of oestrogen, calcium and other environmental factors on bone structure, there is a strong genetic effect on peak bone mass (attained in early adult life), bone loss and fracture rates (5,6). Individuals with an affected first-degree relative have a considerably elevated estimated familial relative risk of fragility fracture of 1.31-4.24 (7,8).

Dual-energy X-ray absorptiometry (DXA) is the most commonly used method for measurement of BMD *in vivo* and diagnosis of osteoporosis. This involves measuring the attenuation of a photon beam by hydroxyapatite, the primary mineral in bone. However, quantitative ultrasound (QUS) is an alternative technology that is used to assess bone properties at peripheral skeletal sites. It has the advantages of being more portable and less expensive than DXA without the use of ionising radiation. Characterisation of bone tissue using QUS commonly utilises the two parameters broadband ultrasound attenuation (BUA) and velocity of sound (VOS). QUS is capable of assessing aspects of the number and principal structural orientation of trabecular elements in bone, which is highly correlated with the mechanical strength of trabecular bone (9). For BUA this is well illustrated by anisotropy (the property of being directionally dependent), wherein the BUA determination is sensitive to the mechanical properties of trabecular bone and the measurement varies according to the orientation; and therefore provides a comprehensive integration of bone “quality” and “quantity”. VOS refers to the division of transmission time of the sound waves by the length

of the body part studied (10) and is thought to provide an integrated assessment of BMD and the elasticity of the bone (11). Studies in cadavers have further confirmed that QUS parameters are significantly correlated with both BMD and micro-architectural parameters of the bone (12). Assessment of bone properties at the calcaneus using QUS has been shown to be a comparable predictor of fracture risk when compared to DXA BMD (13,14) and appears to do so independently (13-15). This suggests that while both techniques assess bone mass, they also capture different properties of the bone tissue to some extent. Heritability estimates for BUA and VOS range from 0.52-0.74 and 0.19-0.61 respectively (16-20).

There have been many previous genome-wide association studies (GWAS) for bone phenotypes, with the vast majority of these focused on DXA BMD measured at the lumbar spine, femoral neck, total hip and forearm sites. Estrada et al. (21) published the results from a large GWAS meta-analysis for BMD conducted in 32,961 individuals, with replication studies in a further 50,933 subjects. They identified 56 BMD loci (32 novel), of which 14 were also associated with fracture risk. A similarly powered recent GWAS meta-analysis including low-frequency and rare variants published by Zheng et al. (22) identified a novel BMD locus, *ENI*, containing multiple low-frequency variants with large effect sizes. In total, at least 76 loci have been reported as associated with BMD and QUS parameters at a high level of confidence (NHGRI GWAS Catalogue (23)). However, the genetic influence on the bone qualities assessed by QUS remains under-studied as a proportion of these. Moayyeri et al. (24) reported a GWAS meta-analysis that successfully identified nine genetic loci associated with QUS parameters at the genome-wide significance level. That study utilised genotype data imputed using the HapMap Phase II reference panel (~2.5 million variants) (24), however this marker coverage is relatively low compared to state-of-the-art protocols such as imputation using the 1000 Genomes Project (1000GP) reference panels which include >4 fold more variants. Therefore, we decided to perform a comprehensive GWAS for the

QUS phenotypes BUA and VOS, spanning the allelic frequency spectrum and including low-frequency variants (minor allele frequency (MAF) ≥ 0.005). For this we used a discovery population comprised of a subset of individuals with whole-genome sequence (WGS) data from the UK10K project (TwinsUK WGS cohort $n=1,268$), with subsequent analyses in two deeply imputed GWAS replication cohorts (TwinsUK GWAS cohort $n=1,610$ and European Prospective Investigation into Cancer and Nutrition study (EPIC) cohort $n=13,749$), and finally with the results meta-analysed in a combined total of 16,627 individuals.

Results

Descriptive statistics for the three cohorts are presented in the Supplementary Material, Table S1. Analysis of the phenotype data in the TwinsUK WGS discovery cohort demonstrated a degree of correlation between BUA and VOS ($\rho=0.52$). Quantile-quantile plots generated for the gender-combined and gender-stratified meta-analyses are shown in the Supplementary Material, Fig. S1 (all λ values <1.05). A total of 9,844,025 variants with a MAF ≥ 0.005 were analysed in the TwinsUK WGS discovery cohort (Table S1). Nineteen variants from the *WNT16* gene region on chromosome 7q31.31 demonstrated genome-wide significant associations ($P < 5.0 \times 10^{-8}$) with BUA in this cohort ($P = 4.9 \times 10^{-8} - 3.0 \times 10^{-9}$, Supplementary Material, Table S2). We analysed these nineteen variants in the combined replication cohorts and found them all to be associated with BUA at $P \leq 3.0 \times 10^{-5}$ ($P = 3.0 \times 10^{-5} - 2.9 \times 10^{-45}$, Supplementary Material, Table S2). For VOS, no genome-wide significant associations were seen in the discovery cohort. We then meta-analysed the results from all three cohorts for BUA and VOS in order to fully exploit the discovery potential of the study (see Fig. 1 for description of study design).

Gender-combined meta-analyses

A total of 11,973,598 variants with a MAF ≥ 0.005 were included in the meta-analysis of all three cohorts. We identified eight genetic loci associated with QUS parameters at the genome-wide significance level in the gender-combined meta-analysis. Five of these loci were significant in both the BUA and VOS analyses and all have a known role in bone biology: 2p16.2 (*SPTBN1*), 6q22.33 (*RSPO3*), 6q25.1 (*CCDC170/ESR1*), 7q31.31 (*WNT16*) and 11q14.2 (*TMEM135*) ($P = 2.0 \times 10^{-8} - 1.2 \times 10^{-51}$, Fig. 2a, Fig. 2b and Table 1). Three novel loci were identified for BUA in this analysis, including 8p23.1 (*PPP1R3B*), 11q23.1 (*LOC387810*) and 22q11.21 (*SEPT5*) ($P = 2.4 \times 10^{-8} - 1.6 \times 10^{-9}$, Fig. 2a, Fig. 3 and Table 1).

Analysis of the genome-wide significant variants for BUA and VOS suggested that 3.8% (95% CI: 2.9-4.7%) and 2.9% (95% CI: 2.1-3.8%) of the phenotype variance was attributable to these variants respectively. Using the less stringent genome-wide suggestive association threshold of $P < 5.0 \times 10^{-7}$, a further seven loci were found to be associated with BUA and/or VOS ($P = 3.9 \times 10^{-7} - 8.2 \times 10^{-8}$, Supplementary Material, Table S3).

Gender-stratified analyses

Females – A total of 10,032 individuals were included in the female-specific meta-analysis of the three cohorts. Four genetic loci were found to be associated with QUS parameters at the genome-wide significance level, including 2p16.2 (*SPTBN1*, BUA), 7q31.31 (*WNT16*, BUA and VOS), 11q14.2 (*TMEM135*, VOS) and 11q23.1 (*LOC387810*, BUA) ($P = 1.4 \times 10^{-8} - 1.6 \times 10^{-29}$, Fig. 2c, Fig. 2d and Table 2). Of these, the findings at 11q23.1 (*LOC387810*) for BUA are novel and were also seen in the gender-combined analysis for this phenotype. An additional seven loci were associated with BUA or VOS at the suggestive level ($P = 4.3 \times 10^{-7} - 1.6 \times 10^{-7}$, Supplementary Material, Table S4).

Males – The TwinsUK WGS discovery cohort is entirely female, so the male-specific meta-analysis only included individuals from the TwinsUK GWAS and EPIC replication cohorts (combined total 6,411 individuals). The two loci 6q25.1 (*CCDC170/ESR1*) and 7q31.31 (*WNT16*) were identified as associated with both BUA and VOS at the genome-wide significance level in this analysis ($P = 3.0 \times 10^{-8} - 3.7 \times 10^{-27}$, Fig. 2c, Fig. 2d and Table 2), with an additional 11 loci associated with BUA or VOS at the suggestive level ($P = 4.8 \times 10^{-7} - 1.2 \times 10^{-7}$, Supplementary Material, Table S4).

Conditional analyses

Conditional analyses were performed on the gender-combined datasets in an effort to identify any independent association signals arising from the genome-wide significant loci. After

inspection of the linkage disequilibrium (LD) structure within each locus, conditional analyses were performed on the 6q25.1 (*CCDC170/ESR1*), 7q31.31 (*WNT16*) and 11q14.2 (*TMEM135*) loci. After conditioning the data on the lead variant from each locus, no secondary signals were identified at $P < 5 \times 10^{-8}$. However, after conditioning on the variants rs1891002 for BUA and rs4869739 for VOS at the 6q25.1 (*CCDC170/ESR1*) locus, suggestive secondary signals were identified led by rs1038304 (BUA, $P = 1.0 \times 10^{-7}$) and rs851984 (VOS, $P = 7.9 \times 10^{-8}$).

Fracture analysis

We analysed the lead variant from each significant locus in the gender-combined analysis for association with fracture rate. The variant pairs rs4869739/rs1891002 (6q25.1) and rs597319/rs511755 (11q14.2) are in strong LD ($r^2 = 0.99$ and 0.94 respectively), so only one variant from each pair was taken forward for analysis. Of the eight variants tested, rs7741021 (*RSPO3*) and rs2908007 (*WNT16*) were found to be significantly associated with fracture rate after correction for multiple testing ($P < 0.006$, Table 3). For rs7741021, each copy of the *C* allele was associated with +0.08 standard deviation for both BUA and VOS, and a 9% reduction in the odds of sustaining a fracture. Each copy of the *G* allele at rs2908007 was associated with +0.17 standard deviation for both BUA and VOS, and an 8% reduction in fracture odds. Another two variants, rs4869739 (*CCDC170/ESR1*) and rs597319 (*TMEM135*), were found to be nominally associated with fracture rate ($P < 0.05$, Table 3). In both of these cases, the allele associated with a reduced QUS measurement was associated with an increased fracture rate.

Bioinformatics analysis

We performed a bioinformatics analysis for the three novel association signals identified for BUA in the gender-combined analysis.

rs367543 (8p23.1) – This variant is in strong LD ($r^2 \geq 0.8$) with three other variants (Supplementary Material, Table S5). All four variants in this LD block are associated with expression of multiple transcripts from the region including *PPP1R3B*, *ERII*, *MFHAS1*, *SGK223*, *FAM86B3P*, *RP11-62H7.2*, *CTA-398F10.2* and *RP11-10A14.5* ($P=3.8 \times 10^{-5} - 5.5 \times 10^{-22}$) (25-27).

rs7952251 (11q23.1) – There are eight variants in strong LD with *rs7952251* (Supplementary Material, Table S5), seven of which are associated with the expression of the GENCODE transcript *RP11-356J5.12* ($P=6 \times 10^{-6} - 1.5 \times 10^{-6}$). The variant *rs7109000* presents strongly as potentially having a regulatory role, residing within an enhancer histone mark in 13 tissues and a DNase hypersensitivity site in 19 tissues.

rs202240051 (22q11.21) – Two variants are in strong LD with *rs202240051* (Supplementary Material, Table S5), both of which are associated with expression of the gene *SEPT5* in skin and adipose tissue ($P=1.3 \times 10^{-5} - 1.5 \times 10^{-10}$). Of these, *rs9606139* may have a regulatory role as it is located in a promoter histone mark in 14 tissues, enhancer histone mark in nine tissues and DNase hypersensitivity site in 28 tissues.

Gene-based association testing

Gender-combined analysis

Gene-based (+/- 50Kb) tests of association were performed on the GWAS results for BUA and VOS using the VEratile Gene-based Association Study 2 (VEGAS2) software (28), with the results from all three cohorts meta-analysed. Ten gene regions were identified as associated with QUS parameters at the genome-wide significance level in the gender-combined analysis (gene-based $P < 1.95 \times 10^{-6}$, Table 4), nine of which are located in and around known bone loci and include *RSPO3* (6q22.33, BUA), *CCDC170* (6q25.1, BUA and VOS), *ESR1* (6q25.1, VOS), *WNT16* (7q31.31, BUA and VOS), *FAM3C* (7q31.31, BUA and

VOS), *CPED1* (7q31.31, BUA and VOS), *TMEM135* (11q14.2, BUA), *GPATCH1* (19q13.11, BUA and VOS) and *WDR88* (19q13.11, VOS). One novel association was identified for BUA in this analysis with *FAM167A* (8p23.1).

Gender-stratified analysis

The female-specific analysis identified four gene regions as associated with BUA and/or VOS at the genome-wide significance level (Table 4). All of these associations are with genes located in known loci and include *WNT16* (7q31.31, BUA and VOS), *FAM3C* (7q31.31, BUA and VOS), *CPED1* (7q31.31, BUA and VOS) and *WDR88* (19q13.11, VOS). Only subjects from the TwinsUK GWAS and EPIC replication cohorts were included in the male-specific analysis, which generated genome-wide significant results for five genes: *CCDC170* (6q25.1, BUA), *WNT16* (7q31.31, VOS), *FAM3C* (7q31.31, BUA and VOS), *CPED1* (7q31.31, VOS) and *DEFB103B* (8p23.1, BUA). Of these, the finding at *DEFB103B* for BUA is novel.

Gene-based fracture analysis

The ten gene-regions identified as associated with QUS parameters at the genome-wide level in the gender-combined analysis were tested for association with fracture rate using the VEGAS2 software, with the results from the three cohorts meta-analysed. Using a multiple-testing corrected significance threshold of $P < 0.005$, none of the gene regions were found to be associated with fracture rate. However, nominally significant associations were observed for the *RSPO3*, *CCDC170* and *WNT16* gene regions ($P = 0.009$, 0.02 and 0.01 respectively).

Discussion

We have identified three novel loci as associated with the QUS parameter BUA at the genome-wide significance level in the single-point gender-combined meta-analysis. The first of these, located at 8p23.1 and represented by the intergenic variant rs367543, is located in close proximity to the protein phosphatase 1 regulatory subunit 3B (*PPP1R3B*) gene and a long non-coding RNA (*LOC101929128*). We assessed this variant in the recently released publicly available GEnetic Factors for OSteoporosis (GEFOS) Consortium datasets (2015 release) (22), which contain GWAS meta-analysis results for three DXA BMD phenotypes in ~32,965 individuals, and observed nominal evidence of association of the *C* allele with a reduced femoral neck BMD ($P=0.048$), consistent with our findings. Our bioinformatics analysis revealed that rs367543 is part of an LD block that is associated with expression levels of several nearby transcripts in multiple tissue types. None of these genes are presently known to have a role in bone biology, however the *PPP1R3B* gene encodes the regulatory subunit of protein phosphatase 1 and has a role in regulating glycogen synthesis in liver and skeletal muscle tissue (29). Multiple GWAS have identified variants in the *PPP1R3B* gene region as associated with circulating lipid levels (30,31) and hepatic steatosis (32). *Ppp1r3b*-null mice display increased glucose tolerance, while prolonged fasting causes them to lose more body weight and display decreased blood glucose levels when compared with wild-type littermates (33). There is evidence to suggest a relationship between blood lipids and BMD (34,35), and it has been suggested that a shared genetic link exists between high density lipoprotein cholesterol (HDL) and BMD (36). It is possible that the *PPP1R3B* locus has pleiotropic effects on blood lipids and bone metabolism.

The second novel locus we identified for BUA was at 11q23.1 and led by the intergenic variant rs7952251. Our bioinformatics analysis suggested that this variant is associated with the expression of the long non-coding RNA *RP11-356J5.12* (*LOC283140*)

($P=1.9\times 10^{-6} - 1.5\times 10^{-6}$ (25)). We also found supporting evidence of a role for this variant in bone from the GEFOS consortium datasets, where the *G* allele was associated with an increased lumbar spine and femoral neck BMD ($P=3.7\times 10^{-4}$ and 0.004 respectively), which is consistent with the allelic effect observed in this study. Of the genes located in the region, there are two that stand out as having an established role in bone biology: *IL18* and *NCAM1*. The interleukin 18 (*IL18*) gene is located ~400Kb upstream of rs7952251 and encodes a pro-inflammatory cytokine that may have a role in inhibiting osteoclast differentiation. *IL18*-transgenic mice present with deformed cortical bone and a decreased turnover rate of lumbar trabecular bone (37), while murine osteoblasts stimulated with IL18 demonstrate increased expression of osteoprotegerin (38), a negative regulator of osteoclastogenesis. We also found this gene to be expressed at a relatively high level in both our human osteoblast and osteoclast microarray gene expression datasets described previously (39). The neural cell adhesion molecule 1 (*NCAM1*) gene is located ~400Kb downstream of rs7952251 and encodes a cell adhesion protein which is a member of the immunoglobulin superfamily. *NCAM1* has been identified as strongly expressed by human osteoblasts *in vivo*, and strong expression of *NCAM1* by myeloma cells has been found to be significantly correlated with the presence of lytic bone lesions (40).

The third novel locus for BUA was identified at 22q11.21, with rs202240051 being the maximally associated variant in the region. This variant is in complete LD with rs9606139, which presents strongly as a potential regulatory variant and is associated with expression of the nearby septin 5 (*SEPT5*) gene in multiple tissue types ($P=1.3\times 10^{-5} - 1.5\times 10^{-10}$ (25)). We assessed this variant in the GEFOS consortium datasets and found the *A* allele to be associated with a reduced lumbar spine and femoral neck BMD ($P=0.02$ and 3.9×10^{-4} respectively), consistent with our findings. The *SEPT5* gene is part of a family of nucleotide binding proteins that appear to regulate cytoskeletal organisation, with *Sept5*

deficiency in mice found to exert pleiotropic effects on a select set of affective behaviours and cognitive processes (41). *SEPT5* does not have an obvious role in bone according to the current literature, however the T-box 1 (*TBX1*) gene is located ~33Kb downstream of *SEPT5* and has an established role in bone biology. *Tbx1*-null mice present with a variety of skeletal abnormalities including short stature, moderate to severe tail kinking or coiling, abnormal craniofacial bone and cervical vertebrae morphology (42-44). Studies in mice have also shown that ablation of *Tbx1* in osteoprogenitor cells impairs bone development and subsequent bone ossification (44). It is possible that variation at rs202240051/rs9606139 influences bone structure through regulatory effects on *TBX1*.

In the gene-based association testing we identified novel associations for BUA in the gender-combined and male-specific analyses for the *FAM167A* and *DEFB103B* genes respectively. Both of these genes are located in the 8p23.1 chromosomal region, which also contains rs367543 (*PPP1R3B*). The *FAM167A* gene is located ~2.2Mb downstream of this locus, with a degree of LD existing between rs367543 and the maximally associated variant in the *FAM167A* gene region, rs4841541 ($r^2=0.19$). The function of the *FAM167A* gene has not yet been fully elucidated. It is possible that the association seen with this gene is due to regulatory effects on this or nearby genes, some of which have a role in bone and include GATA binding protein 4 (*GATA4*), which may have a role in bone mineralisation (45), and cathepsin B (*CTSB*), which appears to have a role in osteoblasts (46). The defensin beta 103B (*DEFB103B*) gene is located ~1.3Mb upstream of rs367543, and the lead variant here, rs143084514, is in linkage equilibrium with rs367543, suggesting independent association signals. The *DEFB103B* gene encodes an antimicrobial agent that is produced by neutrophils and knockout of the mouse ortholog *Defb14* results in increased BMD (47).

When considering the merit of QUS (which can easily be measured in very large numbers of individuals) versus DXA as a phenotype for genetic studies of osteoporosis,

available data show that similar effect sizes on DXA have resulted in similar odds ratios on fracture risk (21), therefore the power of QUS according to our data appears to be similar to that of DXA. The fracture meta-analysis identified significant associations for the variants rs7741021 (*RSPO3*) and rs2908007 (*WNT16*), both of which have been previously reported as associated with QUS parameters and fracture risk (24). The *RSPO3* gene encodes a Wnt signalling agonist and interestingly, like *PPP1R3B*, is also a blood lipid locus (48). The *WNT16* gene is a well-established bone locus and demonstrated the strongest associations with both BUA and VOS in this study, consistent with previous findings (24). The variant rs2908007 is in moderate LD with the maximally associated variants at the *WNT16* locus from the most recent large GWAS meta-analysis for BMD phenotypes, rs3779381 (femoral neck BMD, $r^2=0.53$) and rs7807953 (lumbar spine BMD, $r^2=0.48$) (22). The product of the *WNT16* gene is produced by osteoblast-lineage cells and activates both the canonical (β -catenin dependent) and non-canonical (β -catenin independent) Wnt signalling pathways to inhibit osteoclastogenesis both directly by acting on osteoclast progenitors and indirectly by stimulating increased expression of osteoprotegerin by osteoblasts (49). We also identified nominally significant associations between the variants rs4869739 (*CCDC170/ESR1*) and rs597319 (*TMEM135*) and fracture risk. Variation in the *ESR1* gene has been previously shown to be associated with fracture (50), however as far as we are aware, this is the first study demonstrating associations between variation in the *TMEM135* gene and fracture risk.

In conclusion, GWAS meta-analysis for QUS parameters of bone in a total of 16,627 individuals, the first to use a combination of WGS and deeply imputed genotype data for this phenotype, identified three novel loci as associated with the QUS parameter BUA at the genome-wide significance level in the gender-combined meta-analysis: 8p23.1 (*PPP1R3B*), 11q23.1 (*LOC387810*) and 22q11.21 (*SEPT5*). Gene-based association testing revealed two additional novel loci for BUA, one at *FAM167A* (8p23.1) in the gender-combined analysis

and one at *DEFB103B* (8p23.1) in the male-specific analysis. We also demonstrated significant associations between variation in the *RSPO3* and *WNT16* genes and fracture risk.

Materials and Methods

Cohort descriptions

The discovery cohort used in this study was the TwinsUK WGS cohort, with the replication cohorts being the TwinsUK GWAS and European Prospective Investigation into Cancer and Nutrition study (EPIC) cohorts.

TwinsUK

The TwinsUK WGS and TwinsUK GWAS cohorts are independent sample sets derived from TwinsUK, which is comprised of ~12,000 monozygotic and dizygotic twins recruited without selection for any particular trait from St Thomas' UK Adult Twin Registry (TwinsUK) (www.twinsuk.ac.uk/). This cohort is from Northern European/UK ancestry and has been shown to be representative of singleton populations and the UK population in general (51).

The individuals in the TwinsUK WGS cohort are all unrelated, whereas the TwinsUK GWAS cohort contains related individuals. Any individuals in the TwinsUK GWAS subset who were related to individuals in the TwinsUK WGS group were removed prior to the analysis.

Medical history and lifestyle-factor data were obtained using detailed health questionnaires, including a fracture history (any fracture since 16 years of age). Individuals suffering from rheumatoid arthritis, on oral steroid medication or having had surgical oophorectomy were excluded from the study. Genomic DNA was extracted from whole blood samples obtained for the majority of the cohort at the time of the study visit. QUS measurements were obtained at the left heel using a McCue CUBA Clinical scanner (McCue Ultrasonics, Hampshire, UK). All study participants provided written, informed consent and the research was approved by the Guy's and St Thomas' Hospital Research Ethics Committee.

European Prospective Investigation into Cancer and Nutrition study (EPIC)

The EPIC cohort is part of a large multi-centre population-based study made up of 30,000 men and women aged between 40 and 79 years at baseline from Norfolk, UK. The participants have predominantly Northern European/UK ancestry and were recruited from general practice registers between 1993 and 1997. A second clinic visit was performed for ~15,000 of the participants between 1997 and 2000, during which ultrasound measurements of the left and right calcaneum were obtained, with the mean of these measurements used for analysis (52). QUS measurements were obtained using a McCue CUBA Clinical scanner (McCue Ultrasonics, Hampshire, UK) and a fracture history was obtained for each subject (any fracture excluding those of the skull, fingers or toes). All participants gave signed informed consent and The Norwich District Health Authority Ethics Committee approved the study.

WGS data generation

Generation of the WGS data for the TwinsUK WGS cohort was performed as part of the UK10K project and has been described previously (53). Briefly, low-read depth WGS was performed in 1,990 women from the TwinsUK cohort at the Wellcome Trust Sanger Institute and the Beijing Genomics Institute. Sequencing reads were aligned to the GRCh37 human reference used in Phase 1 of the 1000GP (54). Calling of single nucleotide variants and small insertion/deletions was made using samtools/bcftools (version 0.1.18-r579) (55). SHAPEIT2 (56) was used to re-phase the genotype data. Multi-allelic variants and those with a MAF <0.0002 were removed. Any variants showing sequencing centre batch effects were also removed along with any that were not in Hardy-Weinberg equilibrium ($P < 1 \times 10^{-6}$). Related individuals, those that failed a gender check and those from non-European ancestry were removed.

GWAS genotyping and imputation

A total of 1,610 samples were selected for inclusion in the TwinsUK GWAS cohort. These individuals were unrelated to those in the WGS cohort and had been genotyped using the Illumina HumanHap300, HumanHap610Q, 1M-Duo or 1.2M-Duo arrays as described previously (57,58). Genotype imputation was performed using the combined UK10K/1000GP Phase 1 reference panels as described previously (53). Briefly, QC criteria that were applied to the TwinsUK GWAS cohort included gender check, heterozygosity and zygosity checks. Individuals from non-European ancestry were excluded. SHAPEIT2 (56) was used to re-phase the haplotypes by 3MB chunks with 250KB buffering regions. Imputation was then performed on the same chunks using IMPUTE2 (59). Imputed variants with an info score <0.4 or Hardy Weinberg $P < 1 \times 10^{-6}$ were excluded.

Genotyping in the EPIC cohort was performed using the UK Biobank Axiom Array in conjunction with the Axiom GT1 genotype calling algorithm. Samples with a call rate $<97\%$ were excluded, as were any that failed heterozygosity or gender testing. Variants with a call rate $<95\%$ or Hardy-Weinberg $P < 1 \times 10^{-8}$ were excluded. Genotype imputation was performed using the IMPUTE2 software (59) in conjunction with the 1000GP Phase 3 reference panel, which has been shown to capture $>99\%$ of SNPs with a frequency >0.01 for a variety of ancestries (60).

Statistical analysis

Standardised residuals were generated for the BUA and VOS phenotype data within each cohort for each gender separately using the covariates age, age², height and weight. The EPIC cohort was also adjusted for principal components at this time to control for population stratification. Each set of standardised residuals had a mean of 0 and variance of 1. Outliers were identified (± 3 standard deviations from the mean) and removed. Only variants with a MAF ≥ 0.005 within each cohort were included in the analysis. Assessment of the correlation

between BUA and VOS was performed on the standardised residuals using Spearman's rank correlation coefficient. Association analysis in the TwinsUK WGS and EPIC cohorts was performed using SNPTEST v2 (61), whereas analysis in the TwinsUK GWAS cohort was performed using GEMMA (Genome-wide Efficient Mixed Model Association) (62), which controls for relatedness within the cohort. Meta-analysis of the results from the different cohorts was performed using the GWAMA (Genome-Wide Association Meta-Analysis) software package (63) – a fixed effects model was applied, combining estimates of the allelic effect size and standard error from each cohort. Analysis of the phenotypic variance attributable to the genome-wide significant loci was performed in each cohort separately using the Genome-wide Complex Trait Analysis (GCTA) software package (64). Meta-analysis of the results from the three cohorts was performed using a linear mixed-effects model. Conditional analyses were performed on the meta-analysis results for selected loci using GCTA (64). All variants in the region were conditioned on the maximally associated variant using LD data from the TwinsUK WGS cohort. Genome-wide significant and suggestive thresholds for the single-point analyses were set at 5×10^{-8} and 5×10^{-7} respectively. It should be noted that this genome-wide significant threshold, although being the most commonly used for GWAS, is on the liberal end of the spectrum for deeply imputed genotype data. Therefore any genome-wide significant associations located close to this threshold should be interpreted with caution and likely need further replication to ensure they are correct. Any variants identified as associated with QUS parameters at the genome-wide significance level in the gender-combined analysis were tested for association with fracture rate. Fracture analysis for individual variants was performed in the TwinsUK WGS and GWAS cohorts using logistic regression adjusted for age, age², gender, height and weight (SPSS Statistics 22). Fracture analysis in the EPIC cohort was performed using a binary model in SNPTEST v2 adjusted for the covariates age, age², gender, height, weight and

principal components. Meta-analysis of the fracture odds ratios from each cohort was performed using the “metagen” function in the “meta” package in R v3.1.2 using a fixed effects model. Correction for multiple testing in the fracture analysis was performed according to the Bonferroni method.

Gene-based ($\pm 50\text{Kb}$) tests of association with BUA and VOS were performed on the GWAS results for each cohort separately using the VEGAS2 software (28), which assigns variants to genes and calculates gene-based empirical association p-values while accounting for the LD structure within the gene. Meta-analysis of the results from the three cohorts was performed using the weighted Z-test method (65). A genome-wide Bonferroni-corrected significance threshold of 1.95×10^{-6} for 25,579 gene tests was used, with suggestive significance set at 1×10^{-5} . Any gene regions identified as associated with QUS parameters at the genome-wide significance level in the gender-combined analysis were tested for association with fracture rate. For this, all variants within each gene region ($\pm 50\text{Kb}$) were tested for association with fracture rate while adjusting for the covariates age, age², gender, height and weight (principal components also included as covariates for the EPIC cohort) using either SNPTEST v2 (TwinsUK WGS and EPIC cohorts) or GEMMA (TwinsUK GWAS cohort). Gene-based tests of association were then performed using the VEGAS2 software (28), with meta-analysis of the results performed using the weighted Z-test method (65). Correction for multiple testing in the gene-based fracture analysis was performed using the Bonferroni method.

Bioinformatics analysis

Analysis of the LD surrounding variants of interest was performed using LDlink (1000GP Phase 3 EUR population) (66) and HaploReg v4.1 (67). Prediction of histone marks, DNase hypersensitivity sites and expression quantitative trait locus (eQTL) associations was

performed using HaploReg v4.1 (67) and genomic evolutionary rate profiling (GERP) scores were obtained using GWAVA (Genome Wide Annotation of Variants) (68).

Funding

This work was supported by the Australian National Health and Medical Research Council (Project Grant 1048216) and the iVEC/Pawsey Supercomputing Centre (with funding from the Australian Government and the Government of Western Australia; PN/14/736 and PN/15/1007). The salary of B.H.M. is supported by a Raine Medical Research Foundation Priming Grant.

Acknowledgements

We are grateful to the participants in the cohort studies and the staff involved including interviewers, computer and laboratory technicians, clerical workers, research scientists, volunteers, managers, receptionists and nurses. This study makes use of the data generated by the UK10K Consortium. Funding for UK10K was provided by the Wellcome Trust under award WT091310. A full list of the investigators who contributed to the generation of the data is available at www.UK10K.org. This study presents independent research supported by the National Institute for Health Research (NIHR) Biomedical Research Centre at South London and Maudsley NHS Foundation Trust and King's College London, the IoPPN Genomics & Biomarker Core Facility, King's College London. The views expressed are those of the author(s) and not necessarily those of the NHS, the NIHR, Department of Health or King's College London. Genome-wide summary association statistics for the gender-combined meta-analysis have been submitted to The University of Western Australia's Research Data Online resource (<https://researchdataonline.research.uwa.edu.au>).

Conflict of Interest statement. None declared.

References

1. Kanis, J.A., Melton, L.J., 3rd, Christiansen, C., Johnston, C.C. and Khaltsev, N. (1994) The diagnosis of osteoporosis. *J. Bone Miner. Res.*, **9**, 1137-1141.
2. Heaney, R.P., Recker, R.R., Stegman, M.R. and Moy, A.J. (1989) Calcium absorption in women: relationships to calcium intake, estrogen status, and age. *J. Bone Miner. Res.*, **4**, 469-475.
3. Nordin, B.E., Need, A.G., Morris, H.A., Horowitz, M. and Robertson, W.G. (1991) Evidence for a renal calcium leak in postmenopausal women. *J. Clin. Endocrinol. Metab.*, **72**, 401-407.
4. Robbins, J.A., Biggs, M.L. and Cauley, J. (2006) Adjusted mortality after hip fracture: From the cardiovascular health study. *J. Am. Geriatr. Soc.*, **54**, 1885-1891.
5. Flicker, L., Hopper, J.L., Rodgers, L., Kaymakci, B., Green, R.M. and Wark, J.D. (1995) Bone density determinants in elderly women: a twin study. *J. Bone Miner. Res.*, **10**, 1607-1613.
6. Michaelsson, K., Melhus, H., Ferm, H., Ahlbom, A. and Pedersen, N.L. (2005) Genetic liability to fractures in the elderly. *Arch. Intern. Med.*, **165**, 1825-1830.
7. Deng, H.W., Chen, W.M., Recker, S., Stegman, M.R., Li, J.L., Davies, K.M., Zhou, Y., Deng, H., Heaney, R. and Recker, R.R. (2000) Genetic determination of Colles' fracture and differential bone mass in women with and without Colles' fracture. *J. Bone Miner. Res.*, **15**, 1243-1252.
8. Keen, R.W., Hart, D.J., Arden, N.K., Doyle, D.V. and Spector, T.D. (1999) Family history of appendicular fracture and risk of osteoporosis: a population-based study. *Osteoporos. Int.*, **10**, 161-166.
9. Lin, L., Lin, W. and Qin, Y.X. (2015) Enhanced correlation between quantitative ultrasound and structural and mechanical properties of bone using combined transmission-reflection measurement. *J. Acoust. Soc. Am.*, **137**, 1144-1152.
10. Chin, K.Y. and Ima-Nirwana, S. (2013) Calcaneal quantitative ultrasound as a determinant of bone health status: what properties of bone does it reflect? *International journal of medical sciences*, **10**, 1778-1783.
11. Hans, D., Wu, C., Njeh, C.F., Zhao, S., Augat, P., Newitt, D., Link, T., Lu, Y., Majumdar, S. and Genant, H.K. (1999) Ultrasound velocity of trabecular cubes reflects mainly bone density and elasticity. *Calcif. Tissue Int.*, **64**, 18-23.
12. Padilla, F., Jenson, F., Bousson, V., Peyrin, F. and Laugier, P. (2008) Relationships of trabecular bone structure with quantitative ultrasound parameters: in vitro study on human proximal femur using transmission and backscatter measurements. *Bone*, **42**, 1193-1202.
13. Bauer, D.C., Gluer, C.C., Cauley, J.A., Vogt, T.M., Ensrud, K.E., Genant, H.K. and Black, D.M. (1997) Broadband ultrasound attenuation predicts fractures strongly and independently of densitometry in older women. A prospective study. Study of Osteoporotic Fractures Research Group. *Arch. Intern. Med.*, **157**, 629-634.
14. Hans, D., Dargent-Molina, P., Schott, A.M., Sebert, J.L., Cormier, C., Kotzki, P.O., Delmas, P.D., Pouilles, J.M., Breart, G. and Meunier, P.J. (1996) Ultrasonographic heel measurements to predict hip fracture in elderly women: the EPIDOS prospective study. *Lancet*, **348**, 511-514.
15. Moayyeri, A., Adams, J.E., Adler, R.A., Krieg, M.A., Hans, D., Compston, J. and Lewiecki, E.M. (2012) Quantitative ultrasound of the heel and fracture risk assessment: an updated meta-analysis. *Osteoporos. Int.*, **23**, 143-153.

16. Arden, N.K., Baker, J., Hogg, C., Baan, K. and Spector, T.D. (1996) 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.
17. Danielson, M.E., Cauley, J.A., Baker, C.E., Newman, A.B., Dorman, J.S., Towers, J.D. and Kuller, L.H. (1999) Familial resemblance of bone mineral density (BMD) and calcaneal ultrasound attenuation: the BMD in mothers and daughters study. *J. Bone Miner. Res.*, **14**, 102-110.
18. Howard, G.M., Nguyen, T.V., Harris, M., Kelly, P.J. and Eisman, J.A. (1998) Genetic and environmental contributions to the association between quantitative ultrasound and bone mineral density measurements: a twin study. *J. Bone Miner. Res.*, **13**, 1318-1327.
19. Hunter, D.J., de Lange, M., Andrew, T., Snieder, H., MacGregor, A.J. and Spector, T.D. (2001) Genetic variation in bone mineral density and calcaneal ultrasound: a study of the influence of menopause using female twins. *Osteoporos. Int.*, **12**, 406-411.
20. Karasik, D., Myers, R.H., Hannan, M.T., Gagnon, D., McLean, R.R., Cupples, L.A. and Kiel, D.P. (2002) Mapping of quantitative ultrasound of the calcaneus bone to chromosome 1 by genome-wide linkage analysis. *Osteoporos. Int.*, **13**, 796-802.
21. Estrada, K., Styrkarsdottir, U., Evangelou, E., Hsu, Y.H., Duncan, E.L., Ntzani, E.E., Oei, L., Albagha, O.M., Amin, N., Kemp, J.P. *et al.* (2012) Genome-wide meta-analysis identifies 56 bone mineral density loci and reveals 14 loci associated with risk of fracture. *Nat. Genet.*, **44**, 491-501.
22. Zheng, H.F., Forgetta, V., Hsu, Y.H., Estrada, K., Rosello-Diez, A., Leo, P.J., Dahia, C.L., Park-Min, K.H., Tobias, J.H., Kooperberg, C. *et al.* (2015) Whole-genome sequencing identifies EN1 as a determinant of bone density and fracture. *Nature*, **526**, 112-117.
23. Welter, D., MacArthur, J., Morales, J., Burdett, T., Hall, P., Junkins, H., Klemm, A., Flicek, P., Manolio, T., Hindorff, L. *et al.* (2014) The NHGRI GWAS Catalog, a curated resource of SNP-trait associations. *Nucleic Acids Res*, **42**, D1001-1006.
24. Moayyeri, A., Hsu, Y.H., Karasik, D., Estrada, K., Xiao, S.M., Nielson, C., Srikanth, P., Giroux, S., Wilson, S.G., Zheng, H.F. *et al.* (2014) Genetic determinants of heel bone properties: genome-wide association meta-analysis and replication in the GEFOS/GENOMOS consortium. *Hum. Mol. Genet.*, **23**, 3054-3068.
25. (2015) Human genomics. The Genotype-Tissue Expression (GTEx) pilot analysis: multitissue gene regulation in humans. *Science*, **348**, 648-660.
26. Stranger, B.E., Nica, A.C., Forrest, M.S., Dimas, A., Bird, C.P., Beazley, C., Ingle, C.E., Dunning, M., Flicek, P., Koller, D. *et al.* (2007) Population genomics of human gene expression. *Nat. Genet.*, **39**, 1217-1224.
27. Westra, H.J., Peters, M.J., Esko, T., Yaghootkar, H., Schurmann, C., Kettunen, J., Christiansen, M.W., Fairfax, B.P., Schramm, K., Powell, J.E. *et al.* (2013) Systematic identification of trans eQTLs as putative drivers of known disease associations. *Nat. Genet.*, **45**, 1238-1243.
28. Mishra, A. and Macgregor, S. (2015) VEGAS2: Software for More Flexible Gene-Based Testing. *Twin research and human genetics : the official journal of the International Society for Twin Studies*, **18**, 86-91.
29. Munro, S., Cuthbertson, D.J., Cunningham, J., Sales, M. and Cohen, P.T. (2002) Human skeletal muscle expresses a glycogen-targeting subunit of PP1 that is identical to the insulin-sensitive glycogen-targeting subunit G(L) of liver. *Diabetes*, **51**, 591-598.

30. Teslovich, T.M., Musunuru, K., Smith, A.V., Edmondson, A.C., Stylianou, I.M., Koseki, M., Pirruccello, J.P., Ripatti, S., Chasman, D.I., Willer, C.J. *et al.* (2010) Biological, clinical and population relevance of 95 loci for blood lipids. *Nature*, **466**, 707-713.
31. Waterworth, D.M., Ricketts, S.L., Song, K., Chen, L., Zhao, J.H., Ripatti, S., Aulchenko, Y.S., Zhang, W., Yuan, X., Lim, N. *et al.* (2010) Genetic variants influencing circulating lipid levels and risk of coronary artery disease. *Arterioscler. Thromb. Vasc. Biol.*, **30**, 2264-2276.
32. Speliotes, E.K., Yerges-Armstrong, L.M., Wu, J., Hernaez, R., Kim, L.J., Palmer, C.D., Gudnason, V., Eiriksdottir, G., Garcia, M.E., Launer, L.J. *et al.* (2011) Genome-wide association analysis identifies variants associated with nonalcoholic fatty liver disease that have distinct effects on metabolic traits. *PLoS genetics*, **7**, e1001324.
33. Kelsall, I.R., Rosenzweig, D. and Cohen, P.T. (2009) Disruption of the allosteric phosphorylase a regulation of the hepatic glycogen-targeted protein phosphatase 1 improves glucose tolerance in vivo. *Cell. Signal.*, **21**, 1123-1134.
34. Lidfeldt, J., Holmdahl, L., Samsioe, G., Nerbrand, C., Nyberg, P., Schersten, B. and Agardh, C.D. (2002) The influence of hormonal status and features of the metabolic syndrome on bone density: a population-based study of Swedish women aged 50 to 59 years. The women's health in the Lund area study. *Metabolism*, **51**, 267-270.
35. Yamaguchi, T., Sugimoto, T., Yano, S., Yamauchi, M., Sowa, H., Chen, Q. and Chihara, K. (2002) Plasma lipids and osteoporosis in postmenopausal women. *Endocr. J.*, **49**, 211-217.
36. Ackert-Bicknell, C.L. (2012) HDL cholesterol and bone mineral density: is there a genetic link? *Bone*, **50**, 525-533.
37. Cornish, J., Gillespie, M.T., Callon, K.E., Horwood, N.J., Moseley, J.M. and Reid, I.R. (2003) Interleukin-18 is a novel mitogen of osteogenic and chondrogenic cells. *Endocrinology*, **144**, 1194-1201.
38. Makiishi-Shimobayashi, C., Tsujimura, T., Iwasaki, T., Yamada, N., Sugihara, A., Okamura, H., Hayashi, S. and Terada, N. (2001) Interleukin-18 up-regulates osteoprotegerin expression in stromal/osteoblastic cells. *Biochem. Biophys. Res. Commun.*, **281**, 361-366.
39. Mullin, B.H., Mamotte, C., Prince, R.L. and Wilson, S.G. (2014) Influence of ARHGEF3 and RHOA knockdown on ACTA2 and other genes in osteoblasts and osteoclasts. *PloS one*, **9**, e98116.
40. Ely, S.A. and Knowles, D.M. (2002) Expression of CD56/neural cell adhesion molecule correlates with the presence of lytic bone lesions in multiple myeloma and distinguishes myeloma from monoclonal gammopathy of undetermined significance and lymphomas with plasmacytoid differentiation. *Am. J. Pathol.*, **160**, 1293-1299.
41. Suzuki, G., Harper, K.M., Hiramoto, T., Sawamura, T., Lee, M., Kang, G., Tanigaki, K., Buell, M., Geyer, M.A., Trimble, W.S. *et al.* (2009) Sept5 deficiency exerts pleiotropic influence on affective behaviors and cognitive functions in mice. *Hum. Mol. Genet.*, **18**, 1652-1660.
42. Jerome, L.A. and Papaioannou, V.E. (2001) DiGeorge syndrome phenotype in mice mutant for the T-box gene, Tbx1. *Nat. Genet.*, **27**, 286-291.
43. Vitelli, F., Zhang, Z., Huynh, T., Sobotka, A., Mupo, A. and Baldini, A. (2006) Fgf8 expression in the Tbx1 domain causes skeletal abnormalities and modifies the aortic arch but not the outflow tract phenotype of Tbx1 mutants. *Dev. Biol.*, **295**, 559-570.
44. Funato, N., Nakamura, M., Richardson, J.A., Srivastava, D. and Yanagisawa, H. (2015) Loss of Tbx1 induces bone phenotypes similar to cleidocranial dysplasia. *Hum. Mol. Genet.*, **24**, 424-435.

45. Guemes, M., Garcia, A.J., Rigueur, D., Runke, S., Wang, W., Zhao, G., Mayorga, V.H., Atti, E., Tetradis, S., Peault, B. *et al.* (2014) GATA4 is essential for bone mineralization via ERalpha and TGFbeta/BMP pathways. *J. Bone Miner. Res.*, **29**, 2676-2687.
46. Aisa, M.C., Beccari, T., Costanzi, E. and Maggio, D. (2003) Cathepsin B in osteoblasts. *Biochim. Biophys. Acta*, **1621**, 149-159.
47. Eppig, J.T., Smith, C.L., Blake, J.A., Ringwald, M., Kadin, J.A., Richardson, J.E. and Bult, C.J. (2017) Mouse Genome Informatics (MGJ): Resources for Mining Mouse Genetic, Genomic, and Biological Data in Support of Primary and Translational Research. *Methods Mol. Biol.*, **1488**, 47-73.
48. Surakka, I., Horikoshi, M., Magi, R., Sarin, A.P., Mahajan, A., Lagou, V., Marullo, L., Ferreira, T., Miraglio, B., Timonen, S. *et al.* (2015) The impact of low-frequency and rare variants on lipid levels. *Nat. Genet.*, **47**, 589-597.
49. Moverare-Skrtic, S., Henning, P., Liu, X., Nagano, K., Saito, H., Borjesson, A.E., Sjogren, K., Windahl, S.H., Farman, H., Kindlund, B. *et al.* (2014) Osteoblast-derived WNT16 represses osteoclastogenesis and prevents cortical bone fragility fractures. *Nat. Med.*, **20**, 1279-1288.
50. Ioannidis, J.P., Ralston, S.H., Bennett, S.T., Brandi, M.L., Grinberg, D., Karassa, F.B., Langdahl, B., van Meurs, J.B., Mosekilde, L., Scollen, S. *et al.* (2004) Differential genetic effects of ESR1 gene polymorphisms on osteoporosis outcomes. *JAMA*, **292**, 2105-2114.
51. Spector, T.D. and Williams, F.M. (2006) The UK Adult Twin Registry (TwinsUK). *Twin research and human genetics : the official journal of the International Society for Twin Studies*, **9**, 899-906.
52. Welch, A., Camus, J., Dalzell, N., Oakes, S., Reeve, J. and Khaw, K.T. (2004) Broadband ultrasound attenuation (BUA) of the heel bone and its correlates in men and women in the EPIC-Norfolk cohort: a cross-sectional population-based study. *Osteoporos. Int.*, **15**, 217-225.
53. Taylor, P.N., Porcu, E., Chew, S., Campbell, P.J., Traglia, M., Brown, S.J., Mullin, B.H., Shihab, H.A., Min, J., Walter, K. *et al.* (2015) Whole-genome sequence-based analysis of thyroid function. *Nature communications*, **6**, 5681.
54. Abecasis, G.R., Altshuler, D., Auton, A., Brooks, L.D., Durbin, R.M., Gibbs, R.A., Hurles, M.E. and McVean, G.A. (2010) A map of human genome variation from population-scale sequencing. *Nature*, **467**, 1061-1073.
55. Li, H. (2011) A statistical framework for SNP calling, mutation discovery, association mapping and population genetical parameter estimation from sequencing data. *Bioinformatics*, **27**, 2987-2993.
56. Delaneau, O., Marchini, J. and Zagury, J.F. (2012) A linear complexity phasing method for thousands of genomes. *Nat Methods*, **9**, 179-181.
57. Richards, J.B., Rivadeneira, F., Inouye, M., Pastinen, T.M., Soranzo, N., Wilson, S.G., Andrew, T., Falchi, M., Gwilliam, R., Ahmadi, K.R. *et al.* (2008) Bone mineral density, osteoporosis, and osteoporotic fractures: a genome-wide association study. *Lancet*, **371**, 1505-1512.
58. Moayyeri, A., Hammond, C.J., Valdes, A.M. and Spector, T.D. (2013) Cohort Profile: TwinsUK and healthy ageing twin study. *Int. J. Epidemiol.*, **42**, 76-85.
59. Howie, B.N., Donnelly, P. and Marchini, J. (2009) A flexible and accurate genotype imputation method for the next generation of genome-wide association studies. *PLoS genetics*, **5**, e1000529.

60. Auton, A., Brooks, L.D., Durbin, R.M., Garrison, E.P., Kang, H.M., Korbel, J.O., Marchini, J.L., McCarthy, S., McVean, G.A. and Abecasis, G.R. (2015) A global reference for human genetic variation. *Nature*, **526**, 68-74.
61. Marchini, J., Howie, B., Myers, S., McVean, G. and Donnelly, P. (2007) A new multipoint method for genome-wide association studies by imputation of genotypes. *Nat. Genet.*, **39**, 906-913.
62. Zhou, X. and Stephens, M. (2012) Genome-wide efficient mixed-model analysis for association studies. *Nat. Genet.*, **44**, 821-824.
63. Magi, R. and Morris, A.P. (2010) GWAMA: software for genome-wide association meta-analysis. *BMC Bioinformatics*, **11**, 288.
64. Yang, J., Ferreira, T., Morris, A.P., Medland, S.E., Madden, P.A., Heath, A.C., Martin, N.G., Montgomery, G.W., Weedon, M.N., Loos, R.J. *et al.* (2012) Conditional and joint multiple-SNP analysis of GWAS summary statistics identifies additional variants influencing complex traits. *Nat. Genet.*, **44**, 369-375, S361-363.
65. Zaykin, D.V. (2011) Optimally weighted Z-test is a powerful method for combining probabilities in meta-analysis. *J Evol Biol*, **24**, 1836-1841.
66. Machiela, M.J. and Chanock, S.J. (2015) LDlink: a web-based application for exploring population-specific haplotype structure and linking correlated alleles of possible functional variants. *Bioinformatics*, **31**, 3555-3557.
67. Ward, L.D. and Kellis, M. (2012) HaploReg: a resource for exploring chromatin states, conservation, and regulatory motif alterations within sets of genetically linked variants. *Nucleic Acids Res*, **40**, D930-934.
68. Ritchie, G.R., Dunham, I., Zeggini, E. and Flicek, P. (2014) Functional annotation of noncoding sequence variants. *Nat Methods*, **11**, 294-296.
69. Winkler, T.W., Kutalik, Z., Gorski, M., Lottaz, C., Kronenberg, F. and Heid, I.M. (2015) EasyStrata: evaluation and visualization of stratified genome-wide association meta-analysis data. *Bioinformatics*, **31**, 259-261.
70. Pruim, R.J., Welch, R.P., Sanna, S., Teslovich, T.M., Chines, P.S., Gliedt, T.P., Boehnke, M., Abecasis, G.R. and Willer, C.J. (2010) LocusZoom: regional visualization of genome-wide association scan results. *Bioinformatics*, **26**, 2336-2337.

Figure 1. Description of study design. The three cohorts were meta-analysed for both single-point and gene-based association testing, giving a combined total of 16,627 individuals. Fracture analysis was performed on the eight genome-wide significant loci from the gender-combined single-point meta-analysis and the ten genome-wide significant gene regions from the gender-combined gene-based analysis.

Figure 2. Manhattan plots for the BUA gender-combined (A) and VOS gender-combined (B) meta-analysis results, and Miami plots for the BUA gender-specific (C) and VOS gender-specific (D) meta-analyses. The gender-specific results are presented with the female-specific results in the upper section and male-specific results in the lower section of each image. Each plot depicts the variants genotyped across the 22 autosomes against the $-\log_{10} P$ value from the meta-analysis. The red line represents the genome-wide significance threshold of 5×10^{-8} (69).

Figure 3. Regional association plots of the novel loci identified in the BUA gender-combined meta-analysis for 8p23.1 (A), 11q23.1 (B) and 22q11.21 (C). Genetic variants within 400kb of the lead variant are depicted (x axis) along with their meta-analysis P value ($-\log_{10}$). Variants are colour coded according to their LD (r^2) with the lead variant (1000GP Nov 2014 EUR population). The recombination rate (blue line) and position of genes, their exons and direction of transcription is also indicated (70).

Supplementary Figure S1. Quantile-quantile plots for the BUA gender-combined (A), VOS gender-combined (B), BUA female-specific (C), VOS female-specific (D), BUA male-

specific (E) and VOS male-specific (F) meta-analyses results. Plots display observed vs expected P for each set of results.

Table 1. Genetic loci associated with QUS parameters at the genome-wide significance level ($P < 5 \times 10^{-8}$) in the gender-combined meta-analysis.

Trait	Locus	Variant	Closest gene	EA	OA	EAF	TwinsUK WGS		TwinsUK GWAS		EPIC GWAS		Meta-analysis	
							Beta	P	Beta	P	Beta	P	Beta (SE)	P
BUA	2p16.2	rs11898505	<i>SPTBN1</i>	A	G	0.37	0.09	0.02	0.05	0.23	0.09	6.8E-13	0.09 (0.01)	1.9E-13
	6q22.33	rs7741021	<i>RSPO3</i>	C	A	0.48	0.07	0.09	0.12	0.001	0.07	5.5E-10	0.08 (0.01)	5.3E-12
	6q25.1	rs1891002	<i>CCDC170/</i> <i>ESR1</i>	A	T	0.29	-0.13	0.003	-0.01	0.74	-0.08	7.4E-10	-0.08 (0.01)	1.6E-10
	7q31.31	rs2908007	<i>WNT16</i>	G	A	0.4	0.23	3.5E-8	0.07	0.09	0.18	1.2E-48	0.17 (0.01)	1.2E-51
	8p23.1	rs367543	<i>PPP1R3B</i>	C	T	0.45	-0.12	0.003	-0.09	0.07	-0.06	1.2E-7	-0.07 (0.01)	1.6E-9
	11q14.2	rs511755	<i>TMEM135</i>	C	T	0.32	-0.07	0.08	-0.05	0.28	-0.08	6.0E-11	-0.08 (0.01)	3.5E-11
	11q23.1	rs7952251	<i>LOC387810</i>	G	A	0.47	0.04	0.29	0.06	0.13	0.07	8.6E-9	0.07 (0.01)	5.0E-9
	22q11.21	rs202240051	<i>SEPT5</i>	C	CG	0.1	-0.13	0.04	-0.10	0.11	-0.11	2.9E-7	-0.11 (0.02)	2.4E-8
VOS	2p16.2	rs11898505	<i>SPTBN1</i>	A	G	0.37	0.06	0.16	0.06	0.18	0.07	5.1E-8	0.07 (0.01)	2.0E-8
	6q22.33	rs7741021	<i>RSPO3</i>	C	A	0.48	0.09	0.02	0.05	0.18	0.08	3.7E-12	0.08 (0.01)	5.8E-13
	6q25.1	rs4869739	<i>CCDC170/</i> <i>ESR1</i>	T	A	0.29	-0.07	0.10	-0.11	0.01	-0.08	2.8E-10	-0.08 (0.01)	1.3E-11
	7q31.31	rs2908007	<i>WNT16</i>	G	A	0.4	0.16	8.7E-5	0.14	7.4E-4	0.17	3.5E-45	0.17 (0.01)	1.9E-48
	11q14.2	rs597319	<i>TMEM135</i>	G	A	0.31	-0.08	0.07	-0.07	0.15	-0.09	9.0E-14	-0.09 (0.01)	4.1E-14

EA, effect allele; OA, other allele; EAF, effect allele frequency; SE, standard error.

Table 2. Genetic loci associated with QUS parameters at the genome-wide significance level ($P < 5 \times 10^{-8}$) in the gender-stratified meta-analysis.

							TwinsUK WGS		TwinsUK GWAS		EPIC GWAS		Meta-analysis	
Trait	Locus	Variant	Closest gene	EA	OA	EAF	Beta	P	Beta	P	Beta	P	Beta (SE)	P
Female-specific														
BUA	2p16.2	rs11898505	SPTBN1	A	G	0.36	0.09	0.02	0.06	0.17	0.09	2.4E-8	0.09 (0.01)	1.7E-9
	7q31.31	rs2908007	WNT16	G	A	0.39	0.23	3.5E-8	0.07	0.09	0.16	2.3E-22	0.16 (0.01)	1.3E-27
	11q23.1	rs7109326	LOC387810	G	A	0.45	0.04	0.29	0.05	0.20	0.10	1.2E-9	0.09 (0.01)	2.0E-9
VOS	7q31.31	rs2908007	WNT16	G	A	0.39	0.16	8.7E-5	0.14	1.3E-3	0.17	5.3E-25	0.17 (0.01)	1.6E-29
	11q14.2	rs597319	TMEM135	G	A	0.31	-0.08	0.07	-0.06	0.30	-0.09	4.9E-8	-0.09 (0.02)	1.4E-8
Male-specific														
BUA	6q25.1	rs9383929	CCDC170/ ESR1	A	T	0.29	ND	ND	0.12	0.24	-0.12	2.5E-9	-0.11 (0.02)	3.0E-8
	7q31.31	rs2908007	WNT16	G	A	0.40	ND	ND	0.03	0.75	0.20	1.9E-28	0.20 (0.02)	3.7E-27
VOS	6q25.1	rs66943969	CCDC170/ ESR1	G	-	0.29	ND	ND	-0.06	0.55	-0.12	2.0E-9	-0.12 (0.02)	3.2E-9
	7q31.31	rs3779381	WNT16	G	A	0.25	ND	ND	0.12	0.26	0.20	5.9E-22	0.19 (0.02)	1.2E-21

EA, effect allele; OA, other allele; EAF, effect allele frequency; SE, standard error; ND, no data.

Table 3. Genetic variants tested for association with fracture data.

Locus	Variant	Closest gene	EA	OA	EAF	TwinsUK WGS		TwinsUK GWAS		EPIC GWAS		Meta-analysis	
						OR (95% CI)	P	OR (95% CI)	P	OR (95% CI)	P	OR (95% CI)	P
2p16.2	rs11898505	<i>SPTBN1</i>	A	G	0.37	0.82 (0.67-0.99)	0.04	1.09 (0.87-1.35)	0.46	1.01 (0.96-1.08)	0.65	1.00 (0.95-1.06)	0.99
6q22.33	rs7741021	<i>RSPO3</i>	C	A	0.48	0.84 (0.7-1.0)	0.05	0.98 (0.8-1.2)	0.85	0.91 (0.86-0.97)	0.002	0.91 (0.86-0.96)	4.0E-4
6q25.1	rs4869739	<i>CCDC170/</i> <i>ESR1</i>	T	A	0.29	1.13 (0.94-1.37)	0.20	1.01 (0.8-1.28)	0.91	1.06 (1.0-1.13)	0.06	1.06 (1.0-1.13)	0.03
7q31.31	rs2908007	<i>WNT16</i>	G	A	0.4	0.91 (0.75-1.09)	0.29	0.97 (0.78-1.21)	0.77	0.92 (0.87-0.98)	0.007	0.92 (0.88-0.98)	0.004
8p23.1	rs367543	<i>PPP1R3B</i>	C	T	0.45	1.09 (0.9-1.31)	0.39	0.96 (0.77-1.19)	0.71	1.0 (0.94-1.06)	0.92	1.0 (0.95-1.06)	0.95
11q14.2	rs597319	<i>TMEM135</i>	G	A	0.31	1.08 (0.89-1.3)	0.45	1.04 (0.84-1.3)	0.71	1.08 (1.01-1.15)	0.02	1.08 (1.02-1.14)	0.01
11q23.1	rs7952251	<i>LOC387810</i>	G	A	0.47	0.84 (0.7-1.01)	0.06	0.95 (0.77-1.18)	0.66	0.98 (0.93-1.04)	0.58	0.97 (0.92-1.02)	0.24
22q11.21	rs202240051	<i>SEPT5</i>	C	CG	0.1	0.84 (0.63-1.12)	0.22	0.91 (0.62-1.33)	0.63	1.05 (0.96-1.16)	0.29	1.02 (0.94-1.12)	0.62

EA, effect allele; OA, other allele; EAF, effect allele frequency; OR, odds ratio; CI, confidence interval.

Table 4. Gene regions identified by the VEGAS2 software as associated with QUS parameters at the genome-wide significance level ($P < 1.95 \times 10^{-6}$).

Trait	Locus	TwinsUK WGS gene-based P	TwinsUK GWAS gene-based P	EPIC GWAS gene-based P	Meta-analysis gene-based P
<i>Gender-combined</i>					
BUA	<i>WNT16</i>	3.5E-5	0.18	1.0E-6	6.0E-9
	<i>FAM3C</i>	4.1E-4	0.25	1.0E-6	2.4E-8
	<i>CPED1</i>	5.3E-4	0.56	1.0E-6	1.1E-7
	<i>RSPO3</i>	0.23	0.02	1.0E-6	1.3E-7
	<i>GPATCH1</i>	0.04	0.32	1.0E-6	3.6E-7
	<i>CCDC170</i>	0.005	0.76	1.0E-6	6.8E-7
	<i>TMEM135</i>	0.16	0.25	1.0E-6	7.5E-7
	<i>FAM167A</i>	0.05	0.20	6.0E-6	1.4E-6
VOS	<i>WNT16</i>	0.009	0.02	1.0E-6	1.2E-8
	<i>FAM3C</i>	0.01	0.05	1.0E-6	2.8E-8
	<i>CCDC170</i>	0.08	0.18	1.0E-6	3.0E-7
	<i>CPED1</i>	0.06	0.25	1.0E-6	3.6E-7
	<i>ESR1</i>	0.23	0.18	1.0E-6	7.3E-7
	<i>WDR88</i>	7.4E-05	0.05	2.2E-4	1.0E-6
	<i>GPATCH1</i>	4.7E-05	0.27	7.4E-5	1.2E-6
<i>Female-specific</i>					
BUA	<i>WNT16</i>	3.5E-5	0.13	1.0E-6	1.7E-9
	<i>FAM3C</i>	4.1E-4	0.23	1.0E-6	1.5E-8
	<i>CPED1</i>	5.3E-4	0.63	1.0E-6	1.5E-7
VOS	<i>WNT16</i>	0.009	0.01	1.0E-6	5.3E-9
	<i>FAM3C</i>	0.01	0.05	1.0E-6	2.2E-8
	<i>CPED1</i>	0.06	0.19	1.0E-6	3.7E-7
	<i>WDR88</i>	7.4E-5	0.13	1.0E-4	5.5E-7
<i>Male-specific</i>					
BUA	<i>CCDC170</i>	ND	0.15	1.0E-6	5.8E-7
	<i>FAM3C</i>	ND	0.42	1.0E-6	1.3E-6
	<i>DEFB103B</i>	ND	0.54	1.0E-6	1.8E-6
VOS	<i>FAM3C</i>	ND	0.07	1.0E-6	3.8E-7

<i>WNT16</i>	ND	0.18	1.0E-6	6.4E-7
<i>CPED1</i>	ND	0.21	1.0E-6	7.2E-7

Abbreviations

1000GP – 1000 Genomes Project

BMD – Bone mineral density

BUA – Broadband ultrasound attenuation

DXA – Dual-energy X-ray absorptiometry

EPIC – European Prospective Investigation into Cancer and Nutrition study

eQTL – Expression quantitative trait locus

GEFOS – GEnetic Factors for OSteoporosis

GEMMA – Genome-Wide Efficient Mixed Model Association

GERP – Genomic evolutionary rate profiling

GWAMA – Genome-Wide Association Meta-Analysis

GWAS – Genome-wide association study

GWAVA – Genome Wide Annotation of Variants

HDL – High density lipoprotein

LD – Linkage disequilibrium

MAF – Minor allele frequency

NHGRI – National Human Genome Research Institute

NIHR – National Institute for Health Research

QC – Quality control

QUS – Quantitative ultrasound

VEGAS2 – VErsatile Gene-based Association Study 2

VOS – Velocity of sound

WGS – Whole-genome sequence