The American Journal of Human Genetics, Volume *102*

Supplemental Data

An Osteoporosis Risk SNP at 1p36.12 Acts as an Allele-Specific Enhancer to Modulate *LINC00339* Expression via Long-Range Loop Formation

Xiao-Feng Chen, Dong-Li Zhu, Man Yang, Wei-Xin Hu, Yuan-Yuan Duan, Bing-Jie Lu, Yu Rong, Shan-Shan Dong, Ruo-Han Hao, Jia-Bin Chen, Yi-Xiao Chen, Shi Yao, Hlaing Nwe Thynn, Yan Guo, and Tie-Lin Yang

Supplemental Figures and Legends

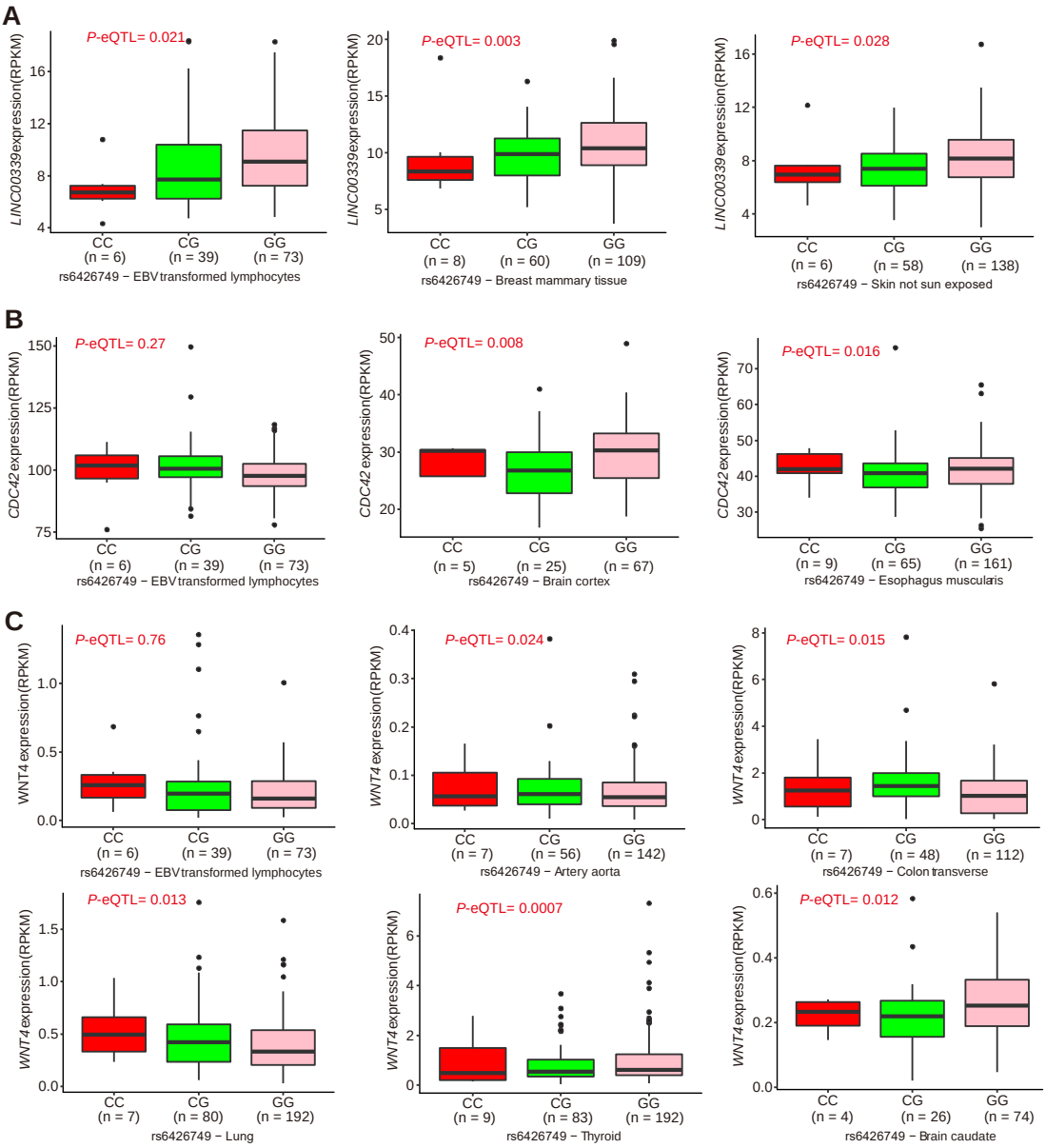

Figure S1. eQTL analyses for rs6426749 with *LINC00339*, *CDC42* and *WNT4* from GTEx

We checked the eQTL association from the Genotype-Tissue Expression (GTEx) database,[1] including 7,051 samples from 449 donors across 44 tissues. The eQTL analysis from GTEx Project[1] between rs6426749 and *LINC00339* (A), *CDC42* (B) or *WNT4* (C) on LCLs and other significant tissues ($P < 0.05$) are shown.

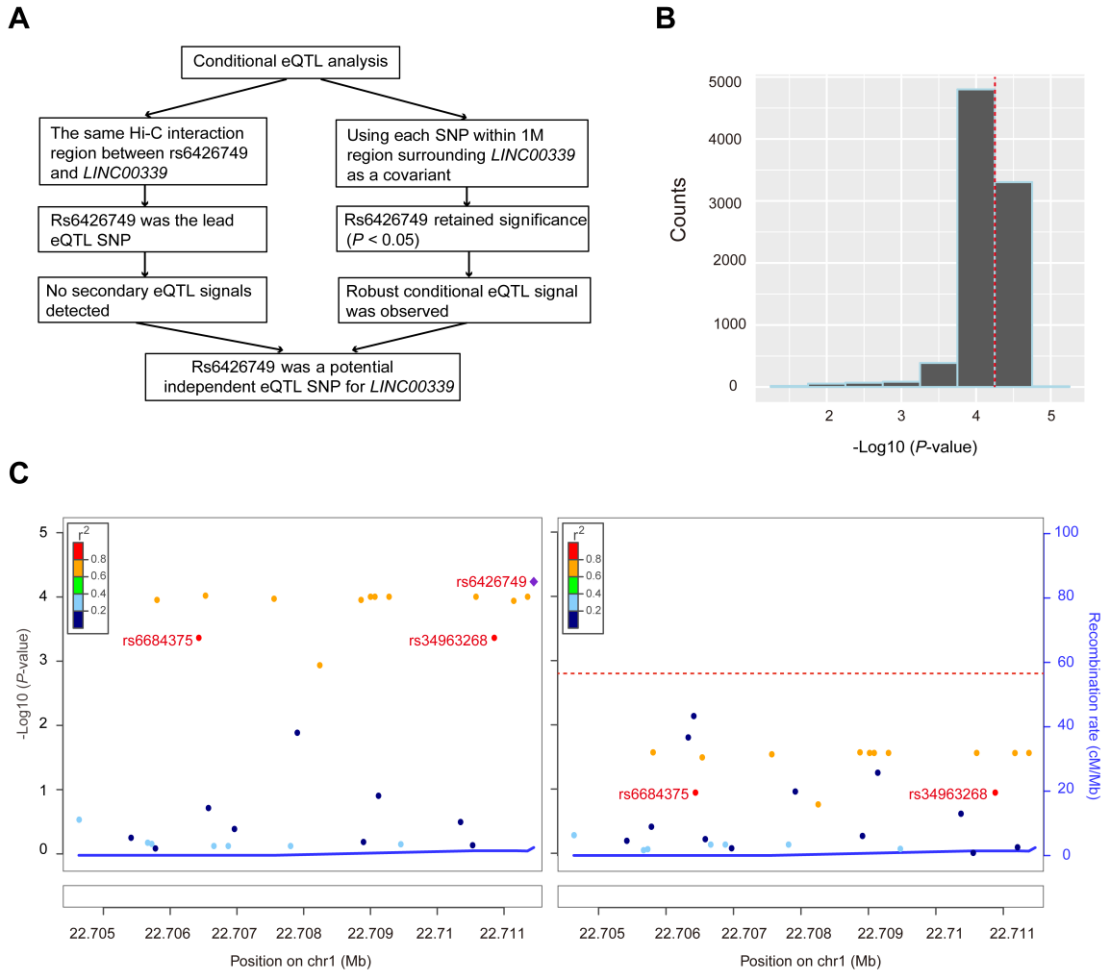

Figure S2. Conditional eQTL analysis for rs6426749

(A) Flowchart of conditional eQTL analysis. (B) Distribution of conditional eQTL signals for rs6426749 on *LINC00339* using each SNP within 1M region surrounding *LINC00339* as a covariant. Raw eQTL signal for rs6426749 with *LINC00339* was marked with dashed red line. (C) Raw (left) and conditional (right) eQTL signal (using rs6426749 as a covariant) for all SNPs within the same Hi-C interaction region with rs6426749. The dashed line represents significant association level of secondary eQTL SNPs (Bonferroni adjusted *P*-value < 0.05). Conditional eQTL analysis was performed by fitting the selected cis-eQTL SNP genotype as a covariate and testing for the secondary association retained using ANNOVA. Bonferroni correction was applied to determine the significance of secondary eQTLs.

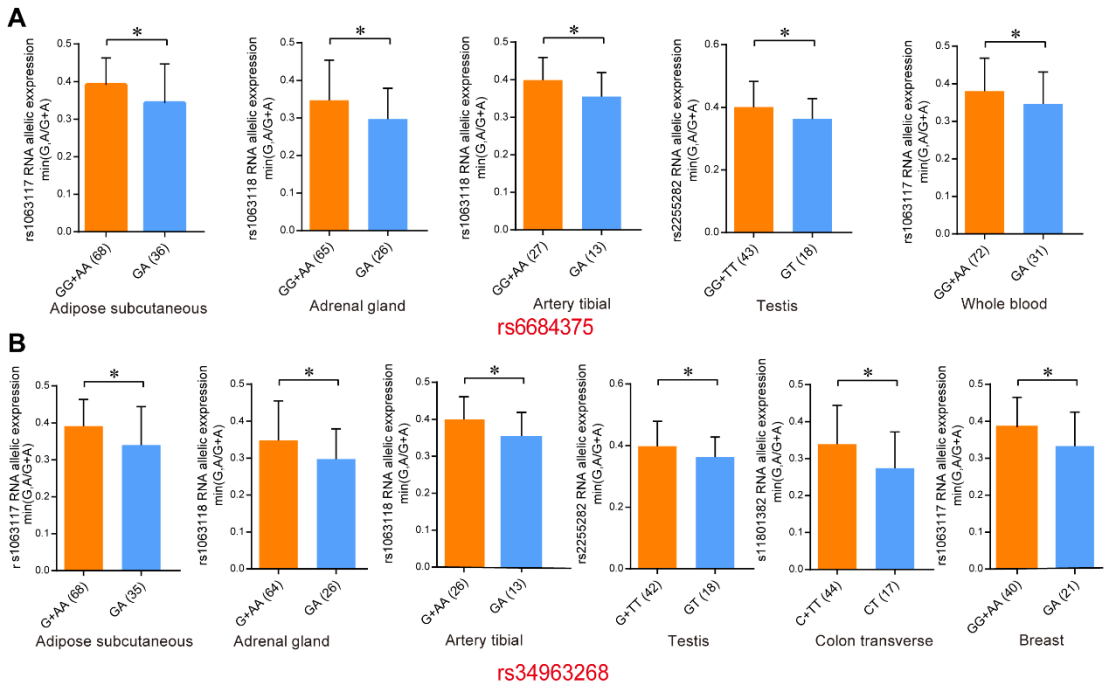

Figure S3. Allele specific expression analysis result on *LINC00339* for rs6684375 and rs34963268

Allele specific expression (ASE) analysis between rs6684375 (A) or rs34963268 (B) and *LINC00339*, using monoallelic gene expression data from GTEx[1]. Only significant tissues ($P < 0.05$) are shown. Error bars, s.d. *$P < 0.05$ as determined by Wilcoxon rank sum test.

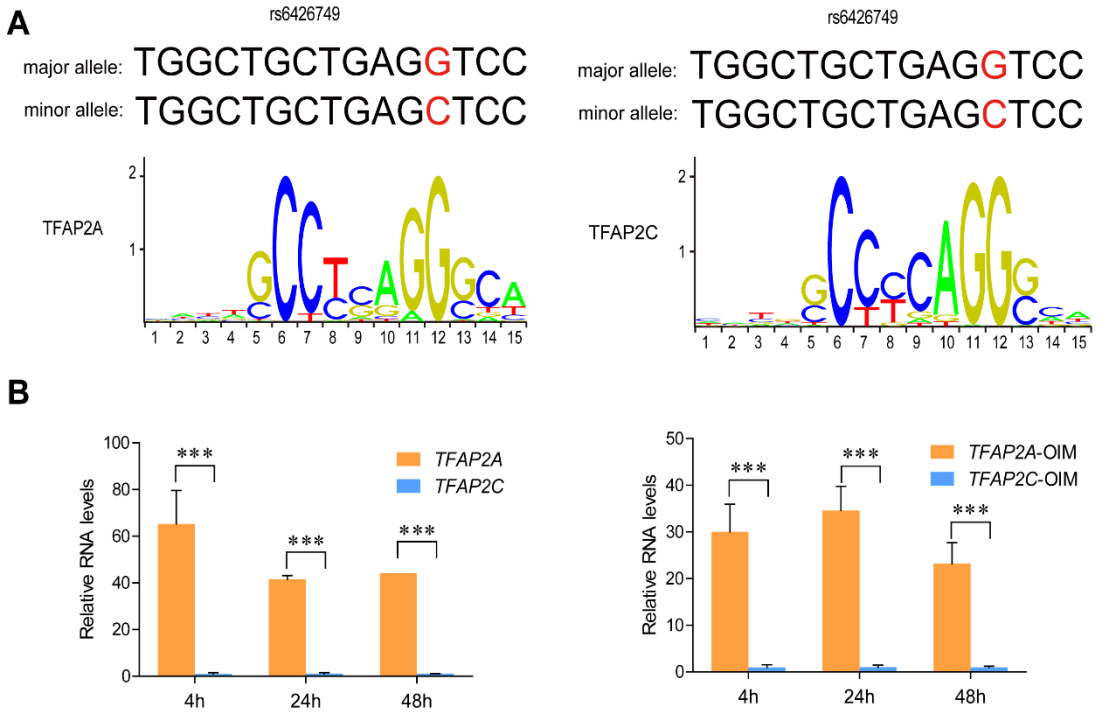

Figure S4. *TFAP2A* is predominantly higher expressed than *TFAP2C* in hFOB1.19 cells

(A) Motif predictions for rs6426749. Two TFAP family motifs (TFAP2A, TFAP2C) were predicted to exclusively bind to major allele (G) of rs6426749. (B) Comparison of mRNA expressions for *TFAP2A* and *TFAP2C* in hFOB1.19 cells undergoing spontaneous differentiation or with the effect of Osteogenic Induction Media (OIM). RNA expression data was extracted from GEO database (GEO: GSE75232).[2] Relative mRNA expression levels were normalized by equalizing *TFAP2C* expression levels to 1. Error bars, s.d. ***$P < 0.001$ as determined by an unpaired, two-tailed Student's t-test.

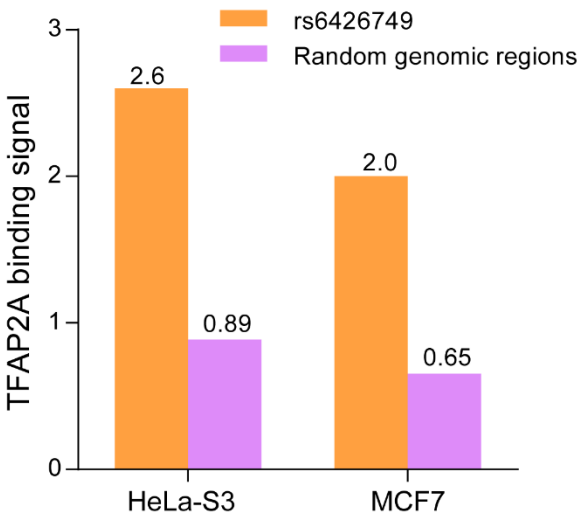

Figure S5. Comparison of TFAP2A binding on rs6426749-locus and random genomic regions

We compared the average TF binding signal surrounding rs6426749 (50-bp) with 1,000,000 randomly chosen genomic regions from 22 autosomes of the same length. Average TFAP2A binding signals for rs6426749-locus and 1,000,000 random genomic regions in HeLa-S3 cells and MCF7 cells were shown, respectively.

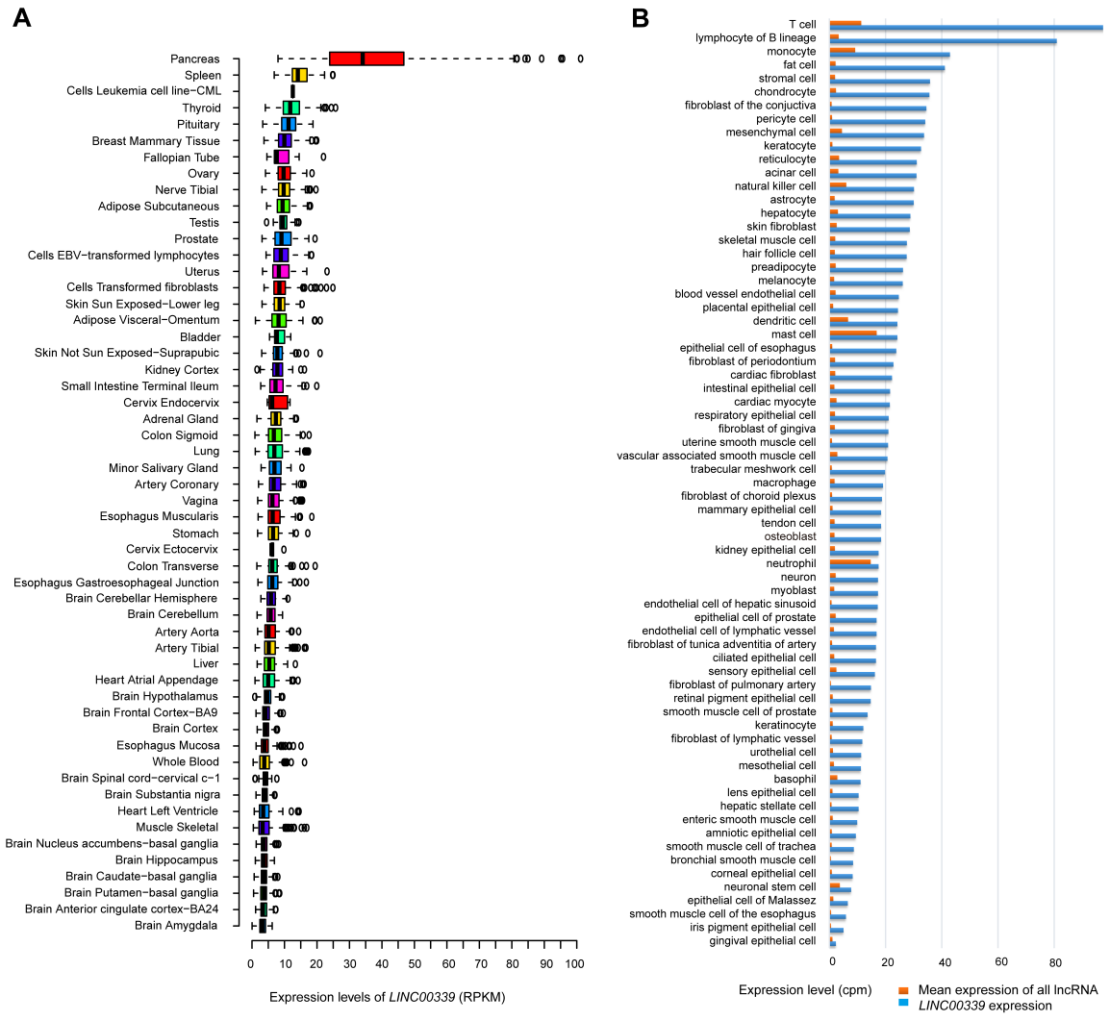

Figure S6. *LINC00339* is extensively expressed in both tissue and cell levels

LINC00339 expression levels were analyzed in 54 diverse tissues from GTEx Project[1] (A) and 69 diverse primary cells from FANMOT5[3] (B), respectively. cpm, counts per million.

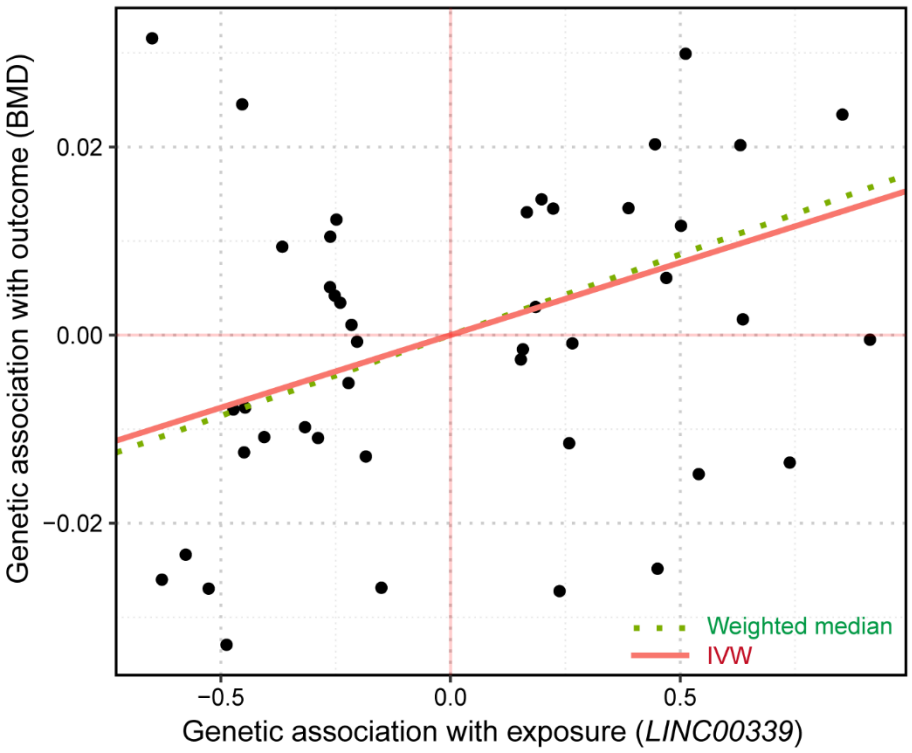

Figure S7. Scatter plot of genetic association with *LINC003339* against association with BMD

Each dot indicated one of 44 total genetic variants used as instrumental variables for multi-instrument based Median randomization analysis. The x and y axis represented coefficients of genetic association with *LINC00339* (eQTL) from GTEx whole blood tissue[1] or genetic association with BMD from UK Biobank,[4] respectively. The red line and green dashed line corresponded to slope from IVW or weighted median, respectively. IVW, inverse-variance weighted.

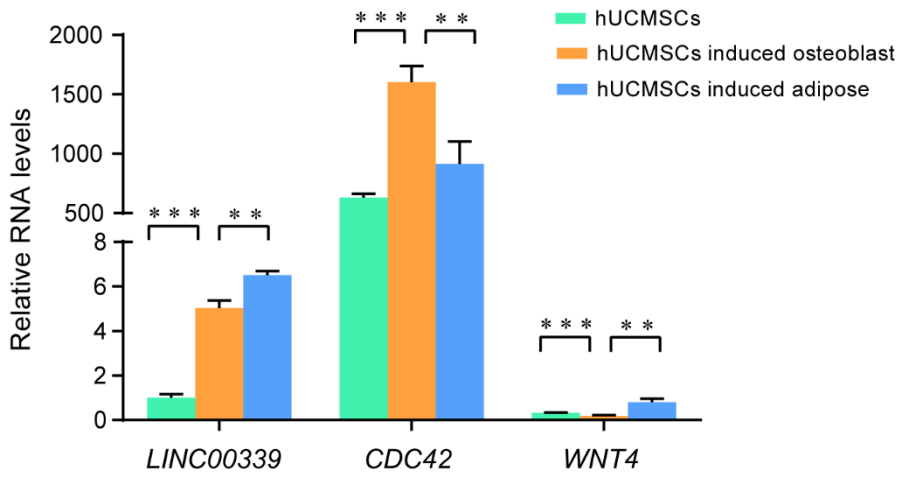

Figure S8. Comparison of gene expression in osteoblast and adipocyte cells induced from the same human mesenchymal stem cells

We obtained umbilical cords from donors with signed informed consent in local hospital and isolated human umbilical cord mesenchymal stem cells (hUCMSCs) as described previously.[5] The hUCMSCs cells were cultured with α-MEM supplemented with 10% FBS, 1% penicillin-streptomycin, and 0.2% cytokine CK2, CK4, CK8, CK9, and maintained at 37°C, 5% CO_2. The osteogenic and adipogenic differentiation was performed by using the OriCell™ hUCMSCs osteogenic differentiation medium kit (HUXUC-90021, Cyagen, China) and the OriCell™ hUCMSCs adipogenic differentiation medium kit (HUXUC-90031, Cyagen, China) according to the manufacturer's instruction, respectively. Cells were maintained in differentiation medium and the medium was changed every 3 days. Two weeks later, cells were harvested for RNA extraction and RT-qPCR. Relative mRNA expression levels were normalized by equalizing *LINC00339* expression levels in human umbilical cord mesenchymal stem cells (hUCMSCs) to 1. Error bars, s.d. $^{**}P < 0.01$, $^{***}P < 0.001$ as determined by an unpaired, two-tailed Student's t-test.

Supplemental Tables

Table S1. Basic characteristics of the Chinese cohort for genetic association analysis

Traits	Chinese cohort (n=1300)
Male/Female	600/700
Age (years)	33.42 (11.32)
Weight (kg)	59.63 (10.41)
Height (cm)	163.94 (8.11)
Lumbar spine BMD (g/cm^2)	0.921 (0.131)
Femoral neck BMD (g/cm^2)	0.815 (0.131)

Note: data are shown as mean (standard deviation, SD).

Table S2. Summary of Hi-C or ChIA-PET data used in this study

Dataset[a]	Data type	Cell[b]	Reference
4DGenome	Hi-C	IMR90	Jin, F. et al.[6]
Cell2014	Hi-C	GM12878	Rao, Suhas S.P. et al.[7]
ChIA-PET^{GEO}	ChIA-PET	GM12878	Tang, Z. et al. [8]
ChIA-PET^{ENCODE}	ChIA-PET	K562, NB4, HCT-116, HeLa-S3, MCF7	Harrow, J. et al.[9]
Cell2016	Capture Hi-C	17 human primary blood cell types	Javierre, B.M. et al.[10]
Mm2015	DNase Hi-C	H1-hESC	Ma, W. et al.[11]
NG2015	Capture Hi-C	GM12878; CD34	Mifsud, B. et al.[12]
TAD	--	IMR90	Dixon, J.R. et al.[13]

Note: [a]Dataset, Hi-C or ChIA-PET data used; [b]Cell, Hi-C data on human healthy cells or ChIA-PET data on all cells were collected; ChIA-PET^{GEO}, ChIA-PET data retrieved from GEO database (GEO: GSE72816); ChIA-PET^{ENCODE}, ChIA-PET data retrieved from UCSC ENCODE download portal.

Table S3. Integrating Hi-C and cis-expression quantitative trait locus (eQTL) analysis for 8 BMD SNPs at 1p36.12

SNP	Target gene predicted from Hi-C	Distance(kb)[a]	Dataset (cell)	Locus1	Locus2	P_{eQTL}	η^2
rs6426749	***LINC00339***	**-359.4**	**4DGenome; IMR90**	**chr1:22341459-22371546**	**chr1:22704394-22711600**	**5.61×10^{-5}**	**0.04**
rs6426749	***LINC00339***	**-359.4**	**Nm2015; H1-hESC**	**chr1:22351460-22356461**	**chr1:22704000-22711000**	**5.61×10^{-5}**	**0.04**
rs6426749	*WNT4*	-242	4DGenome; IMR90	chr1:22434529-22463352	chr1:22704394-22711600	0.45	3.50×10^{-3}
rs6426749	*WNT4*	-242	NG2015; GM12878	chr1:22466747-22478213	chr1:22705031-22711598	0.45	3.50×10^{-3}
rs6426749	*WNT4*	-242	4DGenome; IMR90	chr1:22466749-22528553	chr1:22704394-22711600	0.45	3.50×10^{-3}
rs6426749	*WNT4*	-242	4DGenome; IMR90	chr1:22466749-22478215	chr1:22703397-22711600	0.45	3.50×10^{-3}
rs6426749	*RP1-224A6.3*	-360	4DGenome; IMR90	chr1:22341459-22371546	chr1:22704394-22711600	0.43	6.25×10^{-3}
rs6426749	*ZBTB40*	67	4DGenome; IMR90	chr1:22769969-22795701	chr1:22704394-22711600	0.93	2.96×10^{-4}
rs6426749	*RP1-224A6.9*	-284.4	4DGenome; IMR90	chr1:22425025-22427905	chr1:22703397-22721228	NA	NA
rs34920465	*WNT4*	-229.9	NG2015; GM12878	chr1:22466747-22478213	chr1:22697090-22702342	0.53	2.79×10^{-3}
rs6696981	*WNT4*	-232.4	4DGenome; IMR90	chr1:22440795-22463352	chr1:22702345-22704393	0.63	2.03×10^{-3}
rs6696981	*WNT4*	-232.4	NG2015; CD34	chr1:22466747-22478213	chr1:22702343-22703394	0.63	2.03×10^{-3}
rs7524102	*WNT4*	-228.0	NG2015; GM12878	chr1:22466747-22478213	chr1:22697090-22702342	0.57	2.45×10^{-3}
rs2235529	*HSPG2*	-186.7	ChIA-PET^{ENCODE}; MCF7	chr1:22262651-22265288	chr1:22448523-22450702	0.92	3.63×10^{-4}
rs2235529	*WNT4*	20.0	ChIA-PET^{ENCODE}; K562	chr1:22450461-22453077	chr1:22468006-22470655	0.99	5.56×10^{-5}
rs2235529	*WNT4*	20.0	ChIA-PET^{ENCODE}; MCF7	chr1:22446379-22450951	chr1:22454408-22458465	0.99	5.56×10^{-5}
rs2235529	*WNT4*	20.0	ChIA-PET^{ENCODE}; K562	chr1:22448838-22451961	chr1:22466949-22469506	0.99	5.56×10^{-5}
rs2235529	*RP1-224A6.9*	-2.4	4DGenome; IMR90	chr1:22450003-22488812	chr1:22425025-22427905	NA	NA
rs3765350	*HSPG2*	-183.5	NG2015; GM12878	chr1:22263134-22276213	chr1:22440793-22450000	0.88	5.50×10^{-4}
rs3765350	*WNT4*	23.1	NG2015; GM12878	chr1:22466747-22478213	chr1:22440793-22450000	0.62	2.06×10^{-3}
rs3765350	*WNT4*	23.1	ChIA-PET^{ENCODE}; K562	chr1:22446637-22449453	chr1:22469771-22471584	0.62	2.06×10^{-3}
rs3765350	*WNT4*	23.1	ChIA-PET^{ENCODE}; MCF7	chr1:22446379-22450951	chr1:22454408-22458465	0.62	2.06×10^{-3}
rs3765350	*WNT4*	23.1	NG2015; CD34	chr1:22466747-22478213	chr1:22440793-22450000	0.62	2.06×10^{-3}

Note: Hi-C, Capture Hi-C, DNase Hi-C and ChIA-PET data on over 20 cells summarized in Table S1 were used, with chromatin interaction regions showed in Locus1 and Locus2 (hg19); NA, not available; [a]Distance is the distance between SNP and transcription start site of target gene.

Table S4. Cis-expression quantitative trait locus (eQTL) analysis results for rs6426749, rs6684375, and rs34963268

Gene	rs6426749			rs6684375			rs34963268		
	Distance	*P*-value	η^2	Distance	*P*-value	η^2	Distance	*P*-value	η^2
C1orf213	984.3	0.62	2.07×10^{-3}	989.3	0.72	1.44×10^{-3}	984.9	0.72	1.44×10^{-3}
HNRNPR	959.3	0.49	3.15×10^{-3}	964.4	0.73	1.35×10^{-3}	959.9	0.73	1.35×10^{-3}
RP5-1057J7.1	859.8	0.33	4.81×10^{-3}	864.8	0.26	5.87×10^{-3}	860.4	0.26	5.87×10^{-3}
LUZP1	724.2	0.52	2.84×10^{-3}	729.2	0.40	3.97×10^{-3}	724.7	0.40	3.97×10^{-3}
KDM1A	634.5	0.29	5.36×10^{-3}	639.5	0.47	3.33×10^{-3}	635.1	0.47	3.33×10^{-3}
EPHB2	326.0	0.53	2.72×10^{-3}	331.0	0.44	3.60×10^{-3}	326.6	0.44	3.60×10^{-3}
ZBTB40	67.0	0.93	2.96×10^{-4}	72.0	0.86	6.59×10^{-4}	67.6	0.86	6.59×10^{-4}
WNT4	-242.0	0.45	3.50×10^{-3}	-237.0	0.62	2.08×10^{-3}	-241.4	0.62	2.08×10^{-3}
CDC42	**-332.4**	**4.56×10^{-3}**	**0.023**	**-327.3**	**5.70×10^{-3}**	**0.022**	**-331.8**	**5.70×10^{-3}**	**0.022**
LINC00339	**-359.4**	**5.61×10^{-5}**	**0.042**	**-354.4**	**4.25×10^{-4}**	**0.033**	**-358.9**	**4.25×10^{-4}**	**0.033**
RP1-224A6.3	-360.0	0.43	6.25×10^{-3}	-355.0	0.27	5.75×10^{-3}	-359.4	0.27	5.75×10^{-3}
HSPG2	-447.7	0.26	5.84×10^{-3}	-442.6	0.37	4.26×10^{-3}	-447.1	0.37	4.26×10^{-3}
RP11-26H16.1	-476.9	0.24	3.64×10^{-3}	-471.8	0.38	4.15×10^{-3}	-476.3	0.38	4.15×10^{-3}
LDLRAD2	-572.7	0.40	4.02×10^{-3}	-567.7	0.59	2.26×10^{-3}	-572.1	0.59	2.26×10^{-3}
USP48	-658.4	0.91	4.15×10^{-4}	-653.4	0.69	1.64×10^{-3}	-657.8	0.69	1.64×10^{-3}
NBPF3	-944.9	0.27	5.65×10^{-3}	-939.8	0.43	3.63×10^{-3}	-944.3	0.43	3.63×10^{-3}
HS6ST1P1	-956.7	0.91	4.12×10^{-4}	-951.6	0.85	7.26×10^{-4}	-956.1	0.85	7.26×10^{-4}
NBPF2P	-957.0	0.81	9.39×10^{-4}	-952.0	0.75	1.26×10^{-3}	-956.4	0.75	1.26×10^{-3}
PPP1R11P1	-987.0	0.39	4.04×10^{-3}	-982.0	0.41	3.87×10^{-3}	-986.4	0.41	3.87×10^{-3}

Note: Distance is the distance between SNP and transcription start site of target gene (kb).

Table S5. Co-expression analysis between *LINC00339* and *CDC42*

Tissues	Samples[a]	*CDC42 expression*	*LINC00339* expression	*P*-value	R^{2b}
Thyroid	355	42.34(6.88)	12.22(3.55)	6.15E-15	-0.398
Vagina	97	53.61(9.50)	6.96(2.88)	2.66E-04	-0.362
Ovary	108	42.90(6.86)	10.08(2.81)	2.72E-04	-0.344
Colon transverse	204	53.93(9.69)	6.61(2.43)	2.14E-05	-0.293
Stomach	204	43.82(9.21)	6.71(2.45)	1.36E-03	-0.223
Spleen	118	55.72(8.46)	14.49(3.58)	2.23E-02	-0.21
Small intestine terminal ileum	104	53.27(7.22)	7.90(3.32)	3.78E-02	-0.204
Prostate	119	40.57(6.24)	9.60(3.36)	4.37E-02	-0.185
Liver	137	23.75(7.51)	5.44(2.19)	3.24E-02	-0.183
Colon sigmoid	173	45.29(6.50)	7.32(3.03)	2.40E-02	-0.172
Esophagus mucosa	331	55.43(7.16)	4.22(1.87)	6.68E-03	-0.149
Nerve tibial	335	46.76(6.78)	9.98(2.68)	4.91E-02	-0.108
Brain cortex	128	28.58(5.81)	4.39(1.32)	7.47E-03	0.235
Brain cerebellar hemisphere	115	40.26(9.50)	6.09(1.86)	2.65E-03	0.278
Brain spinal cord (cervical c-1)	76	45.03(13.42)	4.00(1.13)	6.25E-04	0.384
Brain nucleus accumbens (basal ganglia)	123	24.74(8.58)	3.81(1.39)	1.03E-07	0.458
Brain caudate (basal ganglia)	134	25.62(8.06)	3.77(1.32)	5.64E-09	0.477
Brain putamen (basal ganglia)	103	23.44(6.98)	3.71(1.50)	3.21E-08	0.512
Brain frontal cortex (BA9)	117	37.36(10.21)	4.43(1.68)	1.67E-09	0.521
Brain anterior cingulate cortex (BA24)	99	34.32(11.84)	3.65(1.38)	2.10E-10	0.585
Brain hippocampus	103	30.66(10.36)	3.77(1.17)	3.69E-12	0.618
Brain substantia nigra	71	34.63(11.66)	3.96(1.37)	2.75E-10	0.664
Brain amygdala	81	27.83(9.94)	3.26(1.25)	4.44E-16	0.754

Brain hypothalamus	104	42.04(14.33)	4.72(1.69)	$P < 2.20E-12$	0.759
Whole blood	445	83.71(37.45)	4.10(2.24)	$P < 2.20E-12$	0.586
Testis	199	29.88(8.00)	9.68(1.67)	7.93E-12	0.46
Adrenal gland	159	44.93(6.44)	7.55(2.23)	2.63E-02	0.176
Muscle skeletal	469	20.52(6.10)	3.91(2.47)	8.37E-10	0.279
Heart left ventricle	267	25.92(9.09)	3.98(2.17)	3.77E-05	0.249

Note: Co-expression analysis was conducted by Pearson correlation using GTEx RNA expression data[1] in 50 tissues (4 tissues with sample counts less than 20 were excluded). Only significantly correlated tissues ($P < 0.05$) were showed. Expression data was shown as mean (standard deviation, SD); [a]Samples were sample counts without missing *CDC42* or *LINC00339* expression data; [b]R^2 was Pearson Correlation Coefficient.

Table S6. Chromatin interactions between *LINC00339* and *CDC42*

Cell	Validation	Locus1	Gene1	Locus2	Gene2	Score
K562	ChIA-PET^{ENCODE}	chr1:22348117-22354021	*LINC00339*	chr1:22376823-22382698	*CDC42*	15
K562	ChIA-PET^{ENCODE}	chr1:22348152-22354988	*LINC00339*	chr1:22377027-22382698	*CDC42*	18
MCF7	ChIA-PET^{ENCODE}	chr1:22350975-22355164	*LINC00339*	chr1:22378266-22380971	*CDC42*	7
MCF7	ChIA-PET^{ENCODE}	chr1:22351119-22355075	*LINC00339*	chr1:22377954-22381978	*CDC42*	5
HeLa-S3	ChIA-PET^{ENCODE}	chr1:22351928-22352455	*LINC00339*	chr1:22379030-22380021	*CDC42*	3
K562	ChIA-PET^{ENCODE}	chr1:22354998-22357992	*LINC00339*	chr1:22377814-22380762	*CDC42*	2
GM12878	ChIA-PET^{GEO}	chr1:22349358-22349912	*LINC00339*	chr1:22379575-22381183	*CDC42*	5
GM12878	ChIA-PET^{GEO}	chr1:22350646-22354253	*LINC00339*	chr1:22377722-22381071	*CDC42*	61
H1-hESC	Nm2015	chr1:22351460-22356461	*LINC00339*	chr1:22374000-22388000	*CDC42*	NA
IMR90	4DGenome	chr1:22341459-22375876	*LINC00339*	chr1:22375877-22377917	*CDC42*	4.17E-06
IMR90	4DGenome	chr1:22351108-22359290	*LINC00339*	chr1:22379987-22393227	*CDC42*	5.70E-04
IMR90	4DGenome	chr1:22351108-22376206	*LINC00339*	chr1:22377918-22379986	*CDC42*	3.52E-12
IMR90	4DGenome	chr1:22359539-22409777	*LINC00339*	chr1:22351108-22359290	*CDC42*	3.47E-14

Note: Hi-C, DNase Hi-C and ChIA-PET data summarized in Table S1 were used, with chromatin interaction regions showed in Locus1 and Locus2 (hg19); NA, not available; [a]Score: Confidence *P*-value for Hi-C or confidence scores for ChIA-PET chromatin interactions.

Table S7. Summary of primers or siRNA sequences used

Assays	Target	Primers (5'-3')
Luciferase Report-Fusion PCR	rs6426749-F1	GGGGTACCTTTTAGGGAGTTTGAATTGGGCTC (Kpn I)
	rs6426749-R1	AGGCCAGAGGACTATTGTATTTGA
	LINC00339-F1	AATAGTCCTCTGGCCTTGGTTAGCATCTCTGCTTCCTCTA
	LINC00339-R1	CGACGCGTGGACGAGGAAAGATCAGGATAAGA (Mlu I)
Luciferase Report PCR	rs34963268-F1	GGGGTACCAGGCATCTGATAAAGACTCCG (Kpn I)
	rs34963268-R1	CGACGCGTTAAAAGGCCCCAGTAACCC (Mlu I)
	rs6684375-F1	GGGGTACCCCTCATGCCAATGACTCTGGT (Kpn I)
	rs6684375-R1	CGACGCGTATAGCCTGTCCTCATCCTTCCG (Mlu I)
	LINC00339-F2	CGACGCGTTGGTTAGCATCTCTGCTTCCTCTA (Mlu I)
	LINC00339-R2	GAAGATCTGGACGAGGAAAGATCAGGATAAGA (Bgl II)
Luciferase Report-Promoter PCR	*LINC00339*-F	GGGGTACCTGGTTAGCATCTCTGCTTCCTCTA (Kpn I)
	LINC00339-R	CGACGCGTGGACGAGGAAAGATCAGGATAAGA (Mlu I)
Site-directed mutagenesis	rs6426749-F (G-C)	CATACTGGCTGCTGAGCTCCAGGCCAATGGAC
	rs6426749-R (G-C)	GTCCATTGGCCTGGAGCTCAGCAGCCAGTATG
	rs34963268-F (C-G)	CTGGATCGTTGACGTCATTTGAGTGCCTGGAT
	rs34963268-R (C-G)	TGACGTCAACGATCCAGGCACTCAAATGACGT
	rs6684375-F (C-T)	TGGGAATCTGCTCCTCTTCTCTTTTGGGTTGG
	rs6684375-R (C-T)	AGAGGAGCAGATTCCCAGGGGCCCTCCGGCTAAGC
siRNAs (sense)	*CTCF*	UCACCCUCCUGAGGAAUCACCUUAA
	TFAP2A (siRNA-1)	CCGUCUCCGCCAUCCCUAUUAACAA
	TFAP2A (siRNA-2)	AACAUCCCAGAUCAAACUGUA
CRISPR/Cas9	sgRNA 1-F	ACCGTCCTTTCTTCTTTGGACAC
	sgRNA 1-R	AAACGTGTCCAAAGAAGAAAGGA
	sgRNA 2-F	ACCGGCCGCACATTGACATCACC

	sgRNA 2-R	AAACGGTGATGTCAATGTGCGGC
dCas9-KRAB	sgRNA-1-F	ACCGGGGAGCCCTTCCATTCTCG
	sgRNA-1-R	AAACCGAGAATGGAAGGGCTCCC
	sgRNA-2-F	ACCGGCTGATATTAGCAGTGTAC
	sgRNA-2-R	AAACGTACACTGCTAATATCAGC
	sgRNA-3-F	ACCGGCCAATGGGGCATGAGTTG
	sgRNA-3-R	AAACCAACTCATGCCCCATTGGC
RT-qPCR	*LINC00339*-F	GTCCAGATTCCACGAGAGCCTT
	LINC00339-R	GTCTCAGCCACCACCGTCCA
	CDC42-F	GATGGTGCTGTTGGTAAA
	CDC42-R	TAACTCAGCGGTCGTAAT
	CTCF-F	GTGTTCCATGTGCGATTACG
	CTCF-R	TCATGTGCCTTTTCAGCTTG
	TFAP2A-F	GTTCACGCCGATCCATGAAAA
	TFAP2A-R	AGATTGACCTACAGTGCCCAG
ChIP-qPCR	rs6426749-F	ATGTGAAATGCTTACACTGGAGTTC
	rs6426749-R	ATGTGAAATGCTTACACTGGAGTTC

Note: F, forward primer; R, reverse primer; Restriction enzyme site sequences were underlined; For rs6426749, we used fusion PCR[14] to effectively get the long fragment containing both enhancer and *LINC00339* promoter, which was further inserted into the pGL3-basic vector. For rs34963268 and rs6684375, we appended the same restriction enzyme sites to both enhancer and *LINC00339* promoter, which were further inserted into the pGL3-basic vector sequentially.

Table S8. Genetic association with *LINC00339* and BMD for 44 selected SNPs used for multi-instrument based Mendelian randomization analysis

SNP	Chr	Position	eQTL[a]		GWAS[b]	
			P	Beta	*P*	Beta
rs471359	1	21656500	0.010	-0.248	0.006	0.012
rs78885464	1	21807864	0.001	-0.240	0.440	0.003
rs61778393	1	21902436	0.006	0.512	0.003	0.030
rs1130564	1	21952884	0.005	-0.366	0.380	0.009
rs12128206	1	21980091	4.231×10^{-4}	0.631	0.036	0.020
rs60765766	1	22017013	0.010	0.258	0.043	-0.012
rs2315928	1	22189447	0.009	-0.527	0.680	-0.027
rs114537356	1	22214279	0.003	0.739	0.240	-0.014
rs114568494	1	22241660	0.004	-0.650	0.009	0.032
rs6684979	1	22261395	0.008	-0.406	0.240	-0.011
rs35601247	1	22272915	0.004	0.540	0.350	-0.015
rs145444626	1	22287577	2.310×10^{-4}	0.913	0.680	0.000
rs6661287	1	22298481	0.005	0.265	0.740	-0.001
rs12059804	1	22304585	0.002	0.185	0.500	0.003
rs61777960	1	22311348	0.002	-0.184	3.100×10^{-4}	-0.013
rs10917101	1	22314475	1.416×10^{-4}	-0.262	0.036	0.010
rs2865210	1	22342050	0.002	-0.203	0.960	-0.001
rs2255282	1	22352040	1.012×10^{-20}	-0.473	0.011	-0.008
rs116674939	1	22354237	2.753×10^{-7}	0.853	0.044	0.023
rs150153349	1	22355890	0.005	-0.449	0.330	-0.012
rs2473277	1	22361845	1.618×10^{-18}	-0.447	0.012	-0.008
rs2473317	1	22395251	8.522×10^{-4}	-0.252	0.490	0.004
rs16826588	1	22424113	2.482×10^{-4}	0.445	0.074	0.020
rs1046310	1	22443887	2.005×10^{-10}	-0.317	8.900×10^{-4}	-0.010
rs10917161	1	22460208	0.006	-0.454	0.004	0.025
rs113155445	1	22472435	9.761×10^{-4}	-0.262	0.330	0.005
rs4655026	1	22473658	2.811×10^{-8}	-0.288	2.100×10^{-4}	-0.011
rs735475	1	22482230	0.004	0.470	0.680	0.006
rs2807352	1	22495261	3.293×10^{-5}	0.223	6.700×10^{-5}	0.013
rs2982286	1	22506729	0.003	-0.151	5.300×10^{-15}	-0.027
rs140767127	1	22512667	0.008	-0.628	0.170	-0.026
rs115963111	1	22534928	0.001	0.636	0.830	0.002
rs2807331	1	22565967	1.798×10^{-4}	0.198	2.900×10^{-5}	0.014
rs75868741	1	22594676	0.008	0.237	3.300×10^{-8}	-0.027
rs1007243	1	22614839	0.006	0.166	0.001	0.013
rs74816778	1	22641134	0.008	0.451	0.009	-0.025
rs11585537	1	22656868	0.004	-0.222	0.250	-0.005
rs61769163	1	22678805	7.692×10^{-4}	0.388	0.057	0.013
rs4654807	1	22949552	0.009	-0.216	0.730	0.001
rs7549888	1	23004019	0.004	0.153	0.250	-0.003

rs11811882	1	23019404	0.006	0.158	0.570	-0.002
rs75858988	1	23035195	0.006	-0.488	0.011	-0.033
rs111727123	1	23063461	0.007	-0.576	0.300	-0.023
rs76603191	1	23133152	0.003	0.502	0.055	0.012

Note: eQTL[a]: Genetic association with *LINC00339* expression extracted from GTEx whole blood tissue;[1] GWAS[b]: Genetic association with BMD collected from UK Biobank.[4]

Supplemental References

1. Lonsdale, J., Thomas, J., Salvatore, M., Phillips, R., Lo, E., Shad, S., Hasz, R., Walters, G., Garcia, F., Young, N., et al. (2013). The Genotype-Tissue Expression (GTEx) project. Nat Genet 45, 580-585.
2. Thompson, B., Varticovski, L., Baek, S., and Hager, G.L. (2016). Genome-Wide Chromatin Landscape Transitions Identify Novel Pathways in Early Commitment to Osteoblast Differentiation. PLOS ONE 11, e0148619.
3. Hon, C.C., Ramilowski, J.A., Harshbarger, J., Bertin, N., Rackham, O.J., Gough, J., Denisenko, E., Schmeier, S., Poulsen, T.M., Severin, J., et al. (2017). An atlas of human long non-coding RNAs with accurate 5' ends. Nature 543, 199-204.
4. Kemp, J.P., Morris, J.A., Medina-Gomez, C., Forgetta, V., Warrington, N.M., Youlten, S.E., Zheng, J., Gregson, C.L., Grundberg, E., Trajanoska, K., et al. (2017). Identification of 153 new loci associated with heel bone mineral density and functional involvement of GPC6 in osteoporosis. Nat Genet.
5. Lu, L.L., Liu, Y.J., Yang, S.G., Zhao, Q.J., Wang, X., Gong, W., Han, Z.B., Xu, Z.S., Lu, Y.X., Liu, D., et al. (2006). Isolation and characterization of human umbilical cord mesenchymal stem cells with hematopoiesis-supportive function and other potentials. Haematologica 91, 1017-1026.
6. Jin, F., Li, Y., Dixon, J.R., Selvaraj, S., Ye, Z., Lee, A.Y., Yen, C.-A., Schmitt, A.D., Espinoza, C.A., and Ren, B. (2013). A high-resolution map of the three-dimensional chromatin interactome in human cells. Nature 503, 290-294.
7. Rao, Suhas S.P., Huntley, Miriam H., Durand, Neva C., Stamenova, Elena K., Bochkov, Ivan D., Robinson, James T., Sanborn, Adrian L., Machol, I., Omer, Arina D., Lander, Eric S., et al. (2014). A 3D Map of the Human Genome at Kilobase Resolution Reveals Principles of Chromatin Looping. Cell 159, 1665-1680.
8. Tang, Z., Luo, O.J., Li, X., Zheng, M., Zhu, J.J., Szalaj, P., Trzaskoma, P., Magalska, A., Wlodarczyk, J., Ruszczycki, B., et al. (2015). CTCF-Mediated Human 3D Genome Architecture Reveals Chromatin Topology for Transcription. Cell 163, 1611-1627.
9. Harrow, J., Frankish, A., Gonzalez, J.M., Tapanari, E., Diekhans, M., Kokocinski, F., Aken, B.L., Barrell, D., Zadissa, A., Searle, S., et al. (2012). GENCODE: the reference human genome annotation for The ENCODE Project. Genome Res 22, 1760-1774.
10. Javierre, B.M., Burren, O.S., Wilder, S.P., Kreuzhuber, R., Hill, S.M., Sewitz, S., Cairns, J., Wingett, S.W., Várnai, C., Thiecke, M.J., et al. (2016). Lineage-Specific Genome Architecture Links Enhancers and Non-coding Disease Variants to Target Gene Promoters. Cell 167, 1369-1384.e1319.
11. Ma, W., Ay, F., Lee, C., Gulsoy, G., Deng, X., Cook, S., Hesson, J., Cavanaugh, C., Ware, C.B., Krumm, A., et al. (2015). Fine-scale chromatin interaction maps reveal the cis-regulatory landscape of human lincRNA genes. Nat Meth 12, 71-78.
12. Mifsud, B., Tavares-Cadete, F., Young, A.N., Sugar, R., Schoenfelder, S., Ferreira, L., Wingett, S.W., Andrews, S., Grey, W., Ewels, P.A., et al. (2015). Mapping long-range promoter contacts in human cells with high-resolution capture Hi-C. Nat Genet 47, 598-606.
13. Dixon, J.R., Selvaraj, S., Yue, F., Kim, A., Li, Y., Shen, Y., Hu, M., Liu, J.S., and Ren, B. (2012). Topological domains in mammalian genomes identified by analysis of chromatin interactions. Nature 485, 376-380.
14. Horton, R.M., Hunt, H.D., Ho, S.N., Pullen, J.K., and Pease, L.R. (1989). Engineering hybrid genes

without the use of restriction enzymes: gene splicing by overlap extension. Gene 77, 61-68.