Bone 123 (2019) 39–47

Contents lists available at ScienceDirect

Bone

journal homepage: www.elsevier.com/locate/bone

Full Length Article

Functional characterization of the *C7ORF76* genomic region, a prominent GWAS signal for osteoporosis in 7q21.3

Neus Roca-Ayats[a], Núria Martínez-Gil[a], Mónica Cozar[a], Marina Gerousi[a], Natàlia Garcia-Giralt[b], Diana Ovejero[c], Leonardo Mellibovsky[b], Xavier Nogués[b], Adolfo Díez-Pérez[b], Daniel Grinberg[a,1], Susanna Balcells[a,*,1]

[a] Department of Genetics, Microbiology and Statistics, Facultat de Biologia, Universitat de Barcelona, Centro de Investigación Biomédica en Red de Enfermedades Raras (CIBERER), ISCIII, IBUB, IRSJD, Barcelona, Catalonia, Spain
[b] Musculoskeletal Research Group, IMIM (Hospital del Mar Medical Research Institute), Centro de Investigación Biomédica en Red en Fragilidad y Envejecimiento Saludable (CIBERFES), ISCIII, Barcelona, Catalonia, Spain
[c] National Research Council, Institute of Clinical Physiology, Lecce, Italy

ARTICLE INFO

Keywords:
C7ORF76
Osteoporosis
GWAS signal
Non-coding regulatory variant
Enhancer
4C

ABSTRACT

Genome-wide association studies (GWAS) have repeatedly identified genetic variants associated with bone mineral density (BMD) and osteoporotic fracture in non-coding regions of *C7ORF76,* a poorly studied gene of unknown function. The aim of the present study was to elucidate the causality and molecular mechanisms underlying the association. We re-sequenced the genomic region in two extreme BMD groups from the BARCOS cohort of postmenopausal women to search for functionally relevant variants. Eight selected variants were tested for association in the complete cohort and 2 of them (rs4342521 and rs10085588) were found significantly associated with lumbar spine BMD and nominally associated with osteoporotic fracture. *cis*-eQTL analyses of these 2 SNPs, together with SNP rs4727338 (GWAS lead SNP in Estrada et al., *Nat Genet.* 44:491–501, 2012), performed in human primary osteoblasts, disclosed a statistically significant influence on the expression of the proximal neighbouring gene *SLC25A13* and a tendency on the distal *SHFM1*. We then studied the functionality of a putative upstream regulatory element (UPE), containing rs10085588. Luciferase reporter assays showed transactivation capability with a strong allele-dependent effect. Finally, 4C-seq experiments in osteoblastic cell lines showed that the UPE interacted with different tissue-specific enhancers and a lncRNA (*LOC100506136*) in the region.

In summary, this study is the first one to analyse in depth the functionality of *C7ORF76* genomic region. We provide functional regulatory evidence for the rs10085588, which may be a causal SNP within the 7q21.3 GWAS signal for osteoporosis.

1. Introduction

Genome-wide association studies (GWAS) have been successfully used to identify genetic variants associated with complex traits and diseases, such as osteoporosis. In a few cases, the associated SNPs are located within a coding region of a gene, facilitating its functional evaluation. However, the vast majority of associated SNPs lie in non-coding regions, which make it challenging to understand the functional mechanisms underlying the association [1,2]. In addition, it is highly probable that the associated SNPs are in linkage disequilibrium (LD) with the causal variant.

To date, many GWAS have been performed to find genetic association with bone mineral density (BMD) and osteoporotic fracture [3–12]. BMD is a genetically determined, extensively measured quantitative trait (heritability of 0.5–0.85) and, therefore, a good marker for bone status. Low-trauma fracture, the clinical outcome of osteoporosis, is also heritable, albeit to a lesser extent (heritability of 0.54–0.68) [13]. These GWAS have identified > 500 candidate *loci* [12], although the causal variants remain largely unknown. In addition, all the GWAS findings together only explain a small proportion (~20%) of the total

* Corresponding author at: Departament de Genètica, Microbiologia i Estadística, Facultat de Biologia, Universitat de Barcelona. Av. Diagonal, 643. 08028 Barcelona, Catalonia, Spain.
E-mail address: sbalcells@ub.edu (S. Balcells).
[1] Last co-authors.

https://doi.org/10.1016/j.bone.2019.03.014
Received 29 January 2019; Received in revised form 4 March 2019; Accepted 12 March 2019
Available online 14 March 2019

genetic impact on BMD [12]. Some associated *loci* contain genes not previously known to play a role in bone biology, which is the case of *C7ORF76* (AKA *FLJ42280*, currently annotated as transcript variant 6 of gene *SEM1* in GRCh38), in the 7q21.3 genomic region. Several SNPs within this region have been found significantly associated with both lumbar spine (LS) and femoral neck (FN) BMD as well as with osteoporotic fracture in different GWAS and meta-analyses [4,6–8,14–16]. In particular, in the largest meta-analysis carried out so far [7], the lead SNP of this *locus* (rs4727338) was one of the genome-wide top-associated signals. Yet, *C7ORF76* is a poorly studied gene of unknown function and there are other genes within this region that could also be responsible for the associations observed.

In this study, we aimed at elucidating the causality and molecular mechanisms underlying the strong association identified in the *C7ORF76* genomic region. We have deeply re-sequenced the genomic region in two extreme BMD groups of postmenopausal women of the BARCOS cohort and selected some variants to analyse their association in the full cohort. Through a combination of several *in silico* and experimental approaches, we studied a possibly causal SNP located in a regulatory region and demonstrated its functionality.

2. Material and methods

2.1. Study cohort

The BARCOS cohort consisted of 1490 postmenopausal women of Spanish descent from the Barcelona area, monitored at the Hospital del Mar (Barcelona, Spain). Exclusion criteria were any history of bone diseases, metabolic or endocrine disorders, hormone-replacement therapy, or use of drugs that could affect bone mass. BMD of all participants was measured at LS and FN by dual energy X-ray absorptiometry (DXA). The following data were also recorded: age, age of menarche and menopause, number of fractures and anthropometric measures such as weight and height. DNA is available from all samples of the cohort. Details of the cohort and DNA extraction have been described previously [17,18]. Written informed consent was obtained from all patients in accordance with the regulations of the Clinical Research Ethics Committee of Parc de Salut Mar, which approved the study. All experiments and protocols were approved by the Bioethics Committee of Universitat de Barcelona (IRB00003099).

For the re-sequencing of the *C7ORF76 locus*, two extreme LS-BMD groups were selected from the BARCOS cohort, using the statistical *Z*-score. The 50 women with the highest Z-score values (from 2.98 to 0.73) were included in the 50-H group and the 50 women with the lowest Z-score values (from −2.41 to −4.26) were included in the 50-L group.

2.2. Re-sequencing

Re-sequencing of the *C7ORF76* (ENSG00000197851; ENST00000356686.1) genomic region was performed in the 100 individuals of the two extreme subgroups of the BARCOS cohort, according to the *Z*-score. A 28 kb region (chr7:96,108,695-96,136,619; GRCh37), including the *C7ORF76* gene and the 3.8 kb upstream and 2 kb downstream regions of the gene, was amplified in 7 overlapping fragments by Long Range-PCR (Supplementary Table 1). All amplicons were purified and quantified using the Quant-iT PicoGreen dsDNA Reagent and Kit (Life Technologies, Thermo Fisher) before pooling them equimolarly into two groups, one of HBM and another of LBM. Both pools were tagged with a MID adaptor and an emulsion-PCR was carried out prior to massive parallel sequencing at 3600× coverage per pool with Roche's 454 GS Junior System. The massive parallel sequencing was carried out in the Genomics facilities of the Universitat de Barcelona. The raw data obtained were processed to trim the MIDs, using a custom pipeline, and were mapped against the reference genome (GRCh37), using the GS Mapper software (Roche). Mapped reads were filtered, sorted and indexed using SAMtools [19]. Single Nucleotide Variants (SNVs) and indels were identified using GATK standard hard filtering parameters [20]. The variants were filtered according to the following criteria: coverage ≥1.000 reads, variants present in ≥1% of the reads per pool and low strand bias. The number of reads of a variant was normalised with its coverage and the variants were classified according to minor allele frequency (MAF): Common (MAF ≥ 5%), and lower frequency variants (MAF < 5%). The variants were validated either by differential digestion with restriction enzymes or by high resolution melting, using the Light Cycler® 480 ResoLight Dye (Roche).

2.3. In silico functional analyses and motif analysis

In silico functional analyses consisted in annotating the European and Iberian MAF of the variants obtained from dbSNP and 1000 Genomes, when available; predicting the pathogenicity of exonic variants, using SIFT [21], PolyPhen [22] and Mutation Taster [23]; and, for the intronic variants, analysing the DNase I hypersensitivity, histone modifications, transcription factor binding, miRNAs binding, *etc*. of the regions of interest. All the *in silico* data was obtained from ENCODE [24], International Human Epigenome Consortium [25], The Roadmap Epigenomics Project, FANTOM5 [26], HaploReg [27], RegulomeDB [28], miRTarBase [29], miRdSNP [30], MirSNP [31], BioMart, and Ensembl and UCSC Genome Browser. All variants were analysed with the Variant Effect Predictor from Ensembl, the Variant Annotation Integrator from UCSC, and FuncPred from the National Institute of Environmental Health Sciences. Transcription factor binding sites prediction considering the different alleles of the variants was done using MatInspector [32] and the Bioconductor "motifbreakR" package (https://bioconductor.org/packages/release/bioc/html/motifbreakR.html) [33] using default method settings (weighted sum) and a p-value cut-off at 5×10^{-5} for SNPs. Super-enhancer data was obtained from a catalogue of super-enhancers in 86 human cell and tissue samples [34] and from SEdb [35].

2.4. SNV genotyping

Genotyping of 8 selected variants in the complete BARCOS cohort was carried out at LGC Genomics (Hoddesdon, UK). In addition, as BARCOS was included in the replication phase of the meta-analysis by Estrada et al. [7], the genotyping results of the SNP rs4727338 were also available. The genotyping of 6% of the samples was performed in duplicate, as a genotyping quality control, and showed a concordance above 99%.

2.5. Linkage disequilibrium analysis

The Haploview software [36] was used to calculate and represent the degree of linkage disequilibrium between the genotyped common variants using the default parameters.

2.6. Cell culture

The human osteosarcoma cell line Saos-2 was used for luciferase reporter assays and 4C-seq assays. It was obtained from the American Type Culture Collection (ATCC® HTB-85™) and grown in Dulbecco's Modified Eagle Medium (DMEM; Sigma-Aldrich), supplemented with 10% Fetal Bovine Serum (FBS; Gibco, Life Technologies) and 1% penicillin/streptomycin (Gibco, Life Technologies), at 37 °C and 5% of CO_2. Human fetal osteoblasts (hFOB) 1.19 cells were used for 4C-seq assays. They were obtained from ATCC (ATCC® CRL-11372™) and grown in DMEM:F12 (1:1) medium without phenol red (Gibco, Life Technologies), supplemented with 10% FBS and 0.3 mg/ml Geneticin (Gibco, Life Technologies), at 34 °C and 5% of CO_2. Human medulla-derived mesenchymal stem cells (MSCs) were also used for 4C-seq

assays. They were kindly provided by Dr. José Manuel Quesada Gómez, from Instituto Maimónides de Investigación Biomédica, Hospital Universitario Reina Sofía, Córdoba, Spain. They were grown in alpha-MEM medium (Gibco, Life Technologies), supplemented with 10% FBS, 1% penicillin/streptomycin and 1× Glutamax (Gibco, Life Technologies), at 37 °C and 5% of CO_2. Human primary osteoblasts (hOB) were used for eQTL assays. They were obtained from trabecular bone of women who underwent knee replacement due to osteoarthritis and who did not have any other pathology that could affect the bone status. Bony tissue was cut up into small pieces, washed in phosphate buffered saline (PBS; Gibco, Life technologies) to remove non-adherent cells, and placed on a 140 mm culture plate. Samples were cultured in DMEM supplemented with 10% FBS, 100 U/ml penicillin/streptomycin, 0.4% fungizone (Gibco, Life Technologies) and 100 μg/ml ascorbic acid (Sigma-Aldrich). DNA and RNA extractions were performed at maximum passage 2. HeLa and HEK293 cell lines were obtained from ATCC (ATCC® CCL-2™ and ATCC® CRL-1573™, respectively) and grown in DMEM supplemented with 10% FBS and 1% penicillin/streptomycin at 37 °C and 5% CO_2.

2.7. Human primary osteoblasts DNA extraction and sequencing

DNA was extracted from cultured hOBs using the Wizard® Genomic DNA Purification Kit (Promega), according to manufacturer's instructions. The concentration of the purified DNA was analysed in a spectrophotometer (Nanodrop). Genotypes for rs4727338, rs10429035, rs12674052, rs4342521, and rs10085588 were assessed by Sanger sequencing using the BigDye® Terminator v3.1 (Applied Biosystems) in the Genomics facilities of Universitat de Barcelona. Primers (Invitrogen, Thermo Fisher) were designed using the Primer3 Input 0.4.0 (Supplementary Table 2).

2.8. Human primary osteoblasts RNA extraction, retrotranscription and qPCR

RNA was extracted from cultured hOBs using the High Pure RNA Isolation kit (Roche), according to manufacturer's instructions. RNA was quantified using a Nanodrop spectrophotometer and retrotranscribed using the High Capacity cDNA Reverse Transcription Kit (Applied Biosystems, Thermo Fisher), according to the specifications of the manufacturer. qPCR was performed using UPL Probes (Roche) on a LightCycler 480 Instrument II (Roche). Expression of *HMBS* was used as a normalizing control, and fold changes (FC) were calculated by relative quantification, using the 2nd derivative method. Primers used to amplify the neighbouring genes of the *C7ORF76 locus* are summarised in Supplementary Table 3.

2.9. Amplification of an UPstream regulatory Element (UPE)-derived transcript

UPE-derived transcript was amplified from HeLa, HEK293, human primary osteoblasts and Saos-2 cDNA (200 ng) by PCR, using the following primers: 5′-CACTTTTTCAAATCCCACCTG-3′ and 5′-TGAGAGC TGCTTAGAAATGGAA-3'. PCR products were run in a 2% agarose gel.

2.10. Luciferase reporter constructs and site-directed mutagenesis

The 750 bp-fragment containing the UPE was PCR-amplified from human genomic DNA using the following primers: 5′-CACTTTTTCAA ATCCCACCTG-3′ and 5′-TGAGAGCTGCTTAGAAATGGAA-3′, and cloned in both orientations using *Xho*I and *Kpn*I restriction enzymes in the pGL3-Basic vector (Promega). The minor allele of the rs10085588 (A) was introduced with the QuikChange Lightning Site-Directed Mutagenesis Kit (Agilent), following the manufacturer instructions. All the plasmids were validated by Sanger sequencing.

2.11. In vitro luciferase assay

Saos-2 cells were seeded at a density of 3.0×10^5 cells per well in a 6-well plate. After 24 h, they were transfected with 2.2 μg of total DNA per well using FuGENE HD reagent (Promega), according to manufacturer's instructions. Two plasmids were cotransfected in each well: the pGL3-Basic empty or with the UPE fragment cloned upstream of the Firefly Luciferase coding region and the pRL-TK plasmid, containing the Renilla Luciferase gene, in a proportion of 1/10. Forty-eight hours after transfection, cells were rinsed with PBS and lysed. The luciferase activity was measured using a Glomax Multi+ luminometer (Promega), with the Dual-Luciferase® Reporter Assay System reagents (Promega). Each experiment was performed in two biological replicates and was repeated 5 times.

2.12. 4C-seq

4C-seq was carried out at the Functional Genomics Service of the Centro Andaluz de Biología del Desarrollo (Sevilla, Spain). 4C-seq libraries were generated from Saos-2, hFOB 1.19 and hMSCs lines as described previously [37,38]. 4-bp cutters were used as primary (*Dpn*II) and secondary (Csp6I) restriction enzymes. For each cell line, a total of 1.6 μg of library was amplified by PCR (primers used: CTGGAAGAGT CCCAGGGATC and AATGGAAGAGTGGAGATTCAGG; chr7:96,137,244-96,137,535). Samples were sequenced with Illumina Hi-Seq technology according to standard protocols at the Genomics Service of the Centro Nacional de Investigaciones Cardiovasculares (Madrid, Spain). 4C-seq data were analysed as described previously [39]. Briefly, raw sequencing data were demultiplexed and mapped to the corresponding reference genome (GRCh37). Reads located in fragments flanked by two restriction sites of the same enzyme, in fragments smaller than 40 bp or within a window of 10 kb around the viewpoint were filtered out. 4C-seq data were normalised by the total weight of reads within ± 2 Mb around the viewpoint. The experiments were carried out in one biological replicate.

2.13. Topologically associating domain (TAD) analysis

TAD data on different cell types from Dixon et al. [40] was collected from the 3D Genome Browser (http://promoter.bx.psu.edu/hi-c/) [41] and displayed using the UCSC Genome Browser. The 3D Genome Browser was also used to visualise published Hi-C data.

2.14. Statistical methods

Statistical analyses were performed using the R software version 3.4.1. Hardy-Weinberg equilibrium (HWE) was calculated using Chi-square test and p-values < 0.01 were considered significant. Fisher's exact test with Bonferroni correction for multiple testing was used to statistically compare the differences of the genotype frequency of the common variants in each extreme group (HBM and LBM). Linear regression analysis adjusted by years since menopause (YSM) was performed to determine the association between BMD and the genotype of each SNV in the complete BARCOS cohort. Linear regression was also used to assess the association between gene expression levels and genotypes (*cis*-eQTL) in primary osteoblasts. All analyses were performed using the *SNPassoc* package testing the additive, recessive and dominant models. p-values < 0.05 were considered significant. Correction for multiple testing was performed using the Bonferroni's method for the number of SNPs tested.

Relative luciferase units (RLU, *i.e.* the ratio of the firefly luciferase activity over the Renilla luciferase activity) were calculated for each individual measurement and a one-way blocked ANOVA with TukeyHSD post-hoc test was performed. All the data was ascertained for normality, homoscedasticity and atypical data points and p-values < 0.05 were considered significant.

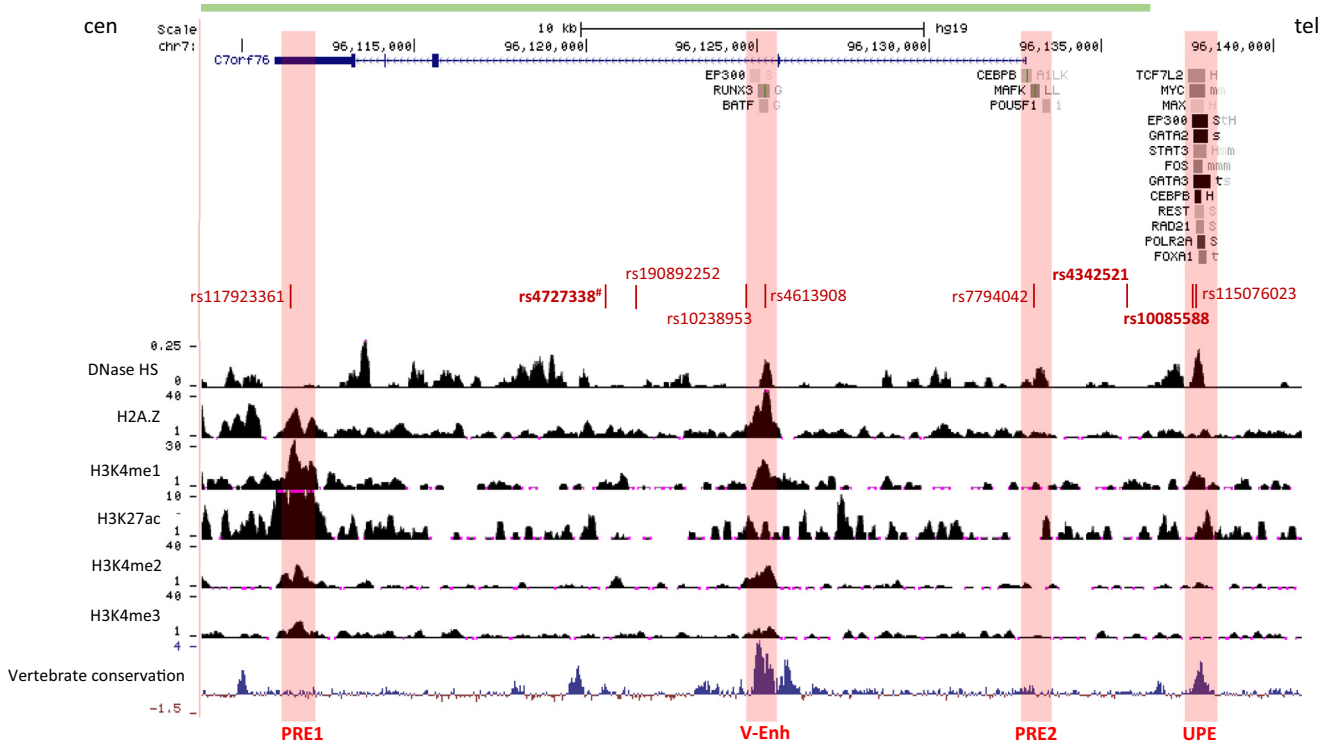

Fig. 1. Schema of *C7ORF76 locus*. In green, region re-sequenced in the present study. In dark red, SNPs genotyped in the BARCOS cohort and in bold, SNPs found associated with BMD; # SNP previously genotyped (Estrada et al. [7]). Transcription factor ChIP-seq, DNase HS, H2A.Z, H3K4me1, H3K27ac, H3K4me2 and H3K4me3 ChIP-seq from osteoblasts (ENCODE), and vertebrate conservation are shown. The red boxes represent putative regulatory elements, according to epigenetic marks as well as transcription factors binding. Note that V-Enh corresponds to an enhancer (hs2311) described in VISTA Enhancers database [42]. PRE1 = putative regulatory element 1; PRE2 = putative regulatory element 2; V-Enh = VISTA enhancer hs2311; UPE = 4 kb upstream putative regulatory element.

3. Results

3.1. Re-sequencing of C7ORF76 locus in extreme BMD groups of the BARCOS cohort and variant functional annotation

We re-sequenced 28 kb of the *C7ORF76* genomic region (chr7:96,108,695-96,136,619; GRCh37) in the 50 women with the highest and 50 women with the lowest LS-BMD of the BARCOS cohort (Fig. 1). Total number and location of single nucleotide variants (SNVs) detected before and after filtering and validating are shown in Table 1. Fifty-one common variants (MAF above 5%) and 59 lower frequency variants (MAF below 5%) were identified. The lower frequency variants were equally distributed between both extreme groups. To better assess the importance of the variants, we explored their functionality using publicly available *in silico* data (transcription factor binding, DNase I hypersensitivity, conservation, miRNAs binding, and histone marks). Twenty-eight variants were found in putative regulatory regions, 4 of which were located in osteoblast regulatory regions (rs9785005, rs10238953, rs4613908 and rs117923361) and 1 of them (rs4613908) was found in an active osteoblast enhancer. One missense variant was also identified but it was predicted to be tolerated by SIFT, to be benign by Polyphen and to be a polymorphism according to MutationTaster. None of the variants was predicted to affect miRNAs binding. Only 1 variant (rs4342521) showed nominal significance between the genotype frequencies of the two extreme groups (Fisher's exact test, p = 0.0382). This SNP is located 3 kb upstream of the *C7ORF76* gene and the minor allele (T) was overrepresented in the LBM group.

3.2. Association of variants with BMD in the complete cohort

We selected 8 interesting variants, taking into consideration both frequency and functionality (Fig. 1), including 2 located in a putative regulatory region 4 kb upstream of the *C7ORF76* transcription start site (TSS), which we named UPE, that was not included in the re-sequencing. These 8 variants were genotyped in the complete BARCOS cohort (n = 1490) to test their association with BMD and osteoporotic fracture. For all of them, we obtained MAF values similar to those found in the 1000 Genomes database for the European or Iberian populations (Supplementary Table 4). Significant differences were obtained with two of the 5′ upstream SNPs, rs4342521 and rs10085588 (Fig. 1), for LS-BMD under additive and recessive models and nominal differences were obtained with the same SNPs for osteoporotic fracture, under additive and dominant models (Table 2). In all cases, the minor allele (T and A, respectively) had a BMD-lowering effect on LS and had a damaging effect on osteoporotic fracture. The two associated SNPs were found to be in linkage disequilibrium with the GWAS hit (rs4727338)

Table 1
Number and location of single nucleotide variants found in this study.

	Raw	Selected by filtering and validating	Coding	Regulatory regions[a]	Putative osteoblast regulatory elements	Active enhancer
Common variants	96	51	0	12	3	1
LFV	24,243	59	1	16	1	0
Total	24,339	110	1	28	4	1

[a] Either present in flanking regions, 5′UTR, 3′UTR or introns; LFV: lower frequency variants (MAF < 0.05).

Table 2
Association between common and LFV variants of *C7ORF76* and LS-BMD, FN-BMD or osteoporotic fracture.

SNP	Genomic position (GRCh37)	Type of variant	p-value LS-BMD			Effect size (β coeff; 95% CI)	p-value FN-BMD			Effect size (β coeff; 95% CI)	p-value osteoporotic fracture			Effect size (OR; 95% CI)
			A	R	D		A	R	D		A	R	D	
rs115076023[a]	7:96137731 G > A	5'UP	–	–	–		–	–	–		–	–	–	
rs10085588	7:96137674 G > A	5'UP	**0.00361**	**0.00343**	0.05032	−0.0324 (−0.0541, −0.0108)	0.12419	0.11909	0.30848		*0.03554*	0.32373	*0.02242*	1.51 (1.05, 2.17)
rs4342521	7:96136005 G > T	5'UP	**0.00314**	**0.00251**	0.05233	−0.0332 (−0.0547, −0.0117)	0.12453	0.09115	0.36111		*0.03911*	0.39100	*0.01974*	1.53 (1.06, 2.19)
rs7794042	7:96132999 C > A	5'UP	0.68070	0.65857	0.63405		0.85483	0.66330	0.89686		0.55917	1.00000	0.30987	
rs4613908	7:96125315 G > A	I	0.06164	0.08889	0.15109		0.55060	0.30356	0.86063		0.58472	0.27134	0.18380	
rs10238953	7:96124975 T > C	I	0.57370	0.49214	0.69433		0.71522	0.57609	0.83130		0.14074	0.48140	0.15002	
rs190892252	7:96121343 A > G	I	0.21722	–	–		0.24861	–	–		0.63625	–	–	
rs4727338[b]	7:96120675 G > C	I	**0.00141**	**0.00207**	*0.02430*	−0.0362 (−0.0592, −0.0132)	0.06943	*0.04540*	0.27708	−0.0175 (−0.0345, −0.0004)	0.07618	0.56459	*0.03270*	1.50 (1.03, 2.18)
rs117923361	7:96111486 C > A	3'UTR	0.82861	0.68376	0.87501		0.08217	0.62988	0.08427		0.30507	1.00000	0.22534	

Values in bold indicate statistical significance; values in italics indicate nominal significance.
A: Additive model; **R**: Recessive model; **D**: Dominant model.
[a] SNP found monomorphic in the BARCOS cohort.
[b] SNP previously genotyped in the BARCOS cohort (Estrada et al. [7]).

Fig. 2. Linkage disequilibrium plot of *C7ORF76* genotyped variants. *GWAS hit from Estrada et al. [7]. The numbers within the squares and the colour scale both refer to D'/LOD values (with bright red: D' = 1 and LOD ≥ 2; white: D' < 1 and LOD < 2; blue: D' = 1 and LOD < 2 and shades of pink/red: D' < 1 and LOD ≥ 2).

from Estrada et al. [7] (Fig. 2), which was also significantly associated with LS-BMD in the BARCOS cohort (Table 2).

3.3. cis-eQTL analyses

To evaluate the functionality of the associated variants (rs4727338 –GWAS hit-, rs4342521 and rs10085588) we first tested them as *cis*-eQTLs in human primary osteoblasts, a cell type unavailable in GTEx. Of note, we failed to detect *C7ORF76* mRNA expression in these cells. None of the SNPs were found to be eQTLs for *DLX5, DLX6*, *DLX6-AS1*, or *SHFM1* (*SEM1* transcript variant 5 in GRCh38), although a tendency was observed for the latter (Table 3). On the contrary, the 3 SNPs were found nominally associated with *SLC25A13* gene expression, where the minor alleles were associated with decreased gene expression. We also tested another SNP of the region (rs10429035) found associated with BMD in a previous GWAS meta-analysis [8] and described in GTEx to be eQTL for *SHFM1* in tibial artery. However, we failed to detect an association of this SNP with gene expression of any of the nearby genes (data not shown).

3.4. Evaluation of the regulatory capability of the UPE

Next, as the associated SNP rs10085588 is located in a putative regulatory region 4 kb upstream of the *C7ORF76* gene (UPE; Fig. 1), we assessed the functionality of the UPE. We performed luciferase reporter

Table 3
cis-eQTL analysis of 3 *C7ORF76* variants associated with BMD.

SNP	p-values				
	DLX5	*DLX6*	*DLX6-AS1*	*SHFM1*	*SLC25A13*
rs10085588	0.14207	0.53849	0.59125	0.05270	*0.03350*
rs4342521[a]	0.09227	0.34167	0.55062	0.05725	*0.01442*
rs4727338[a]	0.09227	0.34167	0.55062	0.05725	*0.01442*

Values in italics indicate nominal significance.
[a] p-values are identical as a reflection of the high LD between the 2 SNPs.

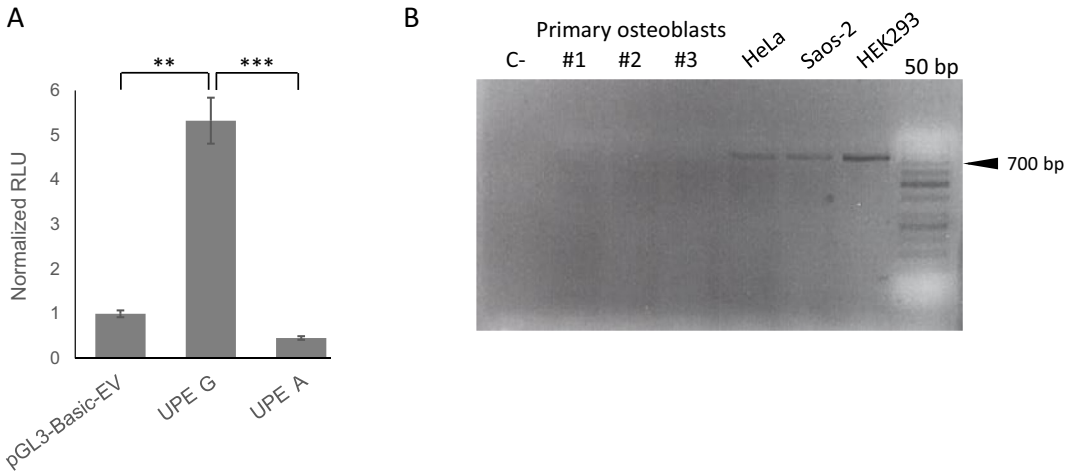

Fig. 3. (A) Luciferase activity of different versions of UPE, containing the G or A allele for the SNP rs10085588, in a forward or inversed orientation, in Saos-2 cells. Results are expressed as mean ± SD. **p-value < 0.01; ***p-value < 0.001 (B) PCR amplification of UPE from cDNA of different cell types. Expected size: 750 bp.

assays in Saos-2 cells to test UPE activity in both orientations and the effects of the two alleles of rs10085588. As shown in Fig. 3A, the forward UPE construct bearing the major allele (G) showed significantly increased luciferase expression compared to the empty vector (FC: 5.3, $p = 0.0097$), and compared to the construct bearing the minor allele (A) (FC: 11.6, $p < 0.001$). No activity was detected when the UPE was tested in the inverse orientation (data not shown). Results of luciferase assays were consistent with eQTL analyses, in the sense that the minor allele (A) fails to activate transcription and is associated with lower expression of *SLC25A13* in osteoblasts (see above). We also evaluated whether the UPE was transcribed (as occurs in many regulatory regions) by performing RT-PCR in cDNA of human primary osteoblasts, Saos-2, HeLa and HEK293 cells. We were able to amplify the UPE sequence from cDNA of HeLa, HEK293 and Saos-2 cells but we failed to amplify it from different cDNA samples of human primary osteoblasts (Fig. 3B). We also interrogated the FANTOM5 Cap Analysis of Gene Expression (CAGE) dataset [26] for evidence of UPE transcription and we could observe a signal for TSS expression in smooth muscle cells.

3.5. Chromatin interactions from the UPE

Finally, we investigated the possible genomic targets of the UPE by examining the 3D chromatin interactions by 4C-seq in different cell types (MSCs, hFOB 1.19 and Saos-2). We detected interaction between UPE and the genomic region spanning approximately 750 kb on each side of it (Fig. 4), and no other interactions were detected elsewhere in the genome. Notably, we found higher interaction levels with most of the tissue-specific enhancers described in the VISTA browser [42,43], especially in the hFOB 1.19 cell line. We could also detect higher interaction with the long non-coding RNA gene *LOC100506136*, upstream of *C7ORF76*. In addition, we analysed the topologically associated domains (TADs) of the region on different cell types using available Hi-C data [40] and the 3D Genome Browser [41] and we observed that in many cell types, the TAD containing the gene *C7ORF76* spanned from approximately 50 kb upstream of *DYNC1I1* TSS to approximately 25 kb upstream of *ACN9* TSS (Fig. 4), consistent with the 4C-seq results.

4. Discussion

Different non-coding variants in the *C7ORF76* genomic region have been previously associated with BMD and osteoporotic fracture in different GWAS [4,6–8,14–16] and at this point, deciphering the functionality of this region would be the next logical step for understanding these associations. In this line, we have analysed the *C7ORF76* region in depth, including re-sequencing, testing variants for association in the BARCOS cohort, and analysing the functionality of the variants and regulatory regions both *in silico* and experimentally. Two upstream variants (rs4342521 and rs10085588) were found significantly associated with LS-BMD and nominally associated with osteoporotic fracture. In addition, both SNPs have been identified as eQTLs of *SLC25A13* in human primary osteoblasts. Moreover, the SNP rs10085588 falls in a regulatory region (UPE) able to stimulate transcription in an allele-dependent manner. This UPE was found to interact with different tissue-specific enhancers and a lncRNA present in the nearby region.

The *C7ORF76* gene is an uncharacterised gene of unknown function without expression data in GTEx. In the current human genome assembly (GRCh38) it is annotated as alternative transcript 6 of *SEM1* (encoding a 26S proteasome complex subunit, whose alternative historical name is *SHFM1* for Split Hand and Foot Malformation 1). However, although RefSeq currently labels it as curated gene, surprisingly few data have been gathered and the real function of *C7ORF76* remains elusive. This is in contrast with the consistent finding of potent GWAS signals for osteoporosis within this gene. Of note, this gene seems to be not expressed in primary osteoblasts, Saos-2, HeLa and SH-SY5Y (data not shown) which suggests that either it might, indeed, not be an osteoblast gene, yet regulate BMD from a different cell type or organ, or, alternatively, other genes in the region could be causal for the association.

Most GWAS variants for complex diseases are located in non-coding regulatory regions (reviewed in Zhang et al. [44]) and several studies have pinpointed the importance of regulatory elements for the susceptibility to osteoporosis [45,46]. Moreover, for some common traits, it has been described that several causal variants exist in a single LD block, located in multiple enhancers that cooperatively influence gene expression (the so called super-enhancers) [34,47]. These enhancer clusters are highly cell type specific. In this respect, although publicly available datasets [34,35] did not consider the *C7ORF76* genomic region to be a super-enhancer, the functional annotation of this region revealed the presence of several putative regulatory elements containing SNPs prone to confer susceptibility to osteoporosis.

To elucidate the molecular mechanisms underlying the strong association, further exploration with high-coverage sequencing to prioritise potentially causal variants was necessary [48]. We performed an extreme-truncate selection of the BARCOS cohort as a discovery phase, prior to genotyping the selected variants in the complete cohort. The efficiency of extreme-truncate selection approach for quantitative trait association studies (*e.g.* BMD) has been widely proven, as well as its utility for detecting rare variants [6,49,50]. The 2 SNPs found associated in the BARCOS cohort in this study (rs4342521 and rs10085588) were previously found associated with LS-BMD, FN-BMD and

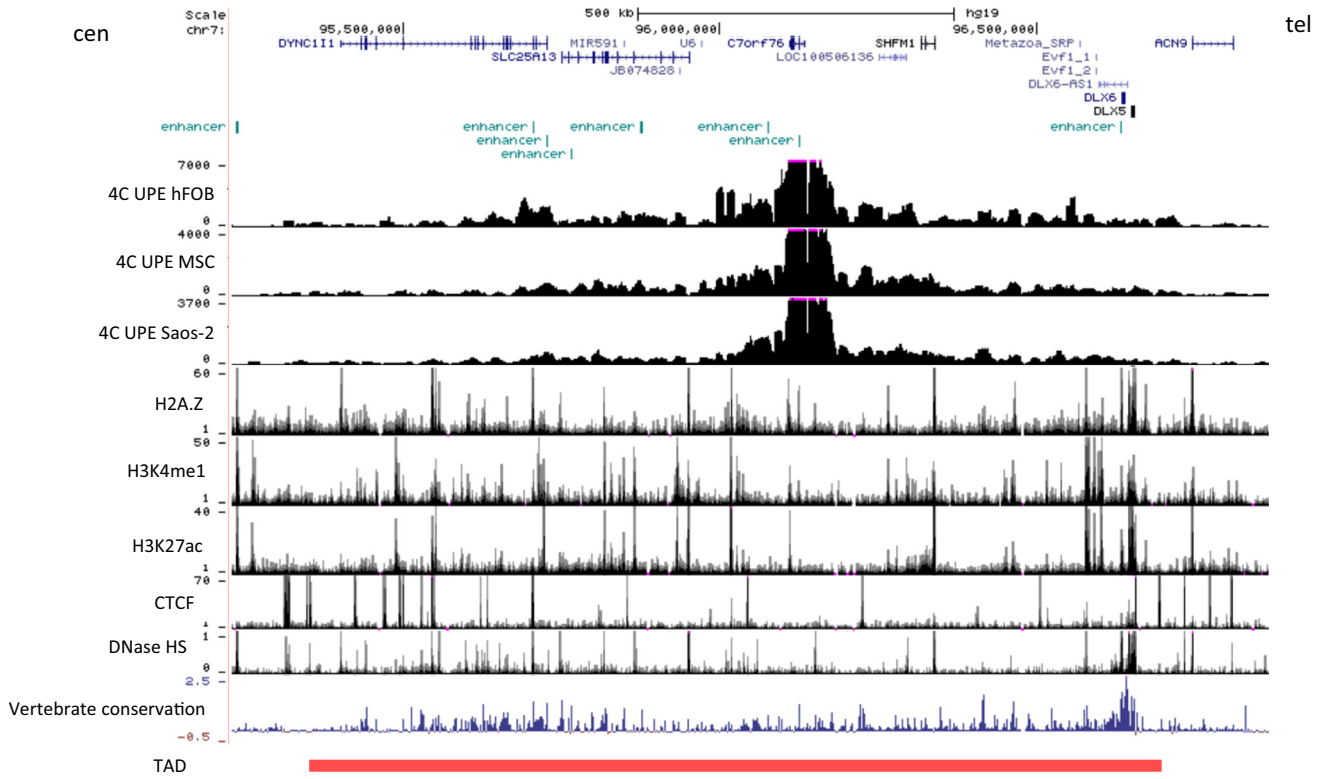

Fig. 4. 4C-seq using UPE as viewpoint in human fetal osteoblast (hFOB) 1.19, mesenchymal stem cells (MSC) and Saos-2 cell line. H2A.Z, H3K4me1, H3K27ac and CTCF ChIP-seq and DNase HS from osteoblasts (ENCODE data), and vertebrate conservation are shown. Experimentally validated active enhancers from VISTA Enhancer Browser [42] are shown in green. In red, topologically associated domain (TAD).

osteoporotic fracture in other GWA studies [4,9]. In contrast, we failed to find association in the rest of the variants interrogated, although other SNPs in this region have been found associated with LS-BMD, FN-BMD, and osteoporotic fracture [6,8,16], as well as with heel BMD [10,12] and total body BMD [11]. Taken together, these association data are consistent with the existence of a large LD block encompassing all the associated variants and the comparatively small sample size of the BARCOS cohort may preclude the detection of some of them.

To further delineate the role of the three associated variants (the two mentioned above and the GWAS SNP in Estrada et al. [7]), we performed *cis*-eQTL analyses in human primary osteoblasts and we detected a nominal association between the minor alleles of the three SNPs and decreased *SLC25A13* gene expression. We also detected a trend for association with decreased expression of *SHFM1*. These results reflect the LD among the SNPs. We failed to find an association between SNP rs10429035 and transcription levels of *SHFM1*, while in GTEx this SNP is described as eQTL for this gene in tibial artery. It could be that it is not eQTL in primary osteoblasts. Alternatively, our limited power (n = 45) and the fact that primary cells are not as homogeneous as cell lines may have prevented us to detect it in primary osteoblasts [51].

Out of the 3 associated SNPs, only one lies within a regulatory region (UPE, see Fig. 1), while the others map to sequences lacking regulatory marks. We set out to experimentally study the UPE, which is a conserved region located approximately 4 kb upstream of *C7ORF76* TSS, with enhancer marks such as H3K27ac and H3K4me1, as well as a DNase hypersensitivity signal. According to ENCODE ChIP-seq data, many transcription factors bind to UPE, among which RAD21, MYC, POLR2A, and the P300 histone acetyltransferase, known to be involved in transcription regulation, initiation and elongation and in enhancer activity [52,53]. Our results, including luciferase assays, RT-PCR and 4C analyses, indicate that, indeed, the UPE acts as a regulatory element, able to activate transcription of a reporter gene. It is well known that many such elements can be transcribed [54–56], producing non-coding RNAs (or eRNAs for enhancer-RNAs) and, in this sense, we found UPE to be transcribed in Saos-2 cells.

Our results also provide clear evidence that SNP rs10085588, within the UPE, is itself functional, since we showed that the minor allele (A) abolished luciferase activation, which was otherwise stimulated by the major allele G. We have performed an *in silico* analysis and found predictions that the histone deacetylase HDAC2 may bind the A allele more probably than the G (p-value: 5.30016e-06), a possible explanation for the reduced expression observed.

The 4C-seq experiments showed that, in the cells included in this study, the UPE interacted with several sites within the TAD where it belongs (see Fig. 4), and nowhere else in the genome. This suggests that the biological function of the GWAS signal should be limited to genes within this TAD, further supported by strong CTCF signals limiting the region. However, interactions with other genomic regions may exist in different tissues or differentiation stages, not tested by us. In this sense and interestingly, a computationally-based characterization of osteoporosis associated SNPs identified an interaction between the GWAS SNP rs4729260 in the *C7ORF76* region and the Xq12 genomic region [57], which contains the androgen receptor gene, known to be involved in bone metabolism [58].

Within the TAD, the UPE enhancer interacted with many other tissue-specific enhancers previously identified [42,43], suggesting that they may act cooperatively or redundantly in regulating gene expression. Several of these enhancers have been shown to affect *DLX5/6* gene expression [43]. However, we have not detected strong interaction signals with the *DLX5/6* region and, in our eQTL study in primary osteoblasts, we did not observe any effect of the associated SNPs on *DLX5/6* gene expression. Likewise, we detected modest interactions between the UPE and the *SHFM1* and the *SLC25A13* coding regions, and negligible signals with that of *DYNC1I1*. In contrast, we did detect a marked interaction with a lncRNA in the close vicinity of *C7ORF76*, namely *LOC100506136*. Interestingly, a recent study found a SNP

within *LOC100506136* to be one of the 2 out of approximately 55,000 genome-wide SNPs in lncRNAs associated with total hip BMD [59]. Another recent publication found a LS-BMD associated signal within *LOC100506136* in a Mexican-Mestizo cohort [60]. These studies suggest that this lncRNA could be involved in osteoporosis pathogenesis.

This work has some limitations including the sample size of the BARCOS cohort, and of the primary osteoblasts used for the eQTL analyses, both of which preclude the detection of variants with smaller effects. In addition, *C7ORF76* was not tested in other bone cells such as osteoclasts. However, these results provide interesting data to understand the functionality of this unexplored region, one of the most frequently found associated with osteoporosis in GWA studies. They also highlight the importance of regulatory variants in bone phenotypes and the usefulness of integrative approaches to uncover their functionality.

5. Conclusions

In summary, this study is the first one to analyse in depth the functionality of the *C7ORF76* genomic region, associated with BMD and osteoporotic fracture in many GWAS. We provide functional regulatory evidence for rs10085588, which may be a causal SNP within this 7q21.3 GWAS signal for osteoporosis.

CRediT authorship contribution statement

Neus Roca-Ayats: Investigation, Formal analysis, Visualization, Writing - original draft, Writing - review & editing. **Núria Martínez-Gil:** Investigation, Formal analysis, Writing - review & editing. **Mónica Cozar:** Investigation, Writing - review & editing. **Marina Gerousi:** Investigation, Writing - review & editing. **Natàlia Garcia-Giralt:** Conceptualization, Writing - review & editing. **Diana Ovejero:** Conceptualization, Writing - review & editing. **Leonardo Mellibovsky:** Conceptualization, Writing - review & editing. **Xavier Nogués:** Conceptualization, Writing - review & editing. **Adolfo Díez-Pérez:** Conceptualization, Writing - review & editing. **Daniel Grinberg:** Conceptualization, Funding acquisition, Formal analysis, Supervision, Writing - original draft, Writing - review & editing. **Susanna Balcells:** Conceptualization, Funding acquisition, Formal analysis, Supervision, Writing - original draft, Writing - review & editing.

Acknowledgements

We thank E. Czwan for relevant technical assistance. This work was supported by the following grants: SAF2014-56562-R, SAF2016-75948-R (Spanish MINECO), 2014SGR932 (Generalitat de Catalunya), CIBERER (U720), and FEIOMM Investigación 2014. NRA is a recipient of an FPU predoctoral fellowship from the Spanish Ministerio de Educación, Cultura y Deporte; NMG is a recipient of a FI predoctoral fellowship from Generalitat de Catalunya.

Declaration of interests

The authors declare they have no conflict of interests.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.bone.2019.03.014.

References

[1] M.T. Maurano, R. Humbert, E. Rynes, R.E. Thurman, E. Haugen, H. Wang, et al., Systematic localization of common disease-associated variation in regulatory DNA, Science 337 (2012) 1190–1195, https://doi.org/10.1126/science.1222794.
[2] Y.G. Tak, P.J. Farnham, Making sense of GWAS: using epigenomics and genome engineering to understand the functional relevance of SNPs in non-coding regions of the human genome, Epigenetics Chromatin 8 (2015) 57, https://doi.org/10.1186/s13072-015-0050-4.
[3] J.B. Richards, F. Rivadeneira, M. Inouye, T.M. Pastinen, N. Soranzo, S.G. Wilson, et al., Bone mineral density, osteoporosis, and osteoporotic fractures: a genome-wide association study, Lancet 371 (2008) 1505–1512, https://doi.org/10.1016/S0140-6736(08)60599-1.
[4] F. Rivadeneira, U. Styrkársdottir, K. Estrada, B.V. Halldórsson, Y.H. Hsu, J.B. Richards, et al., Twenty bone-mineral-density loci identified by large-scale meta-analysis of genome-wide association studies, Nat. Genet. 41 (2009) 1199–1206, https://doi.org/10.1038/ng.446.
[5] U. Styrkarsdottir, B.V. Halldorsson, S. Gretarsdottir, D.F. Gudbjartsson, G.B. Walters, T. Ingvarsson, et al., New sequence variants associated with bone mineral density, Nat. Genet. 41 (2009) 15–17, https://doi.org/10.1038/ng.284.
[6] E.L. Duncan, P. Danoy, J.P. Kemp, P.J. Leo, E. McCloskey, G.C. Nicholson, et al., Genome-wide association study using extreme truncate selection identifies novel genes affecting bone mineral density and fracture risk, PLoS Genet. 7 (2011) e1001372, , https://doi.org/10.1371/journal.pgen.1001372.
[7] K. Estrada, U. Styrkarsdottir, E. Evangelou, Y.H. Hsu, E.L. Duncan, E.E. Ntzani, et al., Genome-wide meta-analysis identifies 56 bone mineral density loci and reveals 14 loci associated with risk of fracture, Nat. Genet. 44 (2012) 491–501, https://doi.org/10.1038/ng.2249.
[8] L. Zhang, H.J. Choi, K. Estrada, P.J. Leo, J. Li, Y.-F. Pei, et al., Multistage genome-wide association meta-analyses identified two new loci for bone mineral density, Hum. Mol. Genet. 23 (2014) 1923–1933, https://doi.org/10.1093/hmg/ddt575.
[9] H.F. Zheng, V. Forgetta, Y.H. Hsu, K. Estrada, A. Rosello-Diez, P.J. Leo, et al., Whole-genome sequencing identifies EN1 as a determinant of bone density and fracture, Nature 526 (2015) 112–117, https://doi.org/10.1038/nature14878.
[10] J.P. Kemp, J.A. Morris, C. Medina-Gomez, V. Forgetta, N.M. Warrington, S.E. Youlten, et al., Identification of 153 new loci associated with heel bone mineral density and functional involvement of GPC6 in osteoporosis, Nat. Genet. 49 (2017) 1468–1475, https://doi.org/10.1038/ng.3949.
[11] C. Medina-Gomez, J.P. Kemp, K. Trajanoska, J. Luan, A. Chesi, T.S. Ahluwalia, et al., Life-course genome-wide association study meta-analysis of total body BMD and assessment of age-specific effects, Am. J. Hum. Genet. 102 (2018) 88–102, https://doi.org/10.1016/j.ajhg.2017.12.005.
[12] J.A. Morris, J.P. Kemp, S.E. Youlten, L. Laurent, J.G. Logan, R. Chai, et al., An atlas of genetic influences on osteoporosis in humans and mice, Nat. Genet. (2018), https://doi.org/10.1038/s41588-018-0302-x.
[13] J.B. Richards, H.F. Zheng, T.D. Spector, Genetics of osteoporosis from genome-wide association studies: advances and challenges, Nat. Rev. Genet. 13 (2012) 576–588, https://doi.org/10.1038/nrg3228.
[14] Y. Guo, J.-T. Wang, H. Liu, M. Li, T.-L. Yang, X.-W. Zhang, et al., Are bone mineral density loci associated with hip osteoporotic fractures? A validation study on previously reported genome- wide association loci in a Chinese population, Genet. Mol. Res. 11 (2012) 202–210, https://doi.org/10.4238/2012.January.31.1.
[15] U. Styrkarsdottir, B.V. Halldorsson, D.F. Gudbjartsson, N.L.S. Tang, J.M. Koh, S.M. Xiao, et al., European bone mineral density loci are also associated with BMD in East-Asian populations, PLoS One 5 (2010) e13217, , https://doi.org/10.1371/journal.pone.0013217.
[16] K. Trajanoska, J.A. Morris, L. Oei, H.-F. Zheng, D.M. Evans, D.P. Kiel, et al., Assessment of the genetic and clinical determinants of fracture risk: genome wide association and mendelian randomisation study, Br. J. Med. 362 (2018) k3225, , https://doi.org/10.1136/bmj.k3225.
[17] M. Bustamante, X. Nogués, L. Mellibovsky, L. Agueda, S. Jurado, E. Cáceres, et al., Polymorphisms in the interleukin-6 receptor gene are associated with bone mineral density and body mass index in Spanish postmenopausal women, Eur. J. Endocrinol. 157 (2007) 677–684, https://doi.org/10.1530/EJE-07-0389.
[18] M. Bustamante, X. Nogués, L. Águeda, S. Jurado, A. Wesselius, E. Cáceres, et al., Promoter 2 -1025 T/C polymorphism in the RUNX2 gene is associated with femoral neck BMD in Spanish postmenopausal women, Calcif. Tissue Int. 81 (2007) 327–332, https://doi.org/10.1007/s00223-007-9069-2.
[19] H. Li, B. Handsaker, A. Wysoker, T. Fennell, J. Ruan, N. Homer, et al., The Sequence Alignment/Map format and SAMtools, Bioinformatics 25 (2009) 2078–2079, https://doi.org/10.1093/bioinformatics/btp352.
[20] M.A. DePristo, E. Banks, R. Poplin, K.V. Garimella, J.R. Maguire, C. Hartl, et al., A framework for variation discovery and genotyping using next-generation DNA sequencing data, Nat. Genet. 43 (2011) 491–498, https://doi.org/10.1038/ng.806.
[21] P. Kumar, S. Henikoff, P.C. Ng, Predicting the effects of coding non-synonymous variants on protein function using the SIFT algorithm, Nat. Protoc. 4 (2009) 1073–1081, https://doi.org/10.1038/nprot.2009.86.
[22] I.A. Adzhubei, S. Schmidt, L. Peshkin, V.E. Ramensky, A. Gerasimova, P. Bork, et al., A method and server for predicting damaging missense mutations, Nat. Methods 7 (2010) 248–249, https://doi.org/10.1038/nmeth0410-248.
[23] J.M. Schwarz, D.N. Cooper, M. Schuelke, D. Seelow, MutationTaster2: mutation prediction for the deep-sequencing age, Nat. Methods 11 (2014) 361–362, https://doi.org/10.1038/nmeth.2890.
[24] ENCODE Project Consortium, The ENCODE (ENCyclopedia Of DNA Elements) Project, Science. 306 (2004) 636–640, https://doi.org/10.1126/science.1105136.
[25] D. Bujold, D.A.de.L. Morais, C. Gauthier, C. Côté, M. Caron, T. Kwan, et al., The International Human Epigenome Consortium Data Portal, Cell Syst. 3 (2016), https://doi.org/10.1016/j.cels.2016.10.019 496–499.e2.
[26] A.R.R. Forrest, H. Kawaji, M. Rehli, J.K. Baillie, M.J.L. De Hoon, V. Haberle, et al., A promoter-level mammalian expression atlas, Nature 507 (2014) 462–470, https://doi.org/10.1038/nature13182.
[27] L.D. Ward, M. Kellis, HaploReg: a resource for exploring chromatin states,

conservation, and regulatory motif alterations within sets of genetically linked variants, Nucleic Acids Res. 40 (2012) D930–D934, https://doi.org/10.1093/nar/gkr917.
[28] A.P. Boyle, E.L. Hong, M. Hariharan, Y. Cheng, M.A. Schaub, M. Kasowski, et al., Annotation of functional variation in personal genomes using RegulomeDB, Genome Res. 22 (2012) 1790–1797, https://doi.org/10.1101/gr.137323.112.
[29] C.H. Chou, S. Shrestha, C.D. Yang, N.W. Chang, Y.L. Lin, K.W. Liao, et al., MiRTarBase update 2018: A resource for experimentally validated microRNA-target interactions, Nucleic Acids Res. 46 (2018) D296–D302, https://doi.org/10.1093/nar/gkx1067.
[30] A.E. Bruno, L. Li, J.L. Kalabus, Y. Pan, A. Yu, Z. Hu, miRdSNP: a database of disease-associated SNPs and microRNA target sites on 3'UTRs of human genes, BMC Genomics 13 (2012) 44, https://doi.org/10.1186/1471-2164-13-44.
[31] C. Liu, F. Zhang, T. Li, M. Lu, L. Wang, W. Yue, et al., MirSNP, a database of polymorphisms altering miRNA target sites, identifies miRNA-related SNPs in GWAS SNPs and eQTLs, BMC Genomics 13 (2012) 661, https://doi.org/10.1186/1471-2164-13-661.
[32] K. Cartharius, K. Frech, K. Grote, B. Klocke, M. Haltmeier, A. Klingenhoff, et al., MatInspector and beyond: promoter analysis based on transcription factor binding sites, Bioinformatics 21 (2005) 2933–2942, https://doi.org/10.1093/bioinformatics/bti473.
[33] S.G. Coetzee, G.A. Coetzee, D.J. Hazelett, motifbreakR: an R/Bioconductor package for predicting variant effects at transcription factor binding sites, Bioinformatics 31 (2015) 3847–3849, https://doi.org/10.1093/bioinformatics/btv470.
[34] D. Hnisz, B.J. Abraham, T.I. Lee, A. Lau, V. Saint-André, A.A. Sigova, et al., Super-enhancers in the control of cell identity and disease, Cell. 155 (2013) 934–947, https://doi.org/10.1016/j.cell.2013.09.053.
[35] Y. Jiang, F. Qian, X. Bai, Y. Liu, Q. Wang, B. Ai, et al., SEdb: a comprehensive human super-enhancer database, Nucleic Acids Res. 47 (2019) D235–D243, https://doi.org/10.1093/nar/gky1025.
[36] J.C. Barrett, B. Fry, J. Maller, M.J. Daly, Haploview: analysis and visualization of LD and haplotype maps, Bioinformatics 21 (2005) 263–265, https://doi.org/10.1093/bioinformatics/bth457.
[37] A. Fernández-Miñán, J. Bessa, J.J. Tena, J.L. Gómez-Skarmeta, Assay for transposase-accessible chromatin and circularized chromosome conformation capture, two methods to explore the regulatory landscapes of genes in zebrafish, in: H.W. Detrich, IIIM. Westerfield, L.I. Zon (Eds.), Methods Cell Biol., Elsevier Inc, 2016, pp. 413–430, , https://doi.org/10.1016/bs.mcb.2016.02.008.
[38] H.J.G. van de Werken, P.J.P. de Vree, E. Splinter, S.J.B. Holwerda, P. Klous, E. de Wit, et al., 4C technology: Protocols and data analysis, in: C. Wu, C.D. Allis (Eds.), Methods Enzymol., Elsevier Inc, 2012, pp. 89–112, , https://doi.org/10.1016/B978-0-12-391938-0.00004-5.
[39] D. Noordermeer, M. Leleu, E. Splinter, J. Rougemont, W. De Laat, D. Duboule, The dynamic architecture of HOX gene clusters, Science 334 (2011) 222–225, https://doi.org/10.1126/science.1207194.
[40] J.R. Dixon, S. Selvaraj, F. Yue, A. Kim, Y. Li, Y. Shen, et al., Topological domains in mammalian genomes identified by analysis of chromatin interactions, Nature 485 (2012) 376–380, https://doi.org/10.1038/nature11082.
[41] Y. Wang, F. Song, B. Zhang, L. Zhang, J. Xu, D. Kuang, et al., The 3D Genome Browser: a web-based browser for visualizing 3D genome organization and long-range chromatin interactions, Genome Biol. 19 (2018) 151, https://doi.org/10.1186/s13059-018-1519-9.
[42] A. Visel, S. Minovitsky, I. Dubchak, L.A. Pennacchio, VISTA Enhancer Browser - a database of tissue-specific human enhancers, Nucleic Acids Res. 35 (2007) D88–D92, https://doi.org/10.1093/nar/gkl822.
[43] R.Y. Birnbaum, D.B. Everman, K.K. Murphy, F. Gurrieri, C.E. Schwartz, N. Ahituv, Functional characterization of tissue-specific enhancers in the DLX5/6 locus, Hum. Mol. Genet. 21 (2012) 4930–4938, https://doi.org/10.1093/hmg/dds336.
[44] F. Zhang, J.R. Lupski, Non-coding genetic variants in human disease, Hum. Mol. Genet. 24 (2015) R102–R110, https://doi.org/10.1093/hmg/ddv259.
[45] S. Yao, Y. Guo, S.S. Dong, R.H. Hao, X.F. Chen, Y.X. Chen, et al., Regulatory element-based prediction identifies new susceptibility regulatory variants for osteoporosis, Hum. Genet. 136 (2017) 963–974, https://doi.org/10.1007/s00439-017-1825-4.
[46] Y. Guo, S.S. Dong, X.F. Chen, Y.A. Jing, M. Yang, H. Yan, et al., Integrating epigenomic elements and GWASs identifies BDNF gene affecting bone mineral density and osteoporotic fracture risk, Sci. Rep. 6 (2016) 30558, https://doi.org/10.1038/srep30558.
[47] O. Corradin, A. Saiakhova, B. Akhtar-Zaidi, L. Myeroff, J. Willis, R. Cowper-Sal lari, et al., Combinatorial effects of multiple enhancer variants in linkage disequilibrium dictate levels of gene expression to confer susceptibility to common traits, Genome Res. 24 (2014) 1–13, https://doi.org/10.1101/gr.164079.113.
[48] Y.H. Hsu, G. Li, C.T. Liu, J.A. Brody, D. Karasik, W.C. Chou, et al., Targeted sequencing of genome wide significant loci associated with bone mineral density (BMD) reveals significant novel and rare variants: the Cohorts for Heart and Aging Research in Genomic Epidemiology (CHARGE) targeted sequencing study, Hum. Mol. Genet. 25 (2016) 5234–5243, https://doi.org/10.1093/hmg/ddw289.
[49] G. Kang, D. Lin, H. Hakonarson, J. Chen, Two-stage extreme phenotype sequencing design for discovering and testing common and rare genetic variants: efficiency and power, Hum. Hered. 73 (2012) 139–147, https://doi.org/10.1159/000337300.
[50] I.J. Barnett, S. Lee, X. Lin, Detecting rare variant effects using extreme phenotype sampling in sequencing association studies, Genet. Epidemiol. 37 (2013) 142–151, https://doi.org/10.1002/gepi.21699.
[51] D.S. Paul, N. Soranzo, S. Beck, Functional interpretation of non-coding sequence variation: concepts and challenges, BioEssays 36 (2014) 191–199, https://doi.org/10.1002/bies.201300126.
[52] P.B. Rahl, R.A. Young, MYC and transcription elongation, Cold Spring Harb. Perspect. Med. 4 (2014) a020990, https://doi.org/10.1101/cshperspect.a020990.
[53] R. Raisner, S. Kharbanda, L. Jin, E. Jeng, E. Chan, M. Merchant, et al., Enhancer activity requires CBP/P300 bromodomain-dependent histone H3K27 acetylation, Cell Rep. 24 (2018) 1722–1729, https://doi.org/10.1016/j.celrep.2018.07.041.
[54] R. Andersson, C. Gebhard, I. Miguel-Escalada, I. Hoof, J. Bornholdt, M. Boyd, et al., An atlas of active enhancers across human cell types and tissues, Nature 507 (2014) 455–461, https://doi.org/10.1038/nature12787.
[55] E. Arner, C.O. Daub, K. Vitting-Seerup, R. Andersson, B. Lilje, F. Drabløs, et al., Transcribed enhancers lead waves of coordinated transcription in transitioning mammalian cells, Science 347 (2015) 1010–1014, https://doi.org/10.1126/science.1259418.
[56] C.C. Hon, J.A. Ramilowski, J. Harshbarger, N. Bertin, O.J.L. Rackham, J. Gough, et al., An atlas of human long non-coding RNAs with accurate 5′ ends, Nature 543 (2017) 199–204, https://doi.org/10.1038/nature21374.
[57] L. Qin, Y. Liu, Y. Wang, G. Wu, J. Chen, W. Ye, et al., Computational characterization of osteoporosis associated SNPs and genes identified by genome-wide association studies, PLoS One 11 (2016) e0150070, , https://doi.org/10.1371/journal.pone.0150070.
[58] S.C. Manolagas, C.A. O'Brien, M. Almeida, The role of estrogen and androgen receptors in bone health and disease, Nat. Rev. Endocrinol. 9 (2013) 699–712, https://doi.org/10.1038/nrendo.2013.179.
[59] Q. Zeng, K.H. Wu, K. Liu, Y. Hu, X.D. Chen, L. Zhang, et al., Genome-wide association study of lncRNA polymorphisms with bone mineral density, Ann. Hum. Genet. 82 (2018) 244–253, https://doi.org/10.1111/ahg.12247.
[60] M. Villalobos-Comparán, R.F. Jiménez-Ortega, K. Estrada, A.Y. Parra-Torres, A. González-Mercado, N. Patiño, et al., A pilot genome-wide association study in postmenopausal Mexican-Mestizo women implicates the RMND1/CCDC170 locus is associated with bone mineral density, Int. J. Genomics 2017 (2017) 5831020, https://doi.org/10.1155/2017/5831020.