Table of Contents

Supplemental methods ........................................................................................................ 2
Whole exome data analysis ............................................................................................... 2
Expression of the ZNF528 in vitro .................................................................................. 2
Protein expression and subcellular localization of ZNF528 ............................................. 5
Cell viability and cytotoxicity assay ................................................................................ 6
Chromatin immunoprecipitation (ChIP) sequencing ......................................................... 7
ChIP sequencing data analysis ......................................................................................... 7
Differential gene expression analyses ............................................................................. 9

Supplemental Figures ......................................................................................................... 11
Supplemental Figure S1 .................................................................................................... 11
Supplemental Figure S2 .................................................................................................... 12
Supplemental Figure S3 .................................................................................................... 13
Supplemental Figure S4 .................................................................................................... 14
Supplemental Figure S5 .................................................................................................... 15
Supplemental Figure S6 .................................................................................................... 16
Supplemental Figure S7 .................................................................................................... 17
Supplemental Figure S8 .................................................................................................... 18
Supplemental Figure S9 .................................................................................................... 19
Supplemental Figure S10 ................................................................................................. 20
Supplemental Figure S11 ................................................................................................. 21
Supplemental Figure S12 ................................................................................................. 22

Supplemental Tables .......................................................................................................... 23
Supplemental Table S1 ..................................................................................................... 23
Supplemental Table S2 ..................................................................................................... 24
Supplemental Table S3 ..................................................................................................... 25
Supplemental Table S4 ..................................................................................................... 25
Supplemental Table S5 ..................................................................................................... 31
Supplemental Table S6 ..................................................................................................... 36
Supplemental methods

Whole exome data analysis

DNA samples from five family members (Supplemental Figure S1) were studied using whole exome sequencing. DNA was fragmented and enriched at the BGI (http://www.genomics.cn/en/index) by using NimbleGen SeqCap EZ Human Exome v3.0 kit. The enriched DNA was subjected to Illumina HiSeq sequencing. Clean reads were then aligned with the reference genome (hg19) and quality controlled according to the BGI Bioinformatics pipeline.

Single nucleotide variants (SNV) or small insertions or deletions (ins/dels) were detected (at least 50x coverage by reads, base call with Phred quality scores greater than 20) and filtered using Chipster. Data was analyzed using BEDTools software. Only private or rare (minor allele frequency (MAF) < 0.01) variants that were shared by the affected individuals and not present in the unaffected individuals or in the in-house exome set (N=71) were studied further. These remaining variants were functionally annotated using wANNOVAR to identify variants estimated to be pathogenic by SIFT, Polyphen-2 and Mutation Taster. Variants without annotations were not excluded. Variants were validated using capillary sequencing with ABI3500xL Genetic Analyzer (Applied Biosystems) from all family members to find alleles co-segregating with the phenotype. The frequencies for co-segregating variants in the general Finnish population were obtained from the The Exome Aggregation Consortium browser (ExAC, https://exac.broadinstitute.org/) and the Genome Aggregation Database (gnomAD, https://gnomad.broadinstitute.org/).

Expression of the ZNF528 in vitro

ZNF528 pcDNA3.1 constructs. To study the effect of the ZNF528 c. 1282C>T/p.Arg428* variant two expression constructs were made: a wild type ZFF528 (V5-ZNF528-WT) and a construct with the identified variant (V5-ZNF528-c.1282C>T). Briefly, the wild type ZNF528 cDNA was commercially acquired (GenScript, Piscataway, NJ, USA) and the c. 1282C>T/p.Arg428* variant was introduced
using QuickChange site directed mutagenesis kit (Stratagene, La Jolla, CA, USA). The cDNA was amplified using PCR, cut with XbaI and BamHI restriction enzymes and ligated into pcDNA3.1. (-) expression vector. V5-tag was introduced to the N-terminus. The constructs were amplified by transformation into XL1Blue *E.coli* strain and culturing overnight at 37°C in 100 ml of LB broth (1g/100ml NaCl, 1g/100ml Tryptone, 0.5g/100ml yeast extract) with thiamine, tetracycline and ampicillin. The constructs were extracted using QIAfilter maxi kit (Qiagen, Chatsworth, CA, USA) and verified by capillary sequencing.

**Cell culture.** Human embryonic kidney (HEK) 293 cells were cultured using 100 mm plates at 37°C, 5% CO2 in DMEM containing 10% FBS, 0.1% penicillin (Sigma-Aldrich, St Louis, MO) and 0.01% Fungizone (Cambrex Bio Science, Walkersville, MD). After 48 hours the cells were divided 1:3 and cultured for 24 hours and then plated on a six-well plate (3.5×10^5 cells/well). Prior to qPCR experiment cells were transiently transfected with V5-ZNF528-WT or V5-ZNF528-c.C1282T constructs using FuGENE HD transfection reagent (Promega, Madison, WI, USA). Empty vector was used as a control. Transfected cells were further cultured for 24 hours prior to the RNA extraction. Two 100 mm plates of each cell line were divided into two portions for RNA extraction creating four replicates of each sample. For western blot Lipofectamin 3000 (Invitrogen, Carlsbad, CA, USA) was used to transfect the HEK293 cells with V5-ZNF528-WT or V5-ZNF528-c.C1282T constructs. The Saos-2 osteosarcoma cells were cultured at 37°C and 5% CO2 in McCoy’s 5a Modified Medium (ATCC, Manassas, VA, USA) containing 15% fetal bovine serum. The cells were cultured on 100 mm plates for five days and then divided 1:2.

**Culturing primary skin fibroblasts.** Primary skin fibroblasts were cultured to obtain total RNA for qPCR analysis to study the expression of ZNF528 target genes in patient cells. Primary skin fibroblasts from two patients, and three age and gender matched unrelated controls were cultured at 37°C and 5% CO2 in Dulbecco’s modified Eagle medium (DMEM) with high glucose and GlutaMAX supplement (Gibco, Grand Island, NY, USA) containing 10% fetal bovine serum (FBS)
HyClone, Logan, UT, USA), non-essential amino acids (Sigma-Aldrich, St Louis, MO, USA), 5 % Sodium pyruvate (Sigma-Aldrich, St Louis, MO, USA), 0.1 % penicillin (Sigma-Aldrich, St Louis, MO, USA), 0.01 % Fungizone (Cambrex Bio Science, Walkersville, MD, USA) and 0.5 μg/ml sodium L-ascorbate (Sigma-Aldrich, St Louis, MO, USA). The primary skin fibroblasts were cultured using 100 mm CellBIND surface plates (Corning, LifeSciences, Lowell, MA, USA). After 48 hours the cells were divided 1:3 and cultured for 24 hours prior to the RNA extraction. Two 100 mm plates of each cell line were divided into two portions for RNA extraction creating four replicates of each sample.

**Real-time quantitative polymerase chain reaction (qPCR).** Total RNA was extracted from HEK293 cells and primary skin fibroblasts using E.Z.N.A Total RNA Kit (Omega Bio-Tek, Norcross, GA, USA) with RNase-free DNase (Omega Bio-Tek, Norcross, GA, USA) treatment and cDNA was synthesized using iScript cDNA Synthesis Kit (Bio-Rad, Hercules, CA). Real-time qPCR was carried out in duplicate using iTaq SYBR Green Supermix kit (Bio-Rad, Hercules, CA) in accordance with the manufacturer’s instructions in a CFX96 Real-Time System instrument (Bio-Rad, Hercules, CA). Data was analyzed using the 2(-Delta Delta C(T)) –method (39). Beta-actin (ACTB) and beta-2-microtubulin (B2M) were used as reference genes for HEK293 cells and hypoxanthine-guanine phosphoribosyltransferase 1 (HPRT1) and succinate dehydrogenase complex flavoprotein subunit A (SDHA) for primary skin fibroblasts.

**Western blot.** HEK293 cells were centrifuged in PBS buffer with 0.6M NaCl, 1 % Triton X-100 and a cocktail of protease inhibitors (Roche, Basel, Switzerland). For nuclear/cytoplasmic separation, we used the NE-PER™ Nuclear and Cytoplasmic Extraction kit (Thermo Scientific, Waltham, MA, USA). Loading buffer (Thermo Scientific, Rockford, IL, USA) was added and the proteins were denatured in 95°C for 5 minutes. Proteins were then separated by 10 % SDS-PAGE and transferred onto 0.45-µm PVDF membrane (Immobilon-P, Millipore, Bedford, MA, USA) using a Semi-Dry transfer cell system (Trans-Blot SD, Bio-Rad, Hercules, CA, USA). The membrane was blocked for
unspecific protein binding using 5 % nonfat milk in TBST (50 mM Tris-HCl, pH 7.5, 150 mM NaCl, 0.05 % Tween-20) for 30 minutes. The membrane was then exposed to antibodies (1:1000 dilution) V5-HRP (Invitrogen, Carlsbad, CA, USA) or β-actin-HRP (Abcam, Cambridge, UK). Chemiluminescence signal was imaged using a LAS-3000 Luminescent Image Analyzer (FujiFilm, Tokyo, Japan). The protein bands were normalized to β-actin and Image J software (National Institute of Mental Health, Bethesda, MD) were used to quantify immunoblots.

**Protein expression and subcellular localization of ZNF528**

*Lentiviral constructs, lentivirus production and infection of Saos-2 cell line.* Saos-2 cell lines stably expressing V5-ZNF528-WT and V5-ZNF528-c.1282C>T were created using lentivirus infection. The fragment of V5-ZNF528-c.1282C>T was inserted to the lentivirus vector pLVET-IRES-GFP, which harbors the GFP gene as an independent expression unit in its backbone. The V5-ZNF528-WT and V5-ZNF528-c.1282 C>T expressing lentiviruses were produced in HEK293T cells using the second-generation lentivirus packaging system. Briefly, 60 % – 80 % confluent HEK293T cells were co-transfected with lentiviral transfer vector (5 µg each), pMD2.G (envelope plasmid, 1.25 µg), pPAX2 (packaging plasmid 3.75 µg), using 20 µl Lipofectamine 2000 (Invitrogen, Carlsbad, CA, USA). The HEK293T cells were grown in low glucose DMEM (Invitrogen, Carlsbad, CA, USA) containing 10 % FBS, 0.1 % penicillin and streptomycin. The growth medium was replaced with fresh medium 24 hours after transfection and the virus-containing medium was collected every 12 hours up to six days. Lentiviruses were passed through 0.45 mm filter, snap frozen in liquid nitrogen and stored. For viral infection, the Saos-2 cells were seeded in 6-well plate 18 hours prior to infection and grown to 60 %–70 % confluency. The medium was replaced with lentivirus-containing medium with 8 mg/ml polybrene (Sigma-Aldrich, St Louis, MO, USA). After 24 hours of incubation the virus-containing medium was removed and replaced with fresh medium. The Saos-2 cell lines were sorted by FACS using BD FACSARia flow cytometer (BD Biosciences, Heidelberg, Germany).
Immunofluorescence. Saos-2 cells overexpressing V5-ZNF528-WT, V5-ZNF528-c.1282C>T were seeded on coverslips. Saos-2 cell line transfected with the empty vector was used as control. The cells were fixed in 4 % paraformaldehyde PBS buffer for 20 minutes, then quenched with 200 mM glycine in PBS for 20 minutes. The cells were further permeabilized with 0.1 % Triton X-100, blocking solution (0.5 % bovine serum albumin, 0.2 % fish gelatin in PBS) for 50 min at room temperature. The cells were incubated with primary V5 antibody (Invitrogen, Carlsbad, CA, USA) and blocking buffer overnight at 4°C and then rinsed three times with blocking buffer at room temperature. Samples were further incubated with fluorescent-conjugated secondary antibody Alexa488 (Invitrogen, Carlsbad, CA, USA) and blocking buffer for 1 hour and washed three times with PBS. The cells were incubated with 2-(4-amidinophenyl)-1H-indole-6-carboxamidine (DAPI) (Sigma-Aldrich, St Louis, MO, USA) to stain nucleus and with TRITC-Phalloidin (Sigma-Aldrich, St Louis, MO, USA) to stain actin filaments. The cells were washed once with PBS before being mounted with Immu-Mount (Thermo Scientific, Waltham, MA, USA). A Zeiss LSM 780 confocal microscope was used for confocal laser scanning images analysis, using a × 40 Plan-Apochromat objective, and analyzed by the ZEN 2011 software (Zeiss, Thornwood, NY, USA).

Cell viability and cytotoxicity assay

The cells of different groups were trypsinized and seeded into 96 well culture plates with 2 x 10³ each well. Cell viability was determined by using Cell Proliferation kit II (11465015001, Roche). We collected the data at the indicated time by measuring the absorbance at 450 nm according to the manufacturer’s instructions. The cytotoxicity of cells were tested by CellTox Green Cytotoxicity Assay kit (G8743, Promega). The cells from each group were resuspended as 4 x 10³ each well. The fluorescence was read with 485nm excitation source and 520nm emission filter. Data was obtained from triplicate wells and statistical significance was calculated by the two-tailed Student’s t test with equal variances.
Chromatin immunoprecipitation (ChIP) sequencing

Saos-2 cells were cross linked in formaldehyde (final concentration one percent) for 10 minutes, and 125 mM glycine was used to stop the reaction. A hypotonic lysis buffer (10 mM KCl, 10% glycerol, 20 mM Tris-Cl, pH 8.0, 2 mM DTT, and a cOmplete EDTA-free protease inhibitor cocktail (Sigma-Aldrich, St Louis, MO, USA) was used to suspend cell pellets followed by incubation at 4 °C for 50 minutes. The nuclei were suspended using SDS lysis buffer (50 mM Tris-HCl, 10 mM EDTA, pH 8.1, 0.5 % SDS, cOmplete EDTA-free protease inhibitor cocktail (Sigma-Aldrich, St Louis, MO, USA)). The chromatin was fragmented to an average size of 300 bp using a Q800R sonicator (Q Sonica, LLC., Newtown, USA). 70 µl of Dynabead protein G (Invitrogen, Carlsbad, CA, USA) was added to each sample. The samples were washed twice with blocking buffer (0.5 % BSA, 150 mM NaCl, 20 mM Tris-HCl, pH 8.0, 2 mM EDTA, 1 % Triton X-100, and cOmplete EDTA-free protease inhibitor cocktail (Sigma-Aldrich, St Louis, MO, USA)) and then incubated with 7 µg of antibodies against the target proteins or control IgG in 1000 ml blocking buffer for 10 hours. The chromatin lysate (250 mg) was mixed with bead/antibody complexes and incubated at 4 °C for 12 hours. The complex of antibodies/beads and chromatin was washed 10 times with RIPA buffer (50 mM Tris-HCl, pH 7.6, 150 mM NaCl, 1 mM EDTA, 0.1 % SDS, 1 % NP40, 0.5 % sodium deoxycholate). DNA-protein complexes were separated from beads by extraction buffer (10 mM Tris-HCl, pH 8.0, 1 mM EDTA, and 1 % SDS) and Proteinase K and NaCl were added followed by incubation overnight at 65 °C. DNA was purified by MinElute PCR Purification Kit (Qiagen, Chatsworth, CA, USA). The ChIP sequencing library was built by using the TruSeq ChIP Sample Preparation kit (Illumina, San Diego, CA, USA). The library was sequenced using a NextSeq550 sequencing system (Illumina, San Diego, CA, USA).

ChIP sequencing data analysis

Motif analysis and peak annotation. FastQC was used to assess the quality of raw sequence reads and Trimmomatic version 0.35 was used to remove adaptors and trim reads. Processed reads were next
aligned to the human genome (hg19) using BWA mapper. The creation of tag directories, peak calling and motif analysis were performed using HOMER. The peak regions across the genome hg19 were annotated and peak-associated genes and differentially expressed genes from RNA sequencing for V5-ZNF528-WT, V5-ZNF528-c.1282C>T were intersected.

*Gene-Ontology enrichment analyses.* Functional enrichment analysis was performed using HOMER. The analysis of distribution of ChIP sequencing peaks across the genome was performed using R package “ChIPseeker” version v1.18.0.

A total of 2260 and 29190 unique peaks for the V5-ZNF528-WT and V5-ZNF528-c.1282C>T respectively, were obtained after excluding 19 common binding sites. Genomic Regions Enrichment of Annotations Tool (GREAT) was employed to predict functions of cis-regulatory regions for ZNF528-WT ChIP-Seq and ZNF528-c.1282C>T ChIP-Seq unique binding sites. Gene regulatory domain was defined with the parameter “single nearest gene” and background regions was selected as “whole genome”. For large data sets, the option “significant By Region-based Binomial” that applies a binomial test over genomic regions was further selected.

In addition, re-analysis was performed for the ChIP sequencing data of wild type ZNF528 that is publicly available (http://zifrc.ccbr.utoronto.ca/) and reported earlier (34). The downloaded data included three different motif data sets: 1) the ZNF528 motifs predicted by bacterial one-hybrid recognition code (B1H-RC motifs), containing 3367 motifs; 2) the de novo motifs identified by MEME, containing 11 841 motifs and 3) the de novo motifs most similar to the B1H-RC motifs trimmed based on their alignment with the B1H-RC motifs (De Novo Similar to B1H-RC Trimmed Motifs), containing 8 009 motifs.

To investigate whether the ChIP sequencing peaks are linked to genes with any functional annotations, ontology analysis was performed using the GREAT. An enrichment analysis was carried out for Gene ontology biological process (GO BP) terms and Mouse and Human phenotypes was
performed. Amongst these pathways, genes that had the ZNF528 binding site and had a previously known connection to bone phenotypes were identified. These genes were considered as potential ZNF528 target genes for further analyses.

**RNA sequencing**

RNA was prepared from Saos-2 cells overexpressing either V5-ZNF528-WT or V5-ZNF528-c.1282C>T and control cells expressing empty vector each with three replicates. The quality and quantity of total RNA was measured by Agilent Bioanalyzer 2100 (Agilent Technologies, Santa Clara, CA, USA) with Eukaryote Total RNA Nano Kit (Agilent Technologies, Santa Clara, CA, USA), and Qubit RNA Broad Range kit (Life Technologies, Carlsbad, CA, USA). Illumina TruSeq Stranded mRNA Library preparation kit (Illumina, San Diego, CA, USA) was used according to the manufacturer's instructions. A Bioanalyzer 2100 was used for the assessment of quantification and quality of libraries with DNA 1000 Kit (Agilent Technologies, Santa Clara, CA, USA), qPCR KAPA Library quantification kit (Roche, Basel, Switzerland) and Qubit Broad Range DNA-kit. Illumina NextSeq550 platform was used to sequence the RNA libraries, with single-ended and 76 cycle mode. The FASTQ data was prepared within BaseSpace (Illumina, San Diego, CA, USA).

**Differential gene expression analyses**

DESeq2 version 1.22.2 was applied to perform differential gene expression analysis in R (version 3.5.2). For calling differentially expressed genes the false discovery rate (FDR) cutoff was set to < 0.01. The read counts were normalized with shifted logarithm transformation for plotting heatmaps using R package pheatmap (version 1.0.12).

Functional enrichment analysis using WikiPathways database was performed through HOMER software. We performed the gene overlapping analysis for genes from ZNF528 WT and ZNF528-c.1282C>T obtained from the overlap of RNA-Seq differentially expressed genes and ChIP-Seq peak annotated genes. Common overlapping genes were excluded and unique genes of ZNF528 WT and
ZNF528-c.1282C>T were then converted to Entrez Gene IDs and followed by the submission to HOMER “findGO.pl” function for the assessment of the enrichment of ontology categories. The “bg” parameter of the “findGO.pl” was set as default.

We checked the common overlapping genes between ZNF528 WT or ZNF528-c.1282C>T with osteoporosis-related genes from OsteoporosAtlas (http://biokb.ncpsb.org/osteoporosis/), which is a manually curated database for human osteoporosis-related genes. We then checked the statistical enrichment of osteoporosis-related genes in the ZNF528 WT or ZNF528-c.1282C>T unique gene lists using Pearson’s Chi-squared test in R (V 3.6.2).
Supplemental Figures

Supplemental Figure S1

Pedigree of the family with primary osteoporosis. Black symbols represent affected individuals. The individuals included in the whole exome sequencing are marked with asterisk.
Supplemental Figure S2

X-ray images and BMD of the affected family members. A) Son II1 B) Son II2 C) Father I2
**Supplemental Figure S3**

Raw RPKM expression correlation among two biological replicates of Saos-2 cell controls and experiments (Saos-2 overexpressing ZNF528 or ZNF528-c.1282C>T), respectively, from Saos-2 RNA-seq data.
Supplemental Figure S4

Western plot showing β-actin expression for the blot presented in Supplemental Figure S4. 1=V5-ZNF528-c.1282C>T, 2=V5-ZNF528-WT, 3=Empty vector control. The 3\textsuperscript{rd} replicate was used in making Figure 1C.
Supplemental Figure S5

Western blot showing protein expression of V5-ZNF528. 1=V5-ZNF528-c.1282C>T, 2=V5-ZNF528-WT, 3=Empty vector control. The 3\textsuperscript{rd} replicate was used in making Figure 1C.
**Supplemental Figure S6**

V5-ZNF528-WT and V5-ZNF528-c.1282C>T expression in nucleus and cytoplasm was observed in HEK293 cells using the Western blot method.

The uncropped Western blots are presented in Supplemental Figures S7-S8.

|                   | Cytoplasm | Nucleus |
|-------------------|-----------|---------|
| Empty vector      | +         | -       | -       | +         | -       | -       |
| V5-ZNF528 WT      | -         | -       | +       | -         | -       | +       |
| V5-ZNF528-c.1282C>T | -         | +       | -       | -         | +       | -       |

| Antibody          | Detection | Molecular Weight |
|-------------------|-----------|------------------|
| Anti-V5           |           | 73.5 KDa         |
| Anti-V5           |           | 47.1 KDa         |
| Anti-actin        |           | 37 KDa           |
Supplemental Figure S7

Western blots showing V5-ZNF528 expression in nucleus and in cytoplasm. Replicate 2 was used in the Supplemental Figure S6. EV=Empty vector, MUT = V5-ZNF528-c.1282C>T, WT=V5-ZNF528-WT
Supplemental Figure S8

Western blots indicating β-actin expression in nucleus and in cytoplasm. Replicate 2 was used in the Supplemental Figure S6. EV=Empty vector, MUT = V5-ZNF528-c.1282C>T, WT=V5-ZNF528-WT
Supplemental Figure S9

Absorbance measurements of the cell viability assay. Absorbance was measured at 450 nm.
Supplemental Figure S10

Results of the cell cytotoxicity assay. The fluorescence was read with 485nm excitation source and 520nm emission filter. Data was obtained from triplicate wells and statistical significance was calculated by the two-tailed Student’s t test.
Supplemental Figure S11

Functional annotation of unique ZNF528-WT (A) and ZNF528-c.1282C>T (B) ChIP-seq peaks. Mouse phenotypes’ ontology category contains data of mouse genotype–phenotype associations, and the x-axis values (in logarithmic scale) corresponds to the binomial raw (uncorrected) p-values.

### A

**GREAT category: Mouse Phenotypes**

| Phenotype                                | -log10(Binomial p value) |
|------------------------------------------|--------------------------|
| increased diameter of long bones         | 5.78                     |
| abnormal enteric neural crest cell migration | 3.79                    |
| absent corpus callosum                    | 3.69                     |

### B

**GREAT category: Mouse Phenotype Single KO**

| Phenotype                                | -log10(Binomial p value) |
|------------------------------------------|--------------------------|
| absent craniofacial bones                | 8.11                     |
| arched palate                            | 8.87                     |
| abnormal mitral cell morphology          | 8.66                     |
| ciliary body coloboma                    | 8.36                     |
| abnormal neural fold elevation formation | 7.81                     |
| failure of conotruncal ridge closure     | 7.71                     |
| increased osteochondroma incidence       | 7.55                     |
| subarterial ventricular septal defect    | 7.31                     |
| decreased spleen iron level              | 7.26                     |
| absence of all nails                     | 7.08                     |
| everted embryonic neuroepithelium        | 6.71                     |
| decreased skin tensile strength          | 6.50                     |
| short metatarsal bones                   | 6.50                     |
| altered response to retinal ischemic injury | 5.50                     |
| absent styloid process                   | 5.04                     |
| absent hyoid bone lesser horns           | 4.87                     |
| absent neutrophils                       | 4.84                     |
| lipodystrophy                            | 4.81                     |
| abnormal trunk neural crest cell morphology | 4.81                    |
| increased trunk neural crest cell apoptosis | 4.81                    |

---

**NOTE:** The data represents the logarithm of the binomial p-value. Higher values indicate a stronger association between the genotype and phenotype.
Supplemental Figure S12

Pathway enrichment analysis for ZNF528-WT unique target genes and statistical significance for osteoporosis genes. (A) Wikipathways enrichment analysis for ZNF528-WT targeted unique genes. X axis was denoted as logarithmic p value. (B) Intersection of osteoporosis genes with ZNF528-WT and ZNF528-c.1282C>T unique genes defined by RNA-Seq and ChIP-Seq. (C) Statistical significance test for the fraction of osteoporosis genes in ZNF528-WT and ZNF528-c.1282C>T by Chi-square test.

### B

|                                      | Genes in total | Overlapping with osteoporosis genes | Non-overlapping with osteoporosis genes |
|--------------------------------------|----------------|-------------------------------------|----------------------------------------|
| ZNF528-WT unique target genes        | 187            | 2                                   | 185                                    |
| defined by RNA-Seq and ChIP-Seq      |                |                                     |                                        |
| ZNF528-c.1282C>T unique target      | 1815           | 93                                  | 1722                                   |
| genes defined by RNA-Seq and ChIP-Seq|                |                                     |                                        |
| Common target genes of ZNF528-WT     | 180            | 12                                  | 168                                    |
| and ZNF528-c.1282C>T defined         |                |                                     |                                        |
| by RNA-Seq and ChIP-Seq              |                |                                     |                                        |
**Supplemental Tables**

**Supplemental Table S1**

Summary of the clinical information.

| ID  | Sex  | Age (years) | BMD lumbar spine z-score | BMD proximal hip z-score | Height (cm) | Weight (kg) | Other illnesses                       | Medications |
|-----|------|-------------|--------------------------|--------------------------|-------------|-------------|---------------------------------------|-------------|
|     |      |             |                          |                          |             |             |                                       |             |
| **Family members** |      |             |                          |                          |             |             |                                       |             |
| I2  | Male | 54          | -0.7                     | 0.3                      | 170         | 76          | None                                  | None        |
| I3  | Female | NA         | NA                       | NA                       | NA          | NA          | NA                                    | NA          |
| II1 | Male | 21          | -4.6                     | -0.9                     | 170         | 64          | None                                  | None        |
| II2 | Male | 23          | -3.1                     | -0.2                     | 170         | 63          | None                                  | None        |
| II3 | Male | 14          | -0.7                     | -0.1                     | 171         | 62          | None                                  | None        |
| I1  | Male | 67          | -0.3                     | -0.2                     | 179         | 74          | NA                                    | NA          |
| I4  | Female | 67         | -0.5                     | -0.5                     | 160         | 45          | Asthma, Celiac disease                | NA          |
| I5  | Female | NA         | NA                       | NA                       | NA          | NA          | NA                                    | NA          |
| **Fibroblast cell lines** |      |             |                          |                          |             |             |                                       |             |
| RHH109 | Male | 28          | NA                       | NA                       | NA          | NA          | None                                  | None        |
| RHH110 | Male | 56          | NA                       | NA                       | NA          | NA          | None                                  | None        |
| BR1334 | Male | 23          | NA                       | NA                       | NA          | NA          | None                                  | NA          |

Abbreviations: NA= Not available
Supplemental Table S2

qPCR primers

| Target  | Primer 1 | Primer 2 |
|---------|----------|----------|
| COL1A2  | CAAGGACAAGAAACACGTCTGGCTAGGAGAAA | CAGGCGCATGAAGCGAAGTTGAGTTG |
| CTNNB1  | TTGTGGCGGCGCCATTCTTGAAG | TCCTCAGACCTCTCTCCGTCC |
| CYLD    | TGGCAACTGGGATGGAAGAT | CTCCCCGCGTCACACT |
| DKK1    | TGACAACACGAGCCGTACC | CAGGCGAGACAGATTTGCA |
| ESR1    | AAAGGATGGGATACGAAAAGACC | CTCATGTCTCCAGCAGACCC |
| JAG1    | CTGCGAGCCAAGGTGTGTG | GGGTGCACTTGGGTCTC |
| LFNG    | CCGCAACAAGGAGATGCAGTTT | ATACTCCACGGCCCATCCTTC |
| LRP5    | ACCGTACAGGGCCCTACATCA | TGGATAGGGGTCTGAGTCC |
| MEF2C   | ACGATTCCGCTAGGTCAAGC | GAGGCGGTGAACAGCAGAAC |
| PLOD2   | GAGAAGCCCTCGAGCATCC | TGAACAGGACCTTCACAGT |
| RSPO3   | GGAACCTTAGTGTCACACAC | TCCAGACTTTTGCTTGAGGT |
| RUNX2   | GAAACCAGAAGGCACACACA | GAGATCGGAATGCGCCCTAA |
| SOX4    | CCTTTCATTCCGAGACGCGA | GTGCCGCGACTTCACCTT |
| SOX5    | CCCACGATGAGCAGAAGAA | GCTGACCTTGAACCTGGATC |
| SOX9    | GCTCTGGGAGACTTCTGAAACGA | CCGTTCTTCACGACTTCT |
| RANKL   | TCCATGTGGCCTGCCCTCC | AGGATCGCATCTTCGGCTCG |
| WLS     | GAGGCTTGATTGTGCTCAGG | CCCAGGCACGAAACCATT |
| ZNF528  | AGTCGCGCCATTTTACATGCAG | TGGGTATGAGCGTGAAC |

24
### Supplemental Table S3

Expression of ZNF528 in bone and connective tissue cell lines.

| Database                                      | Expression                                      |
|-----------------------------------------------|-------------------------------------------------|
| Expression Atlas                              | Ewing's sarcoma cell line, Bone                 |
|                                               | Chondrosarcoma cell line, Cartilage             |
| The Genotype-Tissue Expression (GTEx) v5      | Human Cells-Transformed Fibroblasts              |
| Cell Line Navigator                           | Bone cell lines (MG63, RDES, SJSA1, SW1353)     |

### Supplemental Table S4

Results of enrichment analysis of the Najafabadi et al. 2015 data. Bone related terms are bolded.

| Term name | Binom Raw | Term name | Binom Raw | Term name | Binom Raw |
|-----------|-----------|-----------|-----------|-----------|-----------|
| GO BP     | P-Value   | GO BP     | P-Value   | GO BP     | P-Value   |
| B1hRC motifs | Trimmed De novo motifs | De novo motifs |
| hepaticobiliary system development | cranial suture morphogenesis | positive regulation of chondrocyte differentiation | 2.60E-11 | 4.53E-06 | 4.24E-06 |
| liver development | craniofacial suture morphogenesis | middle ear morphogenesis | 2.61E-10 | 1.14E-05 | 9.83E-06 |
| embryonic cranial skeleton morphogenesis | negative regulation of epidermal cell differentiation | regulation of cell proliferation | 4.76E-10 | 1.68E-04 | 2.78E-05 |
| intermediate filament-based process | negative regulation of epidermis development | cranial suture morphogenesis | 1.04E-08 | 4.14E-04 | 3.31E-05 |
| intermediate filament cytoskeleton organization | cardiac neural crest cell development involved in heart development | negative regulation of epidermal cell differentiation | 6.79E-08 | 6.58E-04 | 5.87E-05 |
| toll-like receptor 3 signaling pathway | type B pancreatic cell differentiation | craniofacial suture morphogenesis | 3.41E-07 | 8.62E-04 | 7.25E-05 |
| potassium ion homeostasis | cardiac neural crest cell differentiation involved in heart development | metanephric nephron morphogenesis | 3.70E-07 | 1.13E-03 | 1.10E-04 |
| MyD88-independent toll-like receptor signaling pathway | positive regulation of stem cell differentiation | | 4.03E-07 | | 1.39E-04 |
| TRIF-dependent toll-like receptor signaling pathway | desmosome organization | | 4.88E-07 | | 2.70E-04 |
| Endochondral bone morphogenesis | 5.81E-07 | Myoblast fusion | 2.99E-04 |
|--------------------------------|----------|----------------|----------|
| Hindbrain morphogenesis        | 2.14E-06 | Cartilage development involved in endochondral bone morphogenesis | 3.81E-04 |
| Toll-like receptor 5 signaling  | 2.37E-06 | Negative regulation of epidermis development | 4.14E-04 |
| Toll-like receptor 10 signaling | 2.37E-06 | Positive regulation of heart rate by epinephrine-norepinephrine | 7.37E-04 |
| Embryonic digestive tract       | 1.87E-05 | Glomerular visceral epithelial cell differentiation | 1.10E-03 |
| Cerebellar cortex formation     | 2.74E-05 | Embryonic foregut morphogenesis | 1.24E-03 |
| **Positive regulation of**      | 3.50E-05 | Glomerular epithelial cell differentiation | 1.35E-03 |
| Chondrocyte differentiation     |          | Positive regulation of the force of heart contraction by chemical signal | 1.39E-03 |
| **Positive regulation of BMP**  | 4.03E-05 | Foregut morphogenesis | 1.67E-03 |
| Signaling pathway               |          | Positive regulation of blood pressure by epinephrine-norepinephrine | 2.99E-03 |
| Term Name                                | Binom Raw P-Value | Term Name                                | Binom Raw P-Value | Term Name                                | Binom Raw P-Value |
|------------------------------------------|-------------------|------------------------------------------|-------------------|------------------------------------------|-------------------|
| arrest of tooth development              | 1.15E-14          | abnormal metatarsal bone morphology      | 7.09E-08          | abnormal lens induction                  | 8.26E-07          |
| decreased cranium height                 | 8.56E-08          | abnormal tendon morphology               | 9.90E-08          | abnormal metatarsal bone morphology      | 4.62E-06          |
| abnormal metacarpal bone morphology     | 2.45E-07          | decreased trabecular bone thickness      | 5.05E-07          | abnormal tendon morphology               | 5.84E-06          |
| chondrodystrophy                         | 5.44E-07          | abnormal lens induction                  | 1.91E-06          | small limb buds                         | 9.82E-06          |
| abnormal Purkinje cell differentiation   | 6.60E-07          | decreased gamma-delta T cell number      | 5.20E-05          | increased vertebrae number               | 3.85E-05          |
| abnormal metatarsal bone morphology     | 7.77E-07          | abnormal cerebellum posterior vermis     | 1.40E-04          | abnormal midbrain-hindbrain boundary     | 4.95E-05          |
| short metatarsal bones                   | 2.25E-06          | increased mean corpuscular hemoglobin    | 1.76E-04          | abnormal ureter smooth muscle morphology | 5.30E-05          |
| Condition                              | p-value | Related Condition                                              | p-value | Related Condition                                              |
|----------------------------------------|---------|-----------------------------------------------------------------|---------|-----------------------------------------------------------------|
| abnormal PP cell morphology            | 5.37E-06| abnormal midbrain-hindbrain boundary morphology                | 2.53E-04| abnormal cochlear nerve compound action potential              |
| failure of palatal shelf elevation     | 6.24E-06| abnormal cerebellum vermis lobule morphology                   | 5.58E-04| abnormal ischium morphology                                     |
| wavy neural tube                       | 8.77E-06| bile duct proliferation                                         | 7.58E-04| conotruncal ridge hypoplasia                                    |
| abnormal lens induction                | 2.73E-05| abnormal PP cell morphology                                    | 1.18E-03| short frontal bone                                              |
| hyperresponsive to tactile stimuli     | 5.04E-05| calcified joint                                                 | 1.23E-03| calcified joint                                                 |
| increased susceptibility to noise-induced hearing loss | 5.04E-05| bleb                                                           | 1.82E-03| short metatarsal bones                                          |
| spinning                               | 5.15E-05| decreased circulating noradrenaline level                      | 2.28E-03| absent femur                                                    |
| abnormal mesenchyme morphology         | 5.82E-05| abnormal melanoblast morphology                                | 2.57E-03| non-pigmented tail tip                                         |
| abnormal foregut morphology            | 5.98E-05| anterior iris synechia                                         | 2.61E-03| decreased cochlear nerve compound action potential            |
| common atrium                          | 8.44E-05| increased presacral vertebrae number                           | 2.88E-03| absent tail                                                     |

29
| Term Name                          | Binom Raw | P-Value |
|-----------------------------------|-----------|---------|
| Abnormal bile duct morphology     | 1.15E-04  |         |
| Abnormal cochlear hair cell       | 1.24E-04  |         |
| Decreased hepatocyte proliferation | 1.46E-04  |         |
| Abnormality of the pubic bones    | 3.62E-06  |         |
| Aplasia/Hypoplasia of the pubic bone | 3.75E-05 |         |

**Abbreviations:** GO BP=Genome ontology Biological Process
Supplemental Table S5

Genes identified in the pathway enrichment analysis for RNA-seq and ChIP-seq data (related to Supplemental Figure 12S B).

| ZNF528-WT unique target genes | Gene name |
|-------------------------------|-----------|
| Osteoporosis related genes (N=2) | LBR, CABIN1 |
| Other genes (N=185) | ADAM10, ANGPT1, SLC25A4, ANXA5, ARNT, KLF9, CCNG2, CDH11, CDKN3, HAPLN1, DPYD, ECT2, MEGF9, EIF2S1, EPS15, GRIK4, ARHGAP35, GUCY1A2, GY1, HLA-H, HMGBl, ID3, IDS, INGI, EIF3E, ITGB1, KCNH1, KPN1A, IPOP5, SMAD2, MAN2A1, MEF2A, MLH1, MYC, NDUF4, NPTX2, PPIC, PPP1R8, PPP2R1B, PKN2, PSMB5, PSMC5, PTTR1J, RARB, RBBP7, ROS1, RPL11, RPL39, RPS29, SIM1, SLC1A4, SNTB1, TBX15, TFAM, TLE4, XPNPEP1, ZNF79, LUZP1, NUP214, BCAR3, B4GALT4, EBAG9, SLC16A7, SRSF11, ARHGAP29, SEC22B, PREPL, PDE4DIP, CEP135, PHF14, CKAP5, ZBTB24, FRY, MBNL2, TRIB1, RCBNA2, STAG1, PAIP1, ARPP21, RAB10, RAPGEF4, KRRI, BNN3A2, MTF2, PP1M1E, ENPP4, ANKR6D, TMCC1, MON2, KDM4C, MAST2, TSPYL4, FNB4P, LARP1, SSBP2, OBPL3, SITRKR5, TES, AS5, MADDHC, USP25, HILPD1, H1N, ZNF639, ACP6, FAM198B, NCKIPS, VPS54, PELO, GAR1, BCA53, TBC1D8B, CDKAL1, ORX1, PRMT6, RHO1, CV3, FAM46A, VPS50, DOC1K0, ZNF407, LRR4C0, SYBU, VAC14, C1orf112, ACSS2, ZC4H2, KIF15, ARFGEF3, HACE1, SHROOM3, RBM26, AT1L2, GNPNAT1, TMEM327, TRAK2, DCTPP1, BHLHE41, GCC1, ZYG11B, CSP1, ZNF676P, RPF1, C1orf21, ANP32E, EIF2A, DYNBP1, ZNF644, RELT, KLH31, ACBD5, MTDH, EFHC1, GBP5, SPIN4, MACROD2, E2F7, LACC1, EFCAB3, FAM75B, TTL, LING02, ZNF782, TMTC3, C1orf54, DCLK2, ZNF800, DXH36, TXLNA, LCCR1, NAAADL2, COL24A1, G1K, ZNF328, TRIM59, CHSY3, ZNF678, EML6, CASC15, GOLGA8B, C4orf47, MROH1, FAM160A1, LVCAT1 |

| ZNF528-c.1282C>T unique target genes | Gene name |
|-------------------------------|-----------|
| Osteoporosis related genes (N=93) | AGT, ANXA2, AR, RHOA, BCL2, BCL2P, BMP6, RUNX3, CD44, CNR1, CPE, CSF1, CTSB, CTSL, CTSS, TSC22D3, LPAR1, EGRF, FRZB, GALNT3, GGCX, GSK3B, GSN, HILLS, HIF1A, HSPA8, HSP90AA1, IBSP, ID4, IDH2, IGF1R, IGFBP3, IGFBP5, IL6R, CXCL18, JUN, ANOS1, LIF, LIFR, LMNA, LRP5, SMAD6, MMP2, NF1, SERPINE1, PLS3, PLXNA2, PPAR, PPAP, PPIB, PTCH1, PTGER4, PTN, PTPN1, PTNP1, RSU1, ATXN1, SH3BP2, SPARC, SPP1, SPTBN1, STAT3, THBS1, TLR4, TXNRD1, VCL, NRP1, RECK, OGT, PDE8B, USP8, HS6ST1, WDR1, HDAC5, SPRY1, CAP1, SEMA4D, KHDRBS1, ADAMTS5, DKK1, KDM4B, SIRT1, ZBTB20, PLEKHO1, VPS53, SOX6, ANKH, MEPE, TWSG1, ARHGAP31, HIVEP3, VPS13B |
| Other genes (N=1722) | AARS, ABCA1, ACACA, ASIC2, ACLY, ACTC1, ACTN4, ACTN1, ARDAB1, ADCY7, ADCY9, ADPRH, PARP4, ADRA1D, ADRA1B, AP2B1, AGA, AGL, AIM1, AK4, ALCAM, AKR1B1, AMFR, AMPH, ANK1, ANXA1, ANXA6, APBP2, BIRC3, BIRC5, APLP1, APO9, ARF6, ARG2, RHOB, ROHOC, RND3, ARSA, STS, ASNS, ASTN1, ZFHX3, RERE, ATP1A1, ATP1B1, ATP1B3, ATP2A2, ATP2A3, ATP2B4, AXL, BACH1, BARD1, BCA1, BCL3, BDNF, BMP3, BMP8B, BNP3, BRCA2, ZF3F5L2, BTD, TSP0, CAPN5, LDRAD4, PTTG1IP, TMEM50B, DDR1, CACNB4, CAD, CALD1, CALU, CANX, CAPN2, CAPZB, CARS, CAV1, CAV2, RUNX1, RUNX1T1, SERPINH1, CCNH, CCNT2, CD86, TNFRSF8, CD36, CD63, LRRB, CDC25B, CDC25C, CDH2, CDK8, CDKN2B, CDKN2C, CLGN, CEBPD, CEBPG, CENPE, CFL2, CTSC, GCC1, FOXL3, CHM, CHRM2, CKB, |
Shared target genes

**Osteoporosis related genes (N=12)**

APP, BMP4, COL1A2, EDN1, FYN, GJA1, GNAQ, HSPG2, IL6ST, NFkB1, NOTCH2, PLOD2

**Other genes (N=168)**

ACTA2, ADCY8, ANK3, DST, CCKAR, CDC27, CHD1, COL5A2, COL11A1, COL12A1, CTBP2, DAP, DCN, DMLX1, DDX10, DOCK3, ECE1, EPAS1, GABRE, GBE1, FOXXN2, IFRD1, JARID2, KPNA2, AFF3, LGALS3BP, LSAMP, MAGEB2, ORC5, PAM, PDE4D, PHKB, PLCB4, PPP2R3A, PRKCA, REST, SATB1, SLC4A3, SLC8A1, SPOCK1, SQUE, SRPK2, NR2F2, TFDP2, TPM1, VLDLR, TRIM26, ZDCA, CDC7, FZD1, DYSK, LGR5, CDC14A, EIF3H, GMPS, EIF2S2, WASL, PDCD5, XPR1, CD83, MAGED1, CLOCK, VGLL4, PHACTR2, TBC1D5, MAFB, PIGK, GPC6, EDIL3, ARL4C, PLXNC1, DSCR3, ANAPC10, FAM3C, PDLIM5, Sep-09, MRPS30, WDR3, DBF4, KERA, FAF1, FSTL1, AKAP11, KLF12, MGAT4A, CNKS2R, LIMCH1, PALLD, TNIK, EFR3A, ADRG1L, NEDD4L, RBFOX2, DDAH1, ORC3, MOX1D1, ZZZ3, TANC2, KCNV1, MTBP, ASAP1, HP1B3, NRR1, SENX7, LSM8, SRBD1, VPS13D, BMP2K, PNRC2, DEPDC1, TENM3, ZNF83, LMBRD1, AJAP1, CCL28, CNNQ5, UGTT1, DPYSL5, SLC39A10, NIPAL3, VANGL2, TEM2, KIDINS220, KLH8, LRRN1, PTBP2, SLC25A19, SPCS3, PAPPA2, SLC22A3, CLSTN2, KCTD15, HMGN5, BORA, FRAS1, DDHD1, ITFG1, FAM107B, CDC7, PLEKHA8, LTV1, ZC3H12C, SCIN, SESTD1, PXYP1L1, OSBP1L8, KCTD12, SKE2IP, PDZD8, AN04, FRMD6, WTIP, B3GAT2, LINC00662, CNTN4, CNKSR3, TMEM65, ZXDB, TMTC2, DGKH, PHACTR1, WDR27, ST6GALNAC3, UBALD2, CCDC141, CCDC18, TMEM170B, LOC101927686
### Supplemental Table S6

Connective tissue related genes with ZNF528 binding motif based on Najafabadi et al. data *

| Gene      | Phenotype                                      | Distance of ZNF528 binding motif to TSS [bp] |
|-----------|------------------------------------------------|---------------------------------------------|
| COL1A2    | OI types II, III, IV                           | -21,896                                    |
| CTNNB1    | BMD, Fracture risk, WNT signaling              | -568,113, -467,306, -467,077, +220,448     |
| CYLD      | NF-kappaB pathway, Osteolysis                  | +104,447                                   |
| DKK1      | WNT signaling, BMD, Fracture risk              | -207,126                                   |
| ESR1      | BMD                                            | +397,102                                   |
| JAG1      | BMD, Fracture risk                             | +77,240, +186,068                          |
| LFNG      | vertebral column development                   | +937                                       |
| LRP5      | BMD, Fracture risk, Osteoporosis-              | -9,131                                      |
|           | pseudoglioma syndrome,                         |                                            |
| ME2FC     | BMD                                            | -694,345, -694,337, -694,322, +14,476, +14,677 |
| PLOD2     | Bruck syndrome type 2                          | -84,033, -83,670                           |
| RSPO3     | BMD                                            | -5,411                                     |
| RUNX2     | Cleidocranial dysplasia                        | -251,531                                   |
| SOX4      | BMD                                            | -917,884, -697,984, -530,334               |
| SOX5      | Kashin-Beck disease                            | -633,819, -517,701, -432,389, -432,649, -257,102, -42,412, +392,340, +515,055, +618,870, +619,029 |
| SOX9      | BMD, Campomelic dysplasia                      | -857,010, -793,260, -504,675, -90,106, +100,912, +101,929, +286,116, +290,959, +388,680 |
| RANKL     | BMD                                            | +322,522                                   |
| WLS       | BMD                                            | -181,295, -145,825, +154,625               |

Abbreviations: bp = base pair, TSS = Transcription Start Site

* Najafabadi HS, Mnaimneh S, Schmitges FW, Garton M, Lam KN, Yang A, Albu M, Weirauch MT, Radovani E, Kim PM, Greenblatt J, Frey BJ, Hughes TR 2015 C2H2 zinc finger proteins greatly expand the human regulatory lexicon. Nat Biotechnol 33:555-562