# Identification of a novel locus on chromosome 2q13 which predisposes to clinical vertebral fractures independently of bone density

# Identification of a novel locus on chromosome 2q13 which predisposes to clinical vertebral fractures independently of bone density

Nerea Alonso[1], Karol Estrada[2], Omar ME Albagha[1,3], Lizbeth Herrera[2], Sjur Reppe[4,5,6], Ole K Olstad[4], Kaare M Gautvik[5,6], Niamh M Ryan[7], Kathryn L Evans[7,8], Carrie M Nielson[9], Yi-Hsiang Hsu[10,11,12], Douglas P Kiel[11,12], George Markozannes[13], Evangelia Ntzani[13,14], Evangelos Evangelou[13,15], Bjarke Feenstra[16], Xueping Liu[16], Mads Melbye[16,17,18], Laura Masi[19], Maria Luisa Brandi[19], Philip Riches[1], Anna Daroszewska[1,20], Jose M Olmos[21], Carmen Valero[21], Jesus Castillo[21], Jose A. Riancho[21], Lise B Husted[22], Bente L Langdahl[22], Matthew A Brown[23], Emma L Duncan[23,24,25], Stephen Kaptoge[26], Kay-Tee Khaw[27], Ricardo Usategui-Martin[28], Javier del Pino[28], Rogelio Gonzalez-Sarmiento[28], Joshua R Lewis[29,30,31], Richard L Prince[29,32], Patrizia D'Amelio[33], Natalia Garcia-Giralt[34], Xavier Nogues[34], Simona Mencej-Bedrac[35], Janja Marc[35], Orit Wolstein[36], John Eisman[36], Ling Oei[2], Carolina Medina-Gomez[2], Katharina E Schraut[37,38], Pau Navarro[39], James F Wilson[37,39], Gail Davies[8], John Starr[8], Ian Deary[8], Toshiko Tanaka[40], Luigi Ferrucci[41], Fernando Gianfrancesco[42], Luigi Gennari[43], Gavin Lucas[44], Roberto Elosua[44], Andre G Uitterlinden[2], Fernando Rivadeneira[2], Stuart H Ralston*[1].

**Author Affiliations**

[1] Rheumatology and Bone Disease Unit, Centre for Genomic and Experimental Medicine, Institute of Genetics and Molecular Medicine, University of Edinburgh, Edinburgh, UK

[2] Departments of Internal Medicine and Epidemiology, Erasmus Medical Centre, Rotterdam, The Netherlands

[3] Qatar Biomedical Research Institute, Hamad Bin Khalifa University, Doha, Qatar

[4] Department of Medical Biochemistry, Oslo University Hospital, Oslo, Norway

[5] Department of Clinical Biochemistry, Lovisenberg Diakonale Hospital, Oslo, Norway

[6] Department of Molecular Medicine, Institute of Basic Medical Sciences, University of Oslo, Oslo, Norway

[7] Centre for Genomic and Experimental Medicine, IGMM, University of Edinburgh, Edinburgh, UK

[8] Centre for Cognitive Ageing and Cognitive Epidemiology, University of Edinburgh, Edinburgh, UK

[9] Department of Public Health and Preventive Medicine, Oregon Health and Science University, Portland, Oregon, USA

[10] Department of Medicine Beth Israel Deaconess Medical Center and Harvard Medical School, Boston, MA, USA

[11] BROAD Institute of MIT and Harvard, Cambridge, MA, USA
[12] Musculoskeletal Research Center, Institute for Aging Research, Harvard Medical School, USA
[13] Department of Hygiene and Epidemiology, University of Ioannina School of Medicine, Ioannina, Greece
[14] Centre for Evidence Synthesis in Health, Department of Health Services, Policy and Practice, School of Public Health, Brown University, RI, USA
[15] Department of Epidemiology and Biostatistics, Imperial College London, London, UK
[16] Department of Epidemiology Research, Statens Serum Institut, Copenhagen, Denmark
[17] Department of Clinical Medicine, Faculty of Health and Medical Sciences, University of Copenhagen, Denmark
[18] Department of Medicine, Standford School of Medicine, Standford, California, USA
[19] Department of Surgery and Translational Medicine, University of Florence, Florence, Italy
[20] Institute of Ageing and Chronic Disease The MRC-Arthritis Research UK Centre for Integrated Research into Musculoskeletal Ageing, University of Liverpool, Liverpool, UK
[21] Department of Internal Medicine, Hospital UM Valdecilla, University of Cantabria, IDIVAL, RETICEF, Santander, Spain
[22] Department of Endocrinology and Internal Medicine THG, Aarhus University Hospital, Aarhus, Denmark
[23] Institute of Health and Biomedical Innovation, Queensland University of Technology, Translational Research Institute, Princes Alexandra Hospital, Brisbane, Australia.
[24] Faculty of Medicine, University of Queensland, Brisbane, Australia
[25] Department of Endocrinology, Royal Brisbane and Women's Hospital, Herston, Brisbane, Australia
[26] Cardiovascular Epidemiology Unit, Department of Public Health and Primary Care, University of Cambridge, Cambridge, UK
[27] Department of Public Health and Primary Care, School of Medicine, University of Cambridge, Cambridge, UK
[28] Molecular Medicine Unit – Department of Medicine and Biomedical Research Institute of Salamanca (IBSAL). University Hospital of Salamanca – University of Salamanca – CSIC.
[29] School of Medicine and Pharmacology, University of Western Australia, Perth, Australia
[30] Centre for Kidney Research, School of Public Health, University of Sidney, Sydney, Australia

[31] School of Medical and Health Sciences, Edith Cowan University, Joondalup, Western Australia, Australia
[32] Department of Endocrinology and Diabetes, Sir Charles Gairdner Hospital, Perth, Australia
[33] Gerontology and Bone Metabolic Diseases Unit, Department of Medical Science, University of Torino, Torino, Italy
[34] Department of Internal Medicine, Hospital del Mar-IMIM, RETICEF, Universitat Autonoma de Barcelona, Barcelona, Spain
[35] Department of Clinical Biochemistry, Faculty of Pharmacy, University of Ljubljana, Ljubljana, Slovenia
[36] Osteoporosis and Bone Biology Program, Garvan Institute of Medical Research, Sydney, Australia
[37] Centre for Global Health Research, Usher Institute for Population Health Sciences and Informatics, University of Edinburgh, Edinburgh, UK
[38] Edinburgh/British Heart Foundation Centre for Cardiovascular Science, QMRI, University of Edinburgh, Edinburgh, UK
[39] MRC Human Genetics Unit, MRC, IGMM, University of Edinburgh, Edinburgh, UK
[40] MedStar Research Institute, Baltimore, USA
[41] InCHIANTI project, Florence, Italy
[42] Institute of Genetics and Biophysics, Naples, Italy
[43] Department of Internal Medicine, Endocrino-metabolic Science and Biochemistry, Polyclinic Le Scotte, Siena, Italy
[44] Grup d'Epidemilogia I Genetica Cardiovascular, IMIM, Barcelona, Spain

**Correspondence to:**
Prof Stuart H Ralston MD
Centre for Genomic and Experimental Medicine
IGMM University of Edinburgh
Western General Hospital
Edinburgh EH4 2XU
E-mail: stuart.ralston@ed.ac.uk
Tel: (+44)-131-651-8741

**Word count:** 2,998

## ABSTRACT

**Objectives:** To identify genetic determinants of susceptibility to clinical vertebral fractures, an important complication of osteoporosis. **Methods:** Here we conduct a genome-wide association study in 1,553 postmenopausal women with clinical vertebral fractures and 4,340 controls, with a 2-stage replication involving 1,028 cases and 3,762 controls. Potentially causal variants were identified using eQTL data from transiliac bone biopsies and bioinformatic studies. **Results:** A locus tagged by rs10190845 was identified on chromosome 2q13 which was significantly associated with clinical vertebral fracture ($p=1.04x10^{-9}$) with a large effect size (odds ratio 1.74, 95% CI 1.06 – 2.6). Bioinformatic analysis of this locus identified several potentially functional SNPs which are associated with expression of the positional candidate genes *TTL* (Tubulin Tyrosine Ligase) and *SLC20A1* (Solute Carrier Family 20 Member 1). Three other suggestive loci were identified on chromosomes 1p31, 11q12 and 15q11. All these loci were novel and had not previously been associated with BMD or clinical fractures. **Conclusion:** We have identified a novel genetic variant that is associated with clinical vertebral fractures by mechanisms that are independent of BMD. Further studies are now in progress to validate this association and evaluate the underlying mechanism.



**KEYWORDS:** Osteoporosis, Gene polymorphism, Bone Mineral Density, *TTL*, *SLC20A1*

## 1. INTRODUCTION

Osteoporosis is a common disease with a strong genetic component. It is characterised by low bone mineral density (BMD), deterioration in the microstructural architecture of bone and an increased risk of fragility fractures. Vertebral fractures are an early and important complication of osteoporosis.[1] They are characterised by loss of height and deformity of the affected vertebrae and associated with increased risk of other fractures.[2] It has been estimated that between 8-30% of patients with radiological evidence of vertebral fractures (so called morphometric fractures) come to medical attention for reasons that are incompletely understood.[3,4] In contrast, other patients with vertebral fractures come to medical attention because of symptoms such as back pain, kyphosis, and height loss, and are defined as having clinical vertebral fractures.[5-7] Clinical vertebral fractures are associated with a markedly increased risk of future fractures and increased mortality.[8] Major advances have been made in identifying genetic variants that regulate BMD and some variants have also been identified that predispose to non-vertebral fractures.[9-20] However, the genetic determinants of vertebral fractures are poorly understood. A previous genome-wide association study (GWAS) published by Oei and colleagues involving a discovery cohort of 8,717 cases and 21,793 controls failed to identify any significant genetic predictors of radiographic vertebral fracture at a genome-wide significant level.[21] However, in this study, the vertebral fractures were defined simply on the basis of morphometric analysis of spinal radiographs. It is well recognised however that the morphometric techniques employed in this study may have identified vertebral deformities that were not fractures.[22] The aim of the present study was to re-evaluate the predictors of clinical vertebral fractures by genome wide association study to try and gain new insights into this important and poorly understood clinical problem.

## 2. PATIENTS AND METHODS

The study involved a discovery phase with 1,553 clinical vertebral fracture cases and 4,340 controls, a first replication phase of 694 cases and 2,105 controls, and a second replication phase of 334 cases and 1,657 controls, as summarised in Supplementary Table 1. The genome wide association study was performed using standard methodology as detailed in the Supplementary Text 1.

## 3. RESULTS

### 3.1. Characteristics of the study populations

The mean (±standard deviation) age of the patients with clinical vertebral fractures was 71.3±9.3 years with a bone mineral density T-score at the lumbar spine of -2.72±1.4; and at the femoral neck of -2.57±1.1. The controls were not matched with the cases by age and did not undergo phenotyping for vertebral fracture on the basis that clinical vertebral fractures are uncommon in the general population (estimated incidence of 9.8/1000 person-years in 75-84 year olds)[23]. While it is possible that clinical vertebral fractures may have occurred in some controls in later life this is unlikely to have substantially affected the results of the analysis, other than to have potentially slightly reduced its power.[24] This approach has been used previously for genome-wide studies in various common diseases including diabetes, Paget's disease, and rheumatoid arthritis.[25,26]

We identified 334 clinical vertebral fracture female cases from the UK Biobank cohort with a mean age (±standard deviation) of 58.8±7.7 years, and they were age-matched with 1,657 female controls from the same cohort.

### 3.2. Genome-wide association analysis of the discovery sample

Since different genotyping platforms were used in the analysis of the different cohorts that constitute the discovery sample, association analysis was conducted following imputation of all genotypes into the CEU panel of HapMap II reference (see Patients and Methods section). Following imputation, we analysed 2,366,456 SNPs and identified 31 with suggestive evidence of association with vertebral fracture ($p \leq 10^{-4}$). Details are summarised in Supplementary Table 2, the Manhattan and quantile-quantile plots are shown in Supplementary Figures 2 and 3. Each study was corrected by genomic control; genomic inflation factors ranged between $\lambda$=1.001 to $\lambda$=1.046 for genotyped SNPs and $\lambda$=1.006 to $\lambda$=1.036 after imputation.

### 3.3. Replication and combined analysis

We analysed the 31 suggestively associated SNPs identified in the discovery cohort (Supplementary Table 4) and seven additional SNPs that had been significantly associated with clinical fractures in a previous GWAS (Supplementary Table 5) in the replication sample.[10] Four SNPs showed nominal association ($p<0.05$) with clinical vertebral fractures at replication (Table 1). The combined discovery and replication analysis corrected for age identified one SNP (rs10190845) on chromosome 2q13 with genome-wide significant evidence of association with clinical vertebral fractures ($p=1.27 \times 10^{-8}$). The predisposing allele had a frequency of 0.034 in cases compared with 0.022 in controls and the odds ratio for susceptibility to fracture was 1.75 [95% CI: 1.44-2.12] (Figure 1). The results were similar without age correction ($p=4.9 \times 10^{-8}$; odds ratio 1.66 [95% CI: 1.38-1.99]). Conditional analysis on rs10190845 did not reveal any secondary association signals at the locus (Supplementary Figure 4). Three other SNPs on

chromosomes 1p31, 11q12 and 15q11 were suggestively associated with vertebral fracture in the combined analysis (Table 1 and Supplementary Figures 5 and 6). None of these regions have previously been found to be associated with BMD or fracture in previous GWAS.[10,13] The top SNP (rs10190845) maps to a region which contains eleven potential candidate genes (Figure 2). This region has previously been implicated as a genetic regulator of bone density by Estrada and colleagues[10] who reported that rs17040773 within *ANAPC1* (Anaphase Promoting Complex Subunit 1) was associated with femoral neck BMD ($p=1.5x10^{-9}$), but not with clinical fractures (p=0.79). rs17040773 is not in linkage disequilibrium with rs10190845 in our population ($r^2=0.006$), and, in keeping with this, when we performed conditional analysis on rs17040773, we confirmed that rs10190845 remained significantly associated with clinical vertebral fractures ($p=2.09x10^{-8}$; odds ratio 1.73 [95% CI: 1.43-2.09]). In order to test whether the variants associated with clinical vertebral fractures played a role in BMD, we tested the rs10190845 variant for association with volumetric vertebral bone mineral density in females on the dataset from Nielson and colleagues.[27] We did not find any association for the variant and BMD (p=0.23). This suggests that rs10190845 constitutes an independent signal which predisposes to clinical vertebral fracture by mechanisms that are independent of an effect on BMD.

**Table 1. Variants showing suggestive or significant association with vertebral fracture**

| | | | | Discovery (n = 5,893) | | | Replication (n= 2,799) | | | Combined* (n= 8,692) | | | | UK Biobank replication (n= 1,991) | | | Total** (n= 10,683) | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **Chr** | **SNP** | **Position** | **A** | **AF** | **p** | **OR (95% CI)** | **AF** | **p** | **OR (95% CI)** | **p** | **OR (95% CI)** | **I2** | **Q p** | **AF** | **p** | **OR (95% CI)** | **p** | **OR (95% CI)** | **$I^2$** | **Q p** |
| 2 | rs10190845 | 112192944 | A | 0.03 | $2.4x10^{-5}$ | 1.70 (1.33-2.17) | 0.05 | $1.60x10^{-4}$ | 1.84 (1.34-2.53) | $1.27x10^{-8}$ | 1.75 (1.45-2.12) | 5.9 | 0.39 | 0.05 | 0.027 | 1.66 (1.06-2.60) | $1.04x10^{-9}$ | 1.75 (1.45-2.12) | 0.0 | 0.48 |
| 11 | rs7121756 | 57980425 | A | 0.29 | $5.2x10^{-5}$ | 1.22 (1.11-1.35) | 0.28 | 0.011 | 1.23 (1.05-1.45) | $1.27x10^{-6}$ | 1.23 (1.13-1.33) | 0.0 | 0.67 | 0.29 | 0.35 | 1.09 (0.91-1.32) | $4.39x10^{-7}$ | 1.22 (1.13-1.32) | 49.0 | 0.03 |
| 15 | rs2290492 | 92464744 | A | 0.23 | $3.4x10^{-5}$ | 1.24 (1.12-1.37) | 0.21 | 0.021 | 1.23 (1.03-1.46) | $1.61x10^{-6}$ | 1.24 (1.13-1.35) | 53.7 | 0.02 | 0.22 | 0.44 | 1.08 (0.88-1.33) | $2.51x10^{-7}$ | 1.23 (1.13-1.33) | 75.6 | $1.1x10^{-5}$ |
| 1 | rs1360181 | 68248452 | C | 0.16 | $8.4x10^{-5}$ | 1.25 (1.12-1.41) | 0.17 | 0.008 | 1.30 (1.07-1.56) | $1.87x10^{-6}$ | 1.26 (1.14-1.41) | 7.7 | 0.57 | 0.17 | 0.38 | 0.90 (0.72-1.14) | $1.09x10^{-5}$ | 1.22 (1.12-1.33) | 32.2 | 0.57 |

The allele (A) and allele frequency (AF) for each of the variants is shown along with the p value for association, odds ratio (OR) and 95% confidence interval (95% CI). Q p values correspond to Cochran's Q p-values. The values shown are adjusted for age but similar results were obtained for unadjusted association tests. Position refers to Human Genome Assembly GRCh38.p11.

*Combined results showed the meta-analysis for discovery and replication stage.

**Total results showed the meta-analysis including the second replication in the UK Biobank cohort.

A second replication for the significant hit on chromosome 2 and suggestive SNPs on chromosomes 1, 11 and 15 was performed in 334 clinical vertebral fracture cases and 1,657 controls from UK Biobank. The top hit (rs10190845) on chromosome 2 was found nominally associated with clinical vertebral fractures (p=0.027, OR=1.66[1.060-2.600], MAF=0.049). No association was found for the suggestive SNPs in this cohort (Table 1).

Meta-analysis of the discovery and the two replication stages showed a combined p-value for rs10190845=$1.04x10^{-9}$ (OR=1.74[1.06-2.6]) with no evidence of heterogeneity between cohorts ($I^2$=0.0, p=0.48) (Table 1).

The SNPs rs7121756 on chromosome 11 and rs2290492 on chromosome 15 showed significant heterogeneity among cohorts (Cochrane's Q<0.05), and a random effect analysis was performed. rs7121756 remained suggestively associated with clinical vertebral fractures ($p=1.01x10^{-6}$), whilst rs2290492 showed a marginal association (p=0.004).

### 3.4.Functional evaluation of chromosome 2q13 locus

This analysis focused on a linkage disequilibrium block of approximately 700kb surrounding the top hit rs10190845. We identified a total of 936 SNPs within the region which were analysed in the GWAS (n=376) or which were in linkage disequilibrium ($r^2$ value of > 0.7) with rs10190845, or which showed suggestive association to clinical vertebral fractures ($p<5x10^{-3}$). We imputed the genotypes for the SNPs within the region of interest using the 1000 Genomes phase 3 panel as reference and tested the SNPs for association with clinical vertebral fractures. We removed 878 of the SNPs since they showed no association with clinical vertebral fractures in our dataset (p>0.05). The remaining 58 candidate SNPs were tested for association with the level of expression of genes within the candidate locus using a bone-derived gene expression dataset (eQTLs)[28] (Tables 2, 3 and Supplementary Figure 7). This resulted in the identification of nine SNPs which were eQTLs for genes within the region. In order to gain insight into the functional basis of the association at 2q13 we used SuRFR[29] which integrates functional annotation and prior biological knowledge to identify potentially causal genetic variants, to assess these 9 SNPs along with the top hit rs10190845 (Table 2 and Supplementary Figure 7).

**Table 2. Functionality of SNPs in 2q13 region, ranked by SuRFR**

| SuRFR Rank | SNP ID | $R^2$ with rs10190845 | A (AF) | GWAS p-value (Discovery cohort only) | OR (95%CI) | Location | GERP Value | DNase HS sit | DNase Foot | Ernst Score | Position Score | MAF Score | Enhancer score | TFBS score | Total score | eQTL | eQTL gene(s) | eQTL p |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | rs35586251 | 0.17 | A (0.02) | $2.09 \times 10^{-4}$ | 1.69 (1.28-2.24) | *Exon FBLN7* | 4.47 | 0 | 0 | 7 | 5 | 0.02 | 0 | 0 | 9.89 | Yes | *TTL* | $6.6 \times 10^{-6}$ |
| 2 | rs77172864 | 0.79 | G (0.03) | $4.96 \times 10^{-5}$ | 1.68 (1.31-2.17) | Intergenic | 0.18 | 0 | 0 | 1 | 3 | 0.02 | 0 | 0 | 8.56 | Yes | *SCL20A1* | 0.0001 |
| 3 | rs10190845 | 1 | A (0.03) | $2.4 \times 10^{-5}$ | 1.70 (1.33-2.17) | Intergenic | 0 | 0 | 0 | 2 | 3 | 0.96 | 0 | 0 | 8.06 | No | - | - |
| 4 | rs77996972 | 0.22 | T (0.02) | $2.11 \times 10^{-4}$ | 1.69 (1.28-2.23) | Intron *FBLN7* | 1.77 | 313 | 0 | 7 | 1 | 0.02 | 0 | 0 | 7.61 | Yes | *TTL*<br>*SLC20A1* | $3.8 \times 10^{-6}$<br>$5.5 \times 10^{-5}$ |
| 5 | rs75814334 | 0.22 | T (0.02) | $2.11 \times 10^{-4}$ | 1.69 (1.28-2.23) | Intron *FBLN7* | 0.43 | 239 | 0 | 8 | 1 | 0.02 | 0 | 0 | 7.56 | Yes | *TTL*<br>*SLC20A1* | $2.1 \times 10^{-6}$<br>$6.6 \times 10^{-5}$ |
| 6 | rs74792868 | 0.22 | A (0.02) | $2.1 \times 10^{-4}$ | 1.69 (1.28-2.24) | Intron *FBLN7* | 0 | 0 | 0 | 9 | 1 | 0.02 | 0 | 0 | 7.5 | Yes | *TTL*<br>*SLC20A1* | $2.0 \times 10^{-5}$<br>$2.8 \times 10^{-5}$ |
| 6 | rs72943913 | 0.29 | G (0.03) | $5.48 \times 10^{-5}$ | 1.67 (1.30-2.14) | Intron *ZC3H8* | 0.15 | 0 | 0 | 3 | 1 | 0.02 | 0 | 0 | 6.46 | Yes | *SLC20A1* | 0.0001 |
| 7 | rs112275607 | 0.22 | A (0.02) | $2.13 \times 10^{-4}$ | 1.69 (1.28-2.24) | Intron *FBLN7* | 0 | 0 | 0 | 8 | 1 | 0.02 | 0 | 0 | 6.83 | Yes | *TTL*<br>*SLC20A1* | $2.8 \times 10^{-6}$<br>$6.2 \times 10^{-5}$ |
| 8 | rs113085288 | 0.06 | T (0.02) | $1.79 \times 10^{-4}$ | 1.70 (1.29-2.24) | Intron *FBLN7* | 0 | 0 | 0 | 7 | 1 | 0.02 | 0 | 0 | 6.08 | Yes | *SLC20A1* | $4.1 \times 10^{-6}$ |
| 9 | rs113428223 | 0.29 | T (0.03) | $4.55 \times 10^{-5}$ | 1.70 (1.31-2.20) | Intron *ZC3H6* | 0 | 0 | 0 | 2 | 1 | 0.02 | 0 | 0 | 5.61 | Yes | *SCL20A1* | 0.0001 |

A (AF): allele (allele frequency); GERP: Genomic evolutionary rate profiling; DNAase HS: DNase hypersensitivity; DNase foot: DNase footprint; Ernst score: classes of chromatin states (recurrent combinations of chromatin marks); MAF: minor allele frequency; TFBS: transcription factor binding site. Gene names: *FBLN7*: Fibulin 7; *ZC3H8*: Zinc Finger CCCH-Type Containing 8; *ZC3H6*: Zinc Finger CCCH-Type Containing 6.

**Table 3. Correlation between genotypes for potentially functional SNP and bone-specific expression of genes in the candidate region**

| RANK | SNP | GENE | PROBE | A1 | A2 | FRQ | BETA | SE | P |
|---|---|---|---|---|---|---|---|---|---|
| 1 | rs35586251 | *TTL* | 224896_s_at | A | G | 0.017 | 0.65 | 0.13 | $6.62 \times 10^{-6}$ |
| 2 | rs77172864 | *SLC20A1* | 230494_at | G | A | 0.013 | -0.46 | 0.11 | 0.00011 |
| 4 | rs77996972 | *TTL* | 224896_s_at | T | C | 0.012 | 0.67 | 0.13 | $3.80 \times 10^{-6}$ |
| | | *SLC20A1* | 230494_at | T | C | 0.012 | -0.49 | 0.11 | $5.50 \times 10^{-5}$ |
| 5 | rs75814334 | *TTL* | 224896_s_at | T | C | 0.013 | 0.67 | 0.13 | $2.10 \times 10^{-6}$ |
| | | *SLC20A1* | 230494_at | T | C | 0.013 | -0.48 | 0.11 | $6.60 \times 10^{-5}$ |
| 6 | rs74792868 | *TTL* | 224896_s_at | A | G | 0.012 | 0.66 | 0.14 | $2.00 \times 10^{-5}$ |
| | | *SLC20A1* | 230494_at | A | G | 0.012 | -0.53 | 0.12 | $2.80 \times 10^{-5}$ |
| 6 | rs72943913 | *SLC20A1* | 230494_at | G | A | 0.013 | -0.46 | 0.11 | 0.00011 |
| 7 | rs112275607 | *TTL* | 224896_s_at | A | G | 0.013 | 0.67 | 0.13 | $2.80 \times 10^{-6}$ |
| | | *SLC20A1* | 230494_at | A | G | 0.013 | -0.48 | 0.11 | $6.02 \times 10^{-5}$ |
| 8 | rs113085288 | *SLC20A1* | 230494_at | T | A | 0.008 | -0.72 | 0.14 | $4.06 \times 10^{-6}$ |
| 9 | rs113428223 | *SLC20A1* | 230494_at | T | C | 0.013 | -0.46 | 0.11 | 0.0001 |

The data shown are only for the associations which were significant after Bonferroni correction (p value for significance $\leq$0.0002). A1: allele 1, A2: Allele 2, FRQ: frequency of allele 1, BETA: effect size on regression analysis referred to A1 allele, SE: standard error of beta estimate, probe IDs obtained from the Affymetrix HG U133 2.0 plus array. Gene names: *TTL*: Tubulin Tyrosine Ligase; *SLC20A1*: Solute Carrier Family 20 Member 1 (also known as *PIT1*).

The top ranking variant identified by SuRFR, rs35586251, located within exon 3 of *FBLN7* is a non-synonymous substitution (p.Val119Met). However, analysis using various *in silico* software tools yielded inconsistent results with regard to functionality of this SNP at the protein level (Supplementary Table 6). The other 9 SNPs are associated with expression of *TTL*, *SCL20A* or both genes. The variant that ranked top by SuRFR, rs35586251, was associated with increased expression of *TTL* ($p=6.6x10^{-6}$). Four other variants were also associated with both increased expression of *TTL* and reduced expression of *SLC20A1* (p-values ranging from $2.1x10^{-6}$ to $10^{-5}$). The second ranking variant, rs77172864, in strong LD with the GWAS top hit ($r^2=0.79$), was associated with reduced expression of *SLC20A1* ($p=10^{-4}$) (Tables 2 and 3).

The variants listed on Table 2 were tested in the UK Biobank cohort for further association with clinical vertebral fractures (Supplementary Table 7). Although none of them was significantly associated with the trait, a trend of significance was found for SNPs rs72943913, rs77172864, and rs113428223 (p=0.06, OR=1.66), and all of them identified as eQTLs for *SLC20A1* gene in bone. These variants showed a lower frequency (MAF=0.03) than the top hit (MAF=0.05), which could require a greater sample size to detect associations with the trait.

### 3.5. Association between clinical vertebral fractures and other osteoporosis related phenotypes

In order to determine if there is overlap between the SNPs identified as associated with lumbar spine BMD in previous GWAS with those associated with clinical vertebral fracture in this study, we evaluated 50 SNPs that have been associated with lumbar spine BMD at a genome-wide significant level in previous studies in our dataset.[10,11,13,30,31] Four variants were nominally associated with clinical vertebral fracture after Bonferroni correction (Table 4). We also analysed 15 variants previously associated with clinical fracture,[13] of which three were associated with clinical vertebral fractures in this study. We also analysed the SNPs identified by Nielson and colleagues[27] as genome-wide significant predictors of volumetric vertebral bone mineral density for association with clinical vertebral fractures in our dataset. Of the six genome-wide significant SNPs identified by Nielson et al, we found that one was significantly associated with clinical vertebral fractures after Bonferroni correction (rs12742784, $p=6.24x10^{-5}$). The BMD-increasing variants in Table 4 conferred a reduced risk of clinical vertebral fractures in our study, whilst the variants associated with appearance of clinical fractures in previous studies were also associated with a higher risk of developing a clinical vertebral fracture in our data.

**Table 4. Association between known genetic determinants of spine BMD and clinical vertebral fractures in the combined GWAS dataset.**

| **Previous studies** | | | | | | | | | **Present study** | |
|---|---|---|---|---|---|---|---|---|---|---|
| Study | SNP | Locus | Candidate gene | Phenotype | Method | Allele | Beta[1] | p | Beta[2] | p |
| Estrada | rs1346004 | 2q24.3 | *GALNT3* | LS-BMD | DXA | A | -0.06 | $3.87x10^{-30}$ | +0.16 | 0.0002 |
| Estrada | rs4727338 | 7q21.3 | *SLC25A13* | LS-BMD | DXA | C | +0.07 | $2.13x10^{-35}$ | -0.15 | 0.0004 |
| Estrada | rs6426749 | 1p36.12 | *ZBTB40* | LS-BMD | DXA | C | +0.1 | $1.86x10^{-44}$ | -0.22 | 0.0003 |
| Styrkarsdottir | rs7524102 | 1p36 | *WNT4* | LS-BMD | DXA | A | -0.11 | $9.2x10^{-9}$ | +0.23 | 0.0002 |
| Estrada | rs4727338 | 7q21.3 | *SLC25A13* | Clinical fracture | Clinical records and X-rays | G | +0.08 | $5.9x10^{-11}$ | +0.14 | 0.0004 |
| Estrada | rs6426749 | 1p36.12 | *ZBTB40* | Clinical fracture | Clinical records and X-rays | G | +0.07 | $3.6x10^{-6}$* | +0.22 | 0.0003 |
| Estrada | rs6959212 | 7p14.1 | *STARD3NL* | Clinical fracture | Clinical records and X-rays | T | +0.05 | $7.2x10^{-5}$* | +0.15 | 0.001 |
| Nielson | rs12742784 | 1p36.12 | *ZBTB40* | Vertebral BMD | qCT imaging | T | +0.09 | $1.05x10^{-10}$ | -0.20 | $6.24x10^{-5}$ |

The variants shown are those that were significant after Bonferroni correction for testing 56 BMD variants (p threshold for association 0.0009) and 16 fracture variants (p threshold for association 0.003). *SNP significantly associated with clinical fracture after Bonferroni correction (p threshold at Estrada et al $5x10^{-4}$).

Beta[1] showed the effect for the previous studies (LS-BMD, clinical fracture and vertebral BMD).

Beta[2] showed the effect for the present study on clinical vertebral fracture

Gene names: *GALNT3*: Polypeptide N-Acetylgalactosaminyltransferase 3); *SLC25A13*: Solute Carrier Family 25 Member 13; *ZBTB40*: Zinc Finger And BTB Domain Containing 40; *WNT4*: Wnt Family Member 4; *STARD3NL*: StAR Related Lipid Transfer Domain Containing 3 N-Terminal Like).

Method column shows the technique used to evaluate the BMD or assess the fracture (DXA: dual energy X-ray absorptiometry, CT: quantitative computerised tomography)

## 4. DISCUSSION

Many advances have been made in defining the genetic determinants of bone mineral density and fractures through large scale genome-wide association studies, genome sequencing studies and linkage studies in rare bone diseases.[32] For example, linkage studies have shown that loss-of-function and gain-of-function variants in *LRP5* cause early onset osteoporosis[33] and high bone mass[34] respectively, whereas loss of function mutations affecting *SOST* and *LRP4* have been identified as causes of high bone mass and osteosclerosis.[35,36] Genome-wide association studies and genome sequencing studies have also been successful in identifying multiple loci that regulate bone mineral density[9-11,30,37] and a smaller number that predispose to clinical fractures.[10,30]

Although vertebral fractures are one of the most common and important complications of osteoporosis, relatively little is known about the genetic determinants of this type of fracture.[38] In a previous study of 8,717 cases and 21,793 controls, Oei and colleagues failed to identify any locus with significant evidence of association with morphometric vertebral fractures.[21] In the present study however, we were successful in identifying one genome-wide significant variant that predisposed to clinical vertebral fractures, which was replicated in several populations. We also detected loci that might play a role in clinical vertebral fractures (showing suggestive association at the genome-wide level), but further studies need to be performed in further cohorts to confirm or refute these associations. A likely reason for the difference between our findings and those of Oei et al, is varying case definition. Here, we studied patients with clinical vertebral fractures as opposed to morphometric vertebral deformities, many of which may not be true fractures.[22] The genome-wide significant SNP identified in the present study, rs10190845, shows one of the largest effect size so far detected in the field of osteoporosis genetics (OR=1.75[1.45-2.12]). Most of the signals associated with BMD or fracture to date showed a very low effect (ORs between 0.90 and 1.10),[12,13] with a few exceptions.[20]

rs10190845 maps to chromosome 2q13, a region previously associated with low femoral neck bone density.[10] However, when conditioning on rs17040773, the previously reported top SNP at the locus,[10] the association with rs10190845 remained significant, indicating that rs10190845 represents a novel signal.

In order to determine if there was an overlap between the results of this study and those previously reported, we analysed 71 SNPs that have previously been associated with either spine BMD or clinical fractures and identified seven variants that were significantly associated with clinical vertebral fracture in this study, after Bonferroni correction (threshold for

significance 0.0009 for BMD and 0.003 for clinical fractures). However, the association for these variants did not reach genome-wide significance, therefore, they were not selected in the GWAS analysis. The SNPs associated with low BMD as well as increased risk of clinical fractures in previous studies were associated with an increased risk of clinical vertebral fractures in this study and those associated with an increased risk of clinical fractures in previous studies were associated with an increased risk of clinical vertebral fractures in this study.

Furthermore, when we analysed six SNPs that were significantly associated with vertebral bone mineral density on quantitative computerised tomography (qCT) analysis[27] one locus on chromosome 1p36, close to *ZBTB40*, was identified and significantly associated with clinical vertebral fracture in this study. These results support the importance of *ZBTB40* as a predictor of clinical fractures and suggest that the mechanism of association is most probably mediated by changes in BMD. The observations in this study, when taken together with the findings of Nielson and Estrada[10,27] indicate that there is a partial overlap between loci that regulate lumbar spine BMD, and clinical vertebral fractures. However, there are some genetic determinants of clinical vertebral fracture which are unique and which operate independently of BMD.

In order to identify the mechanisms by which 2q13 predisposes to vertebral fracture we conducted bioinformatics analyses to determine if rs10190845 or other SNPs nearby were likely to be functional variants. These studies identified several potentially functional SNPs in the same LD block as rs10190845, which might account for the association we observed. The top ranking SNP from SuRFR analysis was rs35586251, which was strongly associated with expression of the *TTL* gene within the candidate locus (Supplementary Figure 8). However, the second ranking SNP, rs77172864 (Supplementary Figure 9), in strong LD with the GWAS top hit, was significantly associated with the expression of *SLC20A1*. Several other SNPs were also significantly associated with expression of *TTL* and/or *SLC20A1*, raising the possibility that alterations in expression of one or both genes might account for the predisposition to clinical vertebral fractures. Association analysis performed using UK Biobank cohort for these SNPs showed a trend of association for markers regulating *SLC20A1* gene, which also showed some degree of linkage disequilibrium, with the GWAS top hit. The lack of significant association might be due to their low allele frequency (MAF=0.03), which means that a larger sample size may be required to detect a strong association. The Tubulin Tyrosine Ligase encoded by *TTL* is involved in regulation of the cytoskeleton. Previous studies have shown that *TTL* is involved in neuronal development[39] and injury signalling,[40] raising the possibility that variants that

regulate *TTL* might be involved in regulating pain perception, which could account for the fact that predisposing variants have not previously been associated with BMD. Other mechanisms might also be possible and further studies need to be performed in order to address the role of *TTL* in clinical vertebral fracture. The other main candidate gene, *SLC20A1*, encodes Pit1, which facilitates the entry of inorganic phosphate into the cytoplasm.[41] Previous studies have shown that *SLC20A1* is involved in mineralisation.[42-45] Altered expression of this gene could convey risk for vertebral fractures via an effect on bone mineralisation. Although *SLC20A1* presents as the candidate gene for association with clinical vertebral fractures in this study, it has not been identified previously as a predictor of BMD or fractures. This opens for alternative mechanisms, or that *TTL* rather than *SLC20A1* is the candidate gene within the 2q13 locus.

Limitations of the study include the fact that the total sample size was relatively small and the power to detect alleles of modest effect size was limited. It is possible that we may have missed associations between rare variants and clinical vertebral fractures since the imputation we performed was against HapMap reference panel rather than larger panels that increase imputation power particularly against low frequency variants. Although case definition was clinically based, there was no significant heterogeneity in the associations we observed across centres.

Strengths of the present study are that it has provided important new information on the genetic determinants of clinical vertebral fracture and that results, despite the sample size, have been validated in two independent replication stages.

### 4.1. Conclusion

Genome wide association analysis identified a significant association between a marker on chromosome 2 and clinical vertebral fractures in postmenopausal women, a finding validated in several independent populations.

It is of interest that the top hit and other suggestive hits identified acted independently of BMD, bringing to attention other bone microarchitectural modalities that determine fracture susceptibility. This suggests that the variants identified might be acting as markers for perception of pain or other factors that are associated with the clinical presentation of vertebral fractures. We also found that some of the variants previously identified as regulators of spine BMD were associated with clinical vertebral fractures, but with effects that were weaker than the top hit and other suggestive hits. Taken together, the data suggest that the genetic basis of clinical vertebral fracture is complex involving variants that act independently of BMD as well

as those that are associated with spine BMD. Further research is now warranted to fully investigate the mechanisms involved.

## 5. ACKNOWLEDGMENTS

The authors are grateful to the patients and controls from the different centres who agreed to participate in this study. We thank Ms Dilruba Kabir, at the Rheumatology and Bone Disease Unit, CGEM-IGMM, Edinburgh, UK, Mr Matt Sims at the MRC Epidemiology Unit, University of Cambridge, UK, Ms Mila Jhamai and Ms Sarah Higgins, at the Genetics Laboratory of Erasmus MC, Rotterdam, The Netherlands, and Ms Johanna Hadler, Ms Kathryn A. Addison, and Ms Karena Pryce of the University of Queensland Centre for Clinical Genomics, Brisbane, Australia, for technical support on the genotyping stage; and Mr Marijn Verkerk and Dr Anis Abuseiris at the Genetics Laboratory of Erasmus MC, Rotterdam, for assistance on the data analysis. We also thank Prof Nick Gilbert and Dr Giovanny Rodriguez-Blanco for their comments and advice on the manuscript preparation.

This research was performed within the Genetic Factors for Osteoporosis (GEFOS) consortium, funded by the European Commission (HEALTH-F2-2008-201865-GEFOS).

This study makes use of data generated by the Wellcome Trust Case Control Consortium. A full list of the investigators who contributed to the generation of the data is available from www.wtccc.org.uk. Funding for the project was provided by the Wellcome Trust under award 076113.

ORCADES was supported by the Chief Scientist Office of the Scottish Government (CZB/4/276, CZB/4/710), the Royal Society, the MRC Human Genetics Unit, Arthritis Research UK and the European Union framework program 6 EUROSPAN project (contract no. LSHG-CT-2006-018947). DNA extractions were performed at the Wellcome Trust Clinical Research Facility in Edinburgh. We would like to acknowledge the invaluable contributions of Lorraine Anderson and the research nurses in Orkney, the administrative team in Edinburgh and the people of Orkney.

CABRIO was supported by the Instituto de Salud Carlos III and Fondos FEDER from the EU (PI 11/1092 and PI12/615).

The AOGC study was funded by the Australian National Health and Medical Research Council (Project grant 511132).

Lothian Birth Cohort 1921 phenotype collection was supported by the UK's Biotechnology and Biological Sciences Research Council (BBSRC), The Royal Society, and The Chief Scientist Office of the Scottish Government. Phenotype collection in the Lothian Birth Cohort 1936 was supported by Age UK (The Disconnected Mind project). Genotyping of the cohorts

was funded by the BBSRC. The work was undertaken by the University of Edinburgh Centre for Cognitive Ageing and Cognitive Epidemiology, part of the cross council Lifelong Health and Wellbeing Initiative (MR/K026992/1). Funding from the BBSRC and Medical Research Council (MRC) is gratefully acknowledged.

Research work on Slovenian case and control samples was funded by Slovenian Research Agency (project no. P3-0298 and J3-2330).

The Danish National Birth Cohort (DNBC) is a result of major grants from the Danish National Research Foundation, the Danish Pharmacists' Fund, the Egmont Foundation, the March of Dimes Birth Defects Foundation, the Augustinus Foundation, and the Health Fund of the Danish Health Insurance Societies. The DNBC biobank is a part of the Danish National Biobank resource, which is supported by the Novo Nordisk Foundation. Dr Bjarke Feenstra is supported by an Oak Foundation Fellowship.

The Framingham Study was funded by grants from the US National Institute for Arthritis, Musculoskeletal and Skin Diseases and National Institute on Aging (R01 AR 41398 and R01 AR061162; 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.

The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

**Author contribution:** Study conception: SHR, NA, AGU, FR; data collection: NA, OMEA, LM, MLB, PR, AD, JMO, CV, JC, JAR, LBH, BLL, MAB, ELD, SK, K-TK, RU-M, JdP, RG-S, JRL, RLP, PD'A, NG-G, XN, SM-B, JM, OW, JE, BF, MM, KES, PN, JFW, GD, JS, ID, TT, LF, FG, LG, GL, RE; Genotyping: AGU, FR; data analysis: NA, OMEA, SR, OKO, KMG, NMR, KLE, CMN, H-YH, DK, GM, EN, VE, XL, BF, MM, KE, LH, LO, CM-G; drafting of the manuscript: NA and SHR; all authors contributed to critically review the article and

approved the final manuscript. NA, SR, CMN, EN and NMR takes responsibility for the data analysis.

## REFERENCES


1. van Staa TP, Dennison EM, Leufkens HG, et al. Epidemiology of fractures in England and Wales. *Bone* 2001;29:517-22.
2. Ismail AA, O'Neill TW, Cooper C, et al. Mortality associated with vertebral deformity in men and women: results from the European Prospective Osteoporosis Study (EPOS). *Osteoporos Int* 1998;8:291-7.
3. Cauley JA, Palermo L, Vogt M, et al. Prevalent vertebral fractures in black women and white women. *J Bone Miner Res* 2008;23:1458-67.
4. Fink HA, Milavetz DL, Palermo L, et al. What proportion of incident radiographic vertebral deformities is clinically diagnosed and vice versa? *J Bone Miner Res* 2005;20:1216-22.
5. Black DM, Delmas PD, Eastell R, et al. Once-yearly zoledronic acid for treatment of postmenopausal osteoporosis. *N Engl J Med* 2007;356:1809-22.
6. Cummings SR, San MJ, McClung MR, et al. Denosumab for prevention of fractures in postmenopausal women with osteoporosis. *N Engl J Med* 2009;361:756-65.
7. Sambrook P, Cooper C. Osteoporosis. *Lancet* 2006;367:2010-8.
8. Cauley JA, Thompson DE, Ensrud KC, et al. Risk of mortality following clinical fractures. *Osteoporos Int* 2000;11:556-61.
9. Zhang L, Choi HJ, Estrada K, et al. Multistage genome-wide association meta-analyses identified two new loci for bone mineral density. *Hum Mol Genet* 2014;23:1923-33.
10. Estrada K, Styrkarsdottir U, Evangelou E, et al. Genome-wide meta-analysis identifies 56 bone mineral density loci and reveals 14 loci associated with risk of fracture. *Nat Genet* 2012;44:491-501.
11. Kung AW, Xiao SM, Cherny S, et al. Association of JAG1 with bone mineral density and osteoporotic fractures: a genome-wide association study and follow-up replication studies. *Am J Hum Genet* 2010;86:229-39.
12. Rivadeneira F, Styrkarsdottir U, Estrada K, et al. Twenty bone-mineral-density loci identified by large-scale meta-analysis of genome-wide association studies. *Nat Genet* 2009;41:1199-206.
13. Duncan EL, Danoy P, Kemp JP, et al. Genome-wide association study using extreme truncate selection identifies novel genes affecting bone mineral density and fracture risk. *PLoS Genet* 2011;7:e1001372.

14. Richards JB, Kavvoura FK, Rivadeneira F, et al. Collaborative meta-analysis: associations of 150 candidate genes with osteoporosis and osteoporotic fracture. *Ann Intern Med* 2009;151:528-37.

15. Styrkarsdottir U, Halldorsson BV, Gretarsdottir S, et al. New sequence variants associated with bone mineral density. *Nat Genet* 2009;41:15-7.

16. Styrkarsdottir U, Halldorsson BV, Gretarsdottir S, et al. Multiple genetic loci for bone mineral density and fractures. *N Engl J Med* 2008;358:2355-65.

17. Xiong DH, Liu XG, Guo YF, et al. Genome-wide association and follow-up replication studies identified ADAMTS18 and TGFBR3 as bone mass candidate genes in different ethnic groups. *Am J Hum Genet* 2009;84:388-98.

18. Hsu YH, Zillikens MC, Wilson SG, et al. An integration of genome-wide association study and gene expression profiling to prioritize the discovery of novel susceptibility Loci for osteoporosis-related traits. *PLoS Genet* 2010;6:e1000977.

19. Zheng HF, Forgetta V, Hsu YH, et al. Whole-genome sequencing identifies EN1 as a determinant of bone density and fracture. *Nature* 2015;526:112-7.

20. Styrkarsdottir U, Thorleifsson G, Sulem P, et al. Nonsense mutation in the LGR4 gene is associated with several human diseases and other traits. *Nature* 2013;497:517-20.

21. Oei L, Estrada K, Duncan EL, et al. Genome-wide association study for radiographic vertebral fractures: a potential role for the 16q24 BMD locus. *Bone* 2014;59:20-7.

22. Ferrar L, Jiang G, Adams J, et al. Identification of vertebral fractures: an update. *Osteoporos Int* 2005;16:717-28.

23. Cooper C, Atkinson EJ, O'Fallon WM, et al. Incidence of clinically diagnosed vertebral fractures: a population-based study in Rochester, Minnesota, 1985-1989. *J Bone Miner Res* 1992;7:221-7.

24. Edwards BJ, Haynes C, Levenstien MA, et al. Power and sample size calculations in the presence of phenotype errors for case/control genetic association studies. *BMC Genet* 2005;6:18.

25. Albagha OM, Visconti MR, Alonso N, et al. Genome-wide association study identifies variants at CSF1, OPTN and TNFRSF11A as genetic risk factors for Paget's disease of bone. *Nat Genet* 2010;42:520-4.

26. The Wellcome Trust Case Control Consortium. Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls. *Nature* 2007;447:661-78.

27. Nielson CM, Liu CT, Smith AV, et al. Novel Genetic Variants Associated With Increased Vertebral Volumetric BMD, Reduced Vertebral Fracture Risk, and Increased Expression of SLC1A3 and EPHB2. *J Bone Miner Res* 2016;31:2085-97.

28. Reppe S, Sachse D, Olstad OK, et al. Identification of transcriptional macromolecular associations in human bone using browser based in silico analysis in a giant correlation matrix. *Bone* 2013;53:69-78.

29. Ryan NM, Morris SW, Porteous DJ, et al. SuRFing the genomics wave: an R package for prioritising SNPs by functionality. *Genome Med* 2014;6:79.

30. Zheng HF, Forgetta V, Hsu YH, et al. Whole-genome sequencing identifies EN1 as a determinant of bone density and fracture. *Nature* 2015.

31. Styrkarsdottir U, Thorleifsson G, Sulem P, et al. Nonsense mutation in the LGR4 gene is associated with several human diseases and other traits. *Nature* 2013.

32. Alonso N, Ralston SH. Unveiling the mysteries of the genetics of osteoporosis. *J Endocrinol Invest* 2014;37:925-34.

33. Gong Y, Slee RB, Fukai N, et al. LDL Receptor-Related Protein 5 (LRP5) Affects Bone Accrual and Eye Development. *Cell* 2001;107:513-23.

34. Little RD, Carulli JP, Del Mastro RG, et al. A mutation in the LDL receptor-related protein 5 gene results in the autosomal dominant high-bone-mass trait. *Am J Hum Genet* 2002;70:11-9.

35. Balemans W, Ebeling M, Patel N, et al. Increased bone density in sclerosteosis is due to the deficiency of a novel secreted protein (SOST). *Hum Mol Genet* 2001;10:537-43.

36. Leupin O, Piters E, Halleux C, et al. Bone overgrowth-associated mutations in the LRP4 gene impair sclerostin facilitator function. *J Biol Chem* 2011;286:19489-500.

37. Reppe S, Wang Y, Thompson WK, et al. Genetic Sharing with Cardiovascular Disease Risk Factors and Diabetes Reveals Novel Bone Mineral Density Loci. *PLoS One* 2015;10:e0144531.

38. Liu CT, Karasik D, Zhou Y, et al. Heritability of prevalent vertebral fracture and volumetric bone mineral density and geometry at the lumbar spine in three generations of the Framingham study. *J Bone Miner Res* 2012;27:954-8.

39. Marcos S, Moreau J, Backer S, et al. Tubulin tyrosination is required for the proper organization and pathfinding of the growth cone. *PLoS One* 2009;4:e5405.

40. Song W, Cho Y, Watt D, et al. Tubulin-tyrosine Ligase (TTL)-mediated Increase in Tyrosinated alpha-Tubulin in Injured Axons Is Required for Retrograde Injury Signaling and Axon Regeneration. *J Biol Chem* 2015;290:14765-75.

41. Saier MH, Jr. A functional-phylogenetic classification system for transmembrane solute transporters. *Microbiol Mol Biol Rev* 2000;64:354-411.

42. Guicheux J, Palmer G, Shukunami C, et al. A novel in vitro culture system for analysis of functional role of phosphate transport in endochondral ossification. *Bone* 2000;27:69-74.

43. Wang D, Canaff L, Davidson D, et al. Alterations in the sensing and transport of phosphate and calcium by differentiating chondrocytes. *J Biol Chem* 2001;276:33995-4005.

44. Palmer G, Bonjour JP, Caverzasio J. Expression of a newly identified phosphate transporter/retrovirus receptor in human SaOS-2 osteoblast-like cells and its regulation by insulin-like growth factor I. *Endocrinology* 1997;138:5202-9.

45. Palmer G, Guicheux J, Bonjour JP, et al. Transforming growth factor-beta stimulates inorganic phosphate transport and expression of the type III phosphate transporter Glvr-1 in chondrogenic ATDC5 cells. *Endocrinology* 2000;141:2236-43.

**Fig 1. Cohort specific association between rs10190845 and clinical vertebral fracture**

The point estimates (squares) and 95% confidence intervals (horizontal lines) for individual studies are shown with the summary indicated by the diamond using a fixed effect model. Summaries are shown for meta-analysis with discovery cohorts only (Summary_discovery), with the first replication cohorts only (Summary_replication), and for the whole 3-stage meta-analysis (Summary_meta-analysis). "BRITISH-WTCCC" shows the results for the combined cohorts CAIFOS, AOGC, DOES, and EPIC, and the control cohort WTCCC2. "Scottish replication" corresponds to EDOS-ORCADES cohorts, "Italian_replication_1" study corresponds to Florence-InCHIANTI cohorts and "Italian_replication_2" study comprises the Turin and Siena cohorts. Cohort sizes are reflected by square dimensions.

**Fig 2. Regional association plots of susceptibility locus for clinical vertebral fracture**

The figure shows the results after imputation using 1000G v3 as reference panel. The SNPs are colour coded according to the extent of LD with the SNP showing the highest association signal from the combined analysis (represented as a purple diamond). The estimated recombination rates (cM/Mb) from HapMap CEU release 22 are shown as light blue lines, and the blue arrows represent known genes in the region. The red line shows the threshold for genome-wide significance ($p = 5 \times 10^{-8}$)