Pre-B-cell leukemia homeobox 1 (*PBX1*) shows functional and possible genetic association with bone mineral density variation

Ching-Lung Cheung[1], Benjamin Y.Y. Chan[1], Vivian Chan[1], Shiro Ikegawa[5], Ikuyo Kou[5], Heidi Ngai[1], David Smith[2], Keith D.K. Luk[3], Qing-Yang Huang[1], Seijiro Mori[6], Pak-Chung Sham[4] and Annie W.C. Kung[1,*]

[1]Department of Medicine, [2]Department of Biochemistry and [3]Department of Orthopaedics and Traumatology, [4]Genome Research Centre, The University of Hong Kong, Pokfulam, Hong Kong, Republic of China, [5]Laboratory for Bone and Joint Disease, Center for Genomic Medicine, RIKEN, Tokyo, Japan and [6]Department of Internal Medicine, Tokyo Metropolitan Geriatric Hospital, Tokyo, Japan

Bone mineral density (BMD) is one of the major determinants of risk for osteoporotic fracture. Multiple studies reveal that peak bone mass is under strong genetic influence. One of the major susceptibility loci for peak spine BMD has been mapped to chromosome 1q21–q23 in the Caucasian population. We have previously replicated this finding in Southern Chinese pedigrees and detected a maximum multipoint log of odds (LOD) score of 2.36 in this region. To further fine-map this region, 380 single-nucleotide polymorphic (SNP) markers were genotyped in 610 sibpairs from 231 families. Several markers were identified in the association analysis as important candidates underlying BMD variation. Among them, successful replication was demonstrated for SNPs in pre-B-cell leukemia homeobox 1 (*PBX1*) gene in two other unrelated case–control cohorts. The functional role of *PBX1* in bone metabolism was examined *in vitro* using human bone-derived cells (HBDC) and murine MC3T3-E1 pre-osteoblasts. *PBX1* mRNA was constitutively expressed in both HBDC and MC3T3-E1 cells. Immunostaining revealed that *PBX1* is localized in the nucleus compartment. Silencing of *PBX1* by RNAi in MC3T3-E1 cells decreased the expression of *Runx2* and *Osterix*, the critical transcription factors for osteogenesis, but accelerated cell proliferation and bone nodule formation. Overall, our data suggest a genetic and functional association of *PBX1* with BMD.

INTRODUCTION

Osteoporosis (MIM 166710) is a major public health problem that is associated with excess disability and mortality and a huge socioeconomic burden due to its complication of bone fractures. It is characterized by low bone mass and microarchitectural deterioration of bone tissue that lead to increased bone fragility and a consequent increased fracture risk (1). As the world's population ages, the prevalence rate of osteoporosis and osteoporotic fractures increase in parallel with approximately 200 million women worldwide suffer from osteoporosis. Vertebral compression fractures lead to pain, deformity and disability whereas hip fractures associated

with falls result in loss of mobility and, in some cases, death. Bone strength is predominantly determined by bone mineral density (BMD) and bone quality (2). BMD predicts fracture risk independently, and clinically it is the best characterized bone parameter that can be precisely measured (1).

Numerous studies have demonstrated that BMD has a strong genetic component (3–6). Up to 85% of the variability of BMD in young Caucasian adults is attributable to genetic factors. Similar studies conducted in Chinese pedigrees also reveal a heritability estimate of 0.73 for spine BMD in women, a figure that remains significant after adjustment for anthropometric measurements and lifestyle factors (7).

*To whom correspondence should be addressed at: Department of Medicine, The University of Hong Kong Pokfulam, Hong Kong, Republic of China. Tel: +86 85228554769; Fax: +86 85228162187; Email: awckung@hkucc.hku.hk

Despite numerous attempts made over previous decades to identify gene(s) involved in osteoporosis, the complex genetic architecture of osteoporosis remains ambiguous, and the genes responsible for BMD variation are largely unknown. The recent development of high throughput technology that allows genotyping of hundreds of thousands of SNPs across the genome has enabled unbiased genome-wide association (GWA) studies to be performed.

Over the past few years, a number of genome-wide linkage scans have searched for genes that underlie BMD variation and several susceptibility loci have been identified (8–11). Among these, chromosome 1q shows the strongest linkage signal with spine BMD; the highest LOD score of 4.3 is observed in the region of 175.62–182.35 cM (11,12). In genome-wide linkage scans in mice, a LOD score as high as 14.02 was observed in its homologous region of human chromosome 1q21–31 (13). This observation has, nonetheless, not been confirmed by other genome-wide linkage scans. We recently performed a collaborative meta-analysis of nine genome-wide scans for BMD that involved 11 842 subjects and detected the strongest evidence of linkage at chromosome 1p13.3–q23.3 ($P = 0.002$) (14). Similarly, our linkage study in 306 Southern Chinese pedigrees revealed a linkage peak region of ~10 cM between 165.62 and 191.52 cM with spine BMD, with the maximum multipoint LOD score of 2.36 (15).

In this present study, fine-mapping of the quantitative trait loci (QTL) in this 6 Mb region was performed using 380 SNPs. Spine BMD association was localized to the *PBX1* gene in two Chinese cohorts and one Japanese cohort. In addition, *PBX1* was identified in human bone-derived cells (HBDC) and mouse MC3T3-E1 osteoprogenitor cells. Silencing of *PBX1* by RNAi decreased the expression of *Runx2* and *Osterix*, the two key factors that control osteoblast differentiation, but accelerated cell proliferation and bone nodule formation. These results suggest that *PBX1* plays a crucial role in osteogenesis and matrix mineralization, and that rs2800791 in this gene is associated with BMD variation.

RESULTS

Initial genetic screening

We genotyped 380 tagging SNPs across a 6 Mb region on chromosome 1q. Among these 380 SNPs, the quality control criteria of a genotyping call rate >90%, MAF >0.05, duplicate error rate <0.02 and Hardy–Weinberg equilibrium (HWE) $P > 0.01$ were achieved for 262 SNPs (Supplementary Material, Table S3). The linkage disequilibrium (LD) pattern in terms of r^2 and D' of these 262 SNPs is summarized in Supplementary Material, Figure S2. Using a BMD spine z-score as the phenotype, eight SNPs showed a significant total association at the empirical P-value of 0.05 through 10 000 permutations. Only two SNPs, i.e. rs2800791 and rs9661977 in the *PBX1* gene, showed consistent associations in the within-association tests (Table 1). All association results (including insignificant results) are provided in Supplementary Material, Table S4. Imputation of untyped SNP genotypes using the Markov Chain framework implemented in Markov Chain Haplotyping software (MaCH) (16) did not improve the association signal (Data not shown).

Table 1. The significant results of single locus test for spine BMD in 610 sib-pairs

SNP	Gene	Physical position	Within association		Total association	
			Beta	Emp. P*	Beta	Emp. P*
rs2800791	*PBX1*	162830048	0.210	0.038	0.270	0.004
rs9661977	*PBX1*	163016317	0.290	0.032	0.240	0.050

*Significance was defined by an empirical P-value of <0.05 as derived from 10 000 permutations.

Replication in two large case–control cohorts

Replication in the Chinese case–control population was achieved with rs2800791. Logistic regression analysis revealed a P-value of 0.007 for rs2800791 and 0.4 for rs9661977 (Table 2). Replication in the Japanese cohort showed a P-value of 0.05 for rs2800791 and 0.344 for rs9661977. The meta-analyses showed a pooled odds ratio (OR) of 0.83 (95% CI: 0.73–0.94; $P = 0.003$) and 0.94 (95% CI: 0.77–1.15; $P = 0.285$) for rs2800791 and rs9661977, respectively.

Resequencing of *PBX1* in 20 high- and 20 low-spine BMD subjects

To determine whether rs2800791 could be in LD with another functional SNP, we resequenced genomic DNA for all exons, 5′-UTR and 3′-UTR from 20 high-spine BMD subjects who carried the protective allele at rs2800791 and 20 low-spine BMD subjects who carried the at risk allele at rs2800791. This resequencing effort identified two novel polymorphisms in the 5′-UTR −776G/C and −668C/A, and one known non-synonymous SNP rs2275558 in exon 2. Genotyping these three SNPs in 835 southern Chinese case–control subjects revealed an extremely low-allele frequency of the 5′-UTR SNPs (MAF < 0.01) and no association between rs2275558 and spine BMD variation ($P = 0.93$; OR = 1.02; 95% CI, 0.69–1.51).

Functionality of *PBX1* in osteoblasts

Semi-quantitative PCR revealed that *PBX1* mRNA expression was consistently present in both HBDCs and MC3T3 mouse pre-osteoblast cells during proliferation and differentiation phases (Fig. 1A and B). *PBX1* protein was localized in the nuclei compartment but not in the cytoplasm (Fig. 1C and D). The expression of *PBX1* mRNA in MC3T3 cells was suppressed 3.7-fold by 100 ng/ml of BMP2 after 4 days of culture (Fig. 1E), while estradiol, PTH and dexamethasone had no effect on *PBX1* gene expression. As rs2800791 lies in intron 2 and that intronic sequence variation might affect mRNA splicing, we studied the mRNA species from HBDC of subjects with different rs2800791 genotypes but did not observe any variation in mRNA species.

Transient silencing of *PBX1* gene with RNAi in MC3T3-E1 pre-osteoblast cells for 4 days was confirmed by real-time PCR and western blot analysis (Supplementary Material, Fig. S3A and B). Transient knock-down of *PBX1* increased the cell proliferation rate by 30% compared with control

Table 2. Replication of association of two SNPs from *PBX1* gene locus with BMD in two independent Southern Chinese and Japanese populations

SNP	Genotype	Southern Chinese Study (n = 835)					Japanese Study (n = 1268)					Meta-analysis (n = 2103)				
		Low BMD (%)	High BMD (%)	OR	95% CI	P-value	Low BMD (%)	High BMD (%)	OR	95% CI	P-value	Low BMD (%)	High BMD (%)	OR	95% CI	P-value
rs2800791	CC	29.0	21.1	0.76	0.61–0.94	0.007	29.6	27.1	0.87	0.74–1.00	0.050	29.3	25.3	0.83	0.73–0.94	0.003
	CT	51.7	54.5				52.4	51.0				52.0	52.0			
	TT	19.3	24.4				18.0	21.9				18.6	22.7			
rs9661977	GG	73.9	75.9	1.04	0.76–1.42	0.4	80.9	79.2	0.88	0.68–1.14	0.344	77.7	78.1	0.94	0.77–1.15	0.285
	GA	24.7	21.6				18.4	19.6				21.1	20.2			
	AA	1.4	2.4				0.7	1.2				1.0	1.6			

Results are genotype frequency. SNP positions according to NCBI dbSNP build 128.
P-value was one-sided since the direction of association was the same as the initial QFAM analysis.
OR, odds ratio; CI: confidence interval.

($P < 0.01$, Fig. 2A). Nonetheless, the gene expression of *Runx2* and *Osterix*, the two master factors that control osteoblast differentiation, was downregulated at Day 4 post-transfection by 3- and 2-fold, respectively (Fig. 2B). Interestingly, matrix mineralization was advanced and mineralized bone nodules were seen at Day 20 after transient knock-down of *PBX1* compared with Day 25 in controls. By Day 30, enhanced mineralization in the *PBX1* knock-down cells was observed (scramble control: $100 \pm 3\%$; RNAi $153 \pm 15\%$, $P < 0.001$, Fig. 2C). These results provide evidence that *PBX1* is involved at different stages of osteoblast proliferation and differentiation.

DISCUSSION

This association and replication study identified *PBX1* as a possible osteoporosis susceptibility gene for spine BMD. According to the HapMap data, Rel 21a/phase II Jan 07, NCBI B36 assembly, rs2800791 of *PBX1* is located in an LD block flanked by rs12739548 and rs12405060 in intron 2. This LD block is present in both Chinese and Japanese populations (Supplementary Material, Fig. S4). It is postulated that either rs2800791 itself or the SNP in high LD is the functional variant accounting for the association with BMD variation. We sequenced the exons and immediate intron–exon boundaries in both homozygous CC and TT carriers of rs2800791 and did not identify any novel sequence variants in homozygous carriers of the variant T alleles with high BMD. Since rs2800791 is located in intron 2, we determined whether the sequence variation might affect mRNA splicing, but we failed to detect variation in mRNA species in subjects with different rs2800791 genotypes. Next, we used TFSearch (TFSEARCH: Searching Transcription Factor Binding Sites, http://mbs.cbrc.jp/research/db/TFSEARCH.html) set at default settings to search for potential alteration in transcription factor binding site. Interestingly, the *c-myc*-binding site was present in the sequence with T allele but not C allele of rs2800791 with a score of 85.6. *C-myc* is a proto-oncogene that functions as a transcription factor. Activation is induced by estrogen through the recruitment of estrogen receptors, such as estrogen receptor alpha (*ESR1*), and specific coactivators to their promoters (17). As a recent study showed that *c-myc* is an effector of estrogen/estrogen receptor signaling, we, therefore, further investigated the epistasis between *ESR1*, estrogen receptor beta (*ESR2*) and *PBX1* using multifactor dimensionality reduction (MDR) (18). The MDR revealed significant epistasis between the promoter SNP T-1213C of *ESR2* and rs2800791 in *PBX1* with a testing balanced accuracy of 0.553 (Supplementary Material, Table S5). The putative *PBX1*-binding site was also investigated in a 2 kb sequence upstream of the first exon of *ESR1* and *ESR2* using JASPAR database (19). Three predicted putative *PBX1*-binding sites were present within a 2 kb region of *ESR1*, but not *ESR2* (Supplementary Material, Table S6). Estrogen and estrogen receptors are well-known mediators in bone metabolism (20), and the relationship underlying estrogen and *PBX1* and probably *c-myc* requires further elucidation. To investigate the possible functional role of rs2800791, we retrieved the expression data from the study
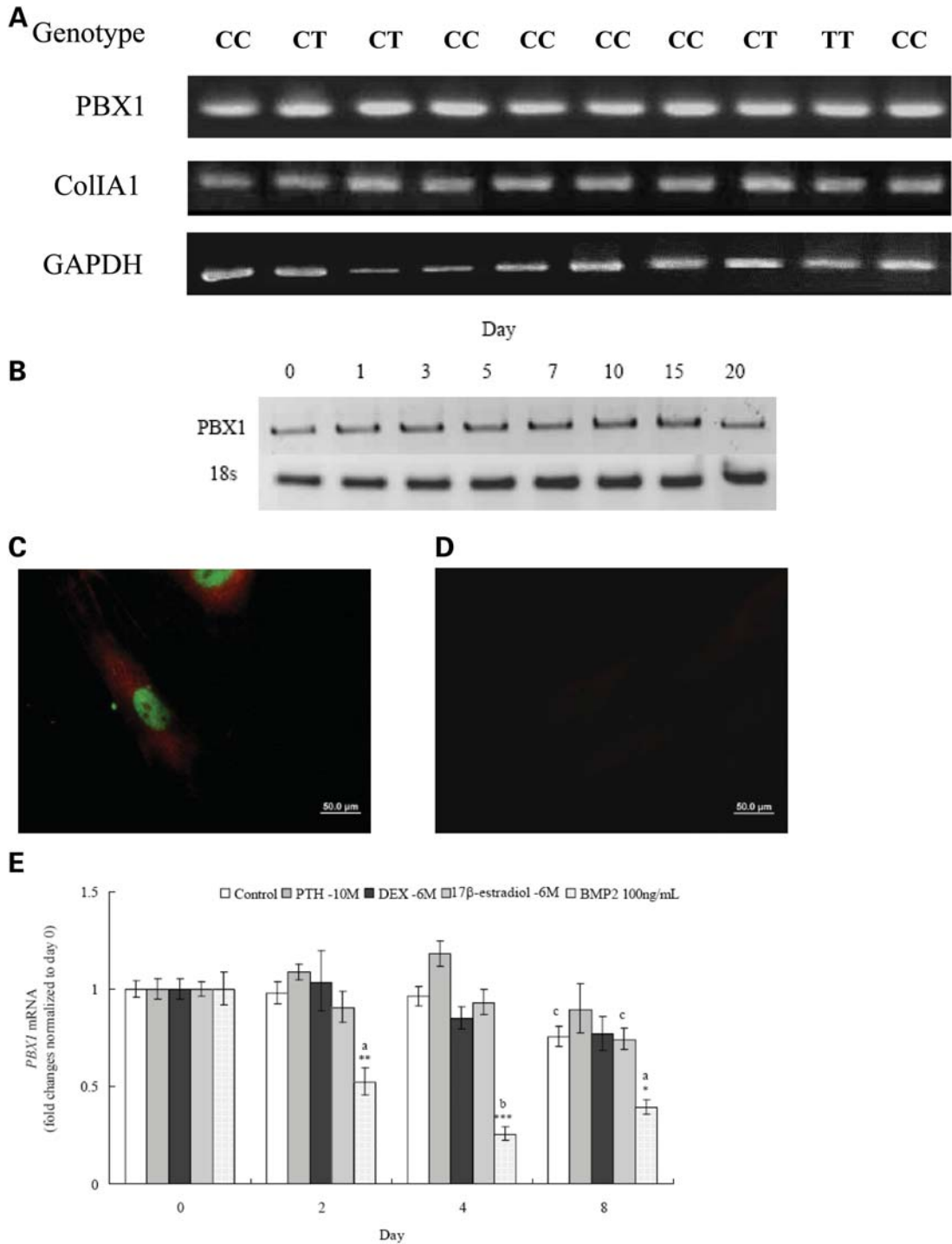
Figure 1. *PBX1* mRNA was consistently expressed in human bone-derived cells (HBDC) (**A**) and murine MC3T3-E1 pre-osteoblast cells (**B**). *PBX1*, pre-B-cell leukemia homeobox 1; col1a1, collagen type 1α1; GAPDH, glyceraldehyde 3-phosphate dehydrogenase. Immunofluorescence staining HBDCs showing *PBX1* (green) in the nucleus and col1aI (red) in the cytoplasm (**C**), with mice sera as negative control (**D**). Effect of PTH (10^{-10} M), Dexamethasone (10^{-6} M), 17β-estradiol (10^{-6} M) and BMP2 (100 ng/ml) on *PBX1* mRNA expression (**E**). Results are expressed as mean ± SD. [a]$P < 0.05$; [b]$P < 0.01$; [c]$P < 0.001$ (versus Day 0); *$P < 0.05$; **$P < 0.01$; ***$P < 0.001$ (versus control).

of Stranger *et al.* (21). Using HapMap CHB+JPT lymphoblastoid cell lines to perform the genotype-expression correlation using linear regression method with adjustment of ethnicity, the genotype rs2800791 (denoted as additive fashion) was not correlated with PBX1 expression. We believe that this nonsignificant correlation does not necessarily negate the possible association of rs2800791 with PBX1 expression as the number of subjects in the HapMap project is small. Thus, future experiment and replication are needed to validate our findings.
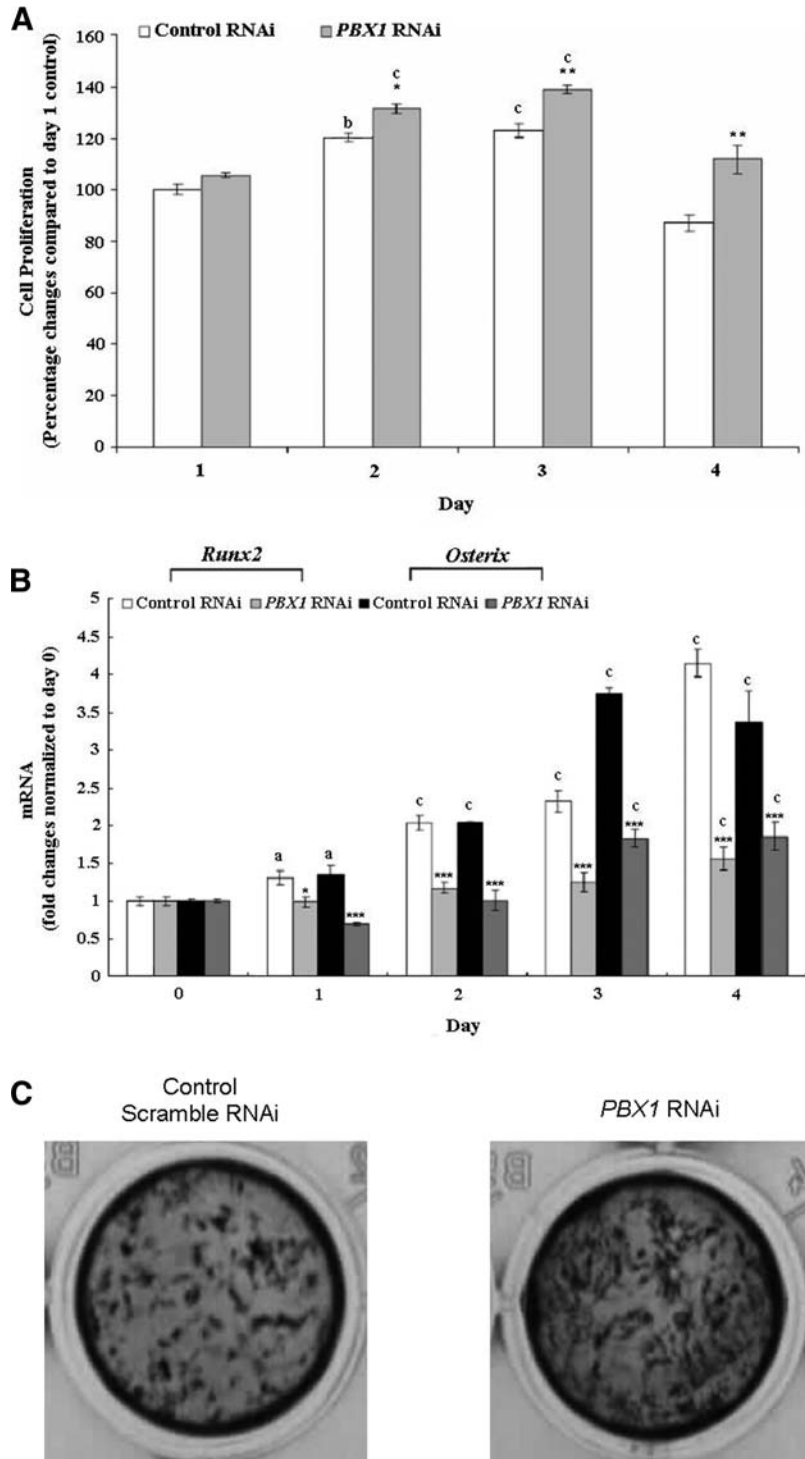
Figure 2. Transient knock-down of *PBX1* significantly enhanced cell proliferation of murine MC3T3-E1 pre-osteoblasts compared with scramble RNAi control (**A**). The two master genes controlling osteoblast differentiation, *Runx2* and *Osterix*, were significantly downregulated by *PBX1* knock-down (**B**). Enhanced matrix mineralization was observed at Day 30 by Von-Kossa silver stain (**C**). Results are expressed as mean ± SD. [a]$P < 0.05$; [b]$P < 0.01$; [c]$P < 0.001$ (versus Day 0); *$P < 0.05$; **$P < 0.01$; ***$P < 0.001$ (versus control).

Two 300K GWA studies were recently published (22,23). In one of these studies of over 10 000 Caucasian subjects using Illumina humanHap 300 chips, several BMD candidate genes, *RANKL, OPG* and *ESR1*, and two novel findings near the major histocompatibility complex and zinc finger and BTB domain containing 40 gene showed association with BMD variation at genome-wide significant level (22). In this GWA study, 52 SNPs were genotyped in *PBX1* gene locus

and 14 showed *P*-value of <0.05 (27%), although rs9661977 and rs2800791 were polymorphic in Caucasian subjects with MAF of 0.273 and 0.4, respectively, neither of them were included in this GWA study. None of the 52 SNPs were in high LD with rs9661977 ($r^2 > 0.8$) based on the LD pattern in Chinese and Japanese populations. Interestingly, the SNP rs2800783, which was in high LD ($r^2 = 0.846$) but different haplotype block with rs2800791, showed a trend toward significant association with spine BMD ($P = 0.062$). The G allele of rs2800783 is associated with a lower risk of low BMD (U. Styrkarsdottir, personal communication), and this is in the same direction of association as the T allele of rs2800791 observed in our study. This further suggest the functional importance of rs2800791 or SNPs in high LD.

We showed in this study that *PBX1* is constitutively expressed in both HBDC and mouse pre-osteoblast cells. *PBX1* gene expression is downregulated by BMP2 during the proliferative phase of the pre-osteoblasts in culture. Transient gene silencing of *PBX1* in pre-osteoblasts led to enhanced cell proliferation and suppression of the activation of *Runx2* and *Osterix* gene expression, the two indispensable factors required for commitment and development of skeletal lineage cells (24,25). The accelerated proliferation and increased cell number was associated with advanced matrix mineralization and enhanced bone nodule formation. These results suggested that *PBX1* is likely acting as a negative regulator for osteoblast proliferation, and also plays a role in terminal differentiation. Indeed, these results are consistent with the previous observation that *PBX1* is required in skeletal patterning and programming. *PBX1* homozygous knockout mice showed several defects in organogenesis and differentiation, including skeletogenesis and chondrocyte differentiation (26). $PBX1^{-/-}$ embryos had thinning, compressions and fusions of the vertebral bodies and neural arches. Defects in proximal but not distal forelimbs and hindlimbs were also observed. $PBX1^{-/-}$ mice also had precocious expression of *Col1a1*, a marker of bone formation, and accelerated endochondral ossification. *PBX1* is thus likely important in developmental programs and plays a role in coordinating the extent and/or timing of chondrocyte and osteoblast proliferation and terminal differentiation. The present *in vitro* observations together with our genetic association data with spine BMD strongly suggest that the *PBX1* gene is important in bone metabolism.

Our pilot study has identified *PBX1* as a novel susceptibility gene for osteoporosis. First, we used informative sib-pairs from the original linkage cohort to identify the susceptibility gene underlying the peak to enhance the chance of identifying the QTL gene. Secondly, successful replications in two independent case–control cohorts further support the association of *PBX1* with BMD variation. Finally, we determined the functional role of *PBX1* in osteoblast biology with *in vitro* cellular studies. There were, nonetheless, several limitations to this study. There exists the possibility that our BMD association data are a false positive finding. Although two cohorts with different designs were used for replication, more replication is necessary to confirm these findings. The replication data in the Japanese population were of marginal significance. This Japanese case–control cohort comprised of subjects with high and low BMD at either hip or spine, and based on the

evidence that genetic control of BMD is site specific (27), it is likely that this cohort is less powerful than the Chinese cohort that comprised of only high- and low-spine BMD subjects. Recent studies suggest sex-specific (28) and menopausal-specific (29) effects for candidate genes underlying BMD variation, but we did not perform sub-group analysis for this study as the original linkage signal was observed in the whole population, not any sub-group (15). For the *in vitro* study, only transient knock-down experiments were done and further stable gene-silencing studies are necessary to unravel the mechanisms whereby *PBX1* affects bone biology.

To summarize, the advantage of our powerful and robust study design ensured a greater chance of identifying the candidate genes in chromosome 1q. Our pilot study that used high-density SNPs on chromosome 1q QTL has narrowed down the list of candidate genes to *PBX1*. The consistent associations between rs2800791 in *PBX1* and BMD variation in almost 3000 subjects, including both sib-pair-based and case–control-based cohorts, together with the expression of *PBX1* in human bone cells and the data obtained from *in vitro PBX1* knockdown analysis, suggest that *PBX1* is a novel susceptibility gene for osteoporosis and in osteoblastogenesis. Further replication study and extensive characterization of the mechanisms of *PBX1* and its variant alleles in BMD variation may aid the development of new management strategies for patients with osteoporosis.

MATERIAL AND METHODS

Study population

All study subjects were drawn from the expanding database being compiled at the Osteoporosis Centre at Queen Mary Hospital, the University of Hong Kong, for the study of genetics and environmental determinants of osteoporosis in Southern Chinese. Subjects were invited to attend the Osteoporosis Center for BMD assessment. Those with BMD *Z*-score of −1.3 or less at either the lumbar spine L1−4 or at the femoral neck (equivalent to the lowest 10th percentile of the population) were defined as proband, and their family members were also invited for the study. A detailed description of subject ascertainment, inclusion and exclusion criteria has been described previously (30). In the previous chromosome 1q linkage study, a total of 306 families with 1459 individuals (293 males and 1166 females) were analyzed (15). These pedigrees contained 1260 sib-pairs, 143 cousin pairs, 2356 parent–child pairs, 522 grandparent–grandchild pairs and 512 avuncular pairs.

To fine-map the region for BMD variation in chromosome 1q, we adopted a new approach to select the most informative sib-pairs based on both phenotype and identity-by-descent (IBD) information. The IBD information at chromosome 1q between all 1260 sib-pairs was obtained from the MERLIN program using the microsatellite data from the linkage study. This selection procedure ensured the retention of most association information for a smaller number of subjects submitted for further genotyping. Using the Genetic Power Calculator (http://pngu.mgh.harvard.edu/~purcell/gpc/), the non-centrality parameter (NCP) of an overall (between and within sib-ship) association test, for all 1260 sib-pairs, was

estimated to be 255 (assuming a QTL that contributes 10% of the phenotypic variance and perfect LD with QTL and marker). A software described elsewhere (Kwan *et al.*, submitted) was used to rank the sib-pairs for information content, given their phenotypes and IBD sharing; genotyping the most informative 610, 464 and 355 sib-pairs retained approximately 90, 80 and 70% of the total (NCP) (Supplementary Material, Fig. S1). The most informative 610 sib-pairs (Supplementary Material, Table S1) were subsequently selected and genotyped using high-density SNP markers in order to retain 90% of the total NCP from the entire sample.

For the second screening and replication, we adopted a case–control design with subjects selected from the opposite extremes of the distribution of BMD trait values. Although cumbersome, replication of association signals in cohorts of extremes is a cost-effective and powerful approach. This replication sample came from another independent unrelated Southern Chinese cohort, with a total of 835 extremely high and extremely low BMD subjects (Supplementary Material, Table S2). Low BMD subjects were arbitrarily defined as those with a BMD T-score -1 or less (equivalent to osteopenia according to WHO definition) at either the spine or hip. High BMD subjects were sex-matched unrelated controls with a BMD T-score of $+1$ or more (approximately equivalent to 85th percentile of the population). The controls were matched as a group with the cases for age and sex. Detailed inclusion and exclusion criteria and a detailed description of the cohort ascertainment were reported previously (16). This case–control cohort had >95% power of detecting a significant association with an SNP that has MAF of 0.3 and contributes to 1% variance in BMD. This assumes that the testing SNP is the true causal polymorphism or in complete LD with the causal polymorphism. The power is >92% if the LD is >0.8.

For the third screening, 1268 case–control Japanese subjects with 703 osteoporotic subjects and 565 normal controls were studied. Osteoporosis was defined according to the Japanese Osteoporosis Society as BMD < 70% of young adult mean at either the spine or femur (31). This cutoff is equivalent to the WHO definition of T-score of less than -2.5. Controls were age-matched subjects with normal BMD of T-score of greater than -1 (cases 74.0 ± 7.2 (mean $\pm$ SD) years; controls 73.0 ± 6.2 years).

Informed consent was obtained from all subjects, and the study was approved by the ethics committee of the University of Hong Kong and conducted according of the Declaration of Helsinki.

DNA collection, marker selection and genotyping

Genomic DNA extraction from peripheral blood leukocytes was performed using standard phenol/chloroform protocols. In our previous study, the linkage peak was centered on marker D1S196. For the purpose of fine-mapping, the QTL was defined by a drop of 1 in LOD score from the peak, i.e. from 162.5 to 168.5 Mb. According to the NCBI Map Viewer B36.2, 60 genes were located in this region. In the designated region, SNPs were selected from Perlegen genome browser (32) and HAPMAP (33) Chinese population. Gene-based tagging was performed using CLUSTAG (34,35)

with r^2 threshold of 0.8 and 0.6 for known gene and hypothetical gene/pseudogene, respectively. A total of 380 SNPs were selected as tagging SNPs and force tagging was done for the SNPs located in the coding regions and promoter region.

Genotyping was performed using the SEQUENOM genotyping platform. The allele frequencies used in the study were estimated by direct counting. SNPs with minor allele frequency (MAF) <0.05, call rate <0.9 and HWE <0.01 were excluded from analysis. The physical positions used were from the NCBI B36.2 assembly, database SNP B128.

LD analysis

Pairwise LD was calculated as r^2 and D' for all SNP–pair combinations using LDMAX routine in software GOLD (36). Graphical overviews of r^2 and D' together with block structure were generated from GOLD software (36) and HAPLOVIEW (37) software, respectively.

Sib-pairs association analyses

Single locus association analyses were carried out using QFAM implemented in PLINK software (38). Total and within sib-pairs association tests were performed. The total test extracted all association information from the sib-pairs sample controlling for relatedness using a permutation procedure that separated out the between-sib-ships and the within-sib-ship components. This is more powerful than a within-sib-ship test but is not robust to population stratification. Therefore, the stringent within-sib-ship test, which is robust to population stratification, was also performed. The QFAM analyses were performed based on the presumption of an additive model. A total of 10 000 permutations were analyzed to correct for relatedness. A nominal *P*-value of 0.05 was considered a significant association. SNPs with $P \leq 0.05$ in both total- and within-association tests were replicated in an independent case–control cohort.

Statistical analysis for case–control replication study

To assess the association of SNPs and BMD status (high versus low), OR and 95% CI were determined using binary logistic regression with or without adjustment for age, sex, height and weight. The additive model of 'a' allele was used to evaluate the association. Genotypes AA, Aa and aa were coded additively as 0, 1, 2, whereby 'A' represents the major allele and 'a' represents the minor allele. SPSS 14.0 for Windows was used for statistical calculations. Meta-analysis (random effects model) was performed to analyze the results from southern Chinese and Japanese (39).

HBDC and MC3T3-E1 cells

Human bone-derived cells from healthy subjects were generated as described previously (40). The isolated adherent HBDCs were harvested for mRNA extraction or fixed with 10% paraformaldehyde for immunohistochemistry study. MC3T3-E1 mouse pre-osteoblast cells were grown in 100 mm diameter tissue culture dish until sub-confluence and passaged with 0.05% trypsin and 0.53 mM EDTA. For

transfection experiments, cells were plated at a density of 2×10^4 cells cm^{-2}. All treatments were initiated after 18 h of incubation, and tissue culture mediums were replenished every 2 days. All cells were maintained at 37°C in a humidified 5% CO$_2$ environment.

Interfering RNA transfection

Murine MC3T3-E1 pre-osteoblast cells (clone 4; ATCC) were maintained in α-MEM medium supplemented with 10% fetal bovine serum (40). Oligonucleotides to produce small interfering RNA (RNAi) specific for *PBX1* were transfected into MC3T3-E1 cells at a final concentration of 50 nM with Lipofectamine 2000 according to the manufacturer's instructions (Invitrogen). The *PBX1* Stealth™ RNAi negative control was created based on the following sequence: CGAGGCG-GAAGAGACGGAATTTCAA. RNA with medium GC content was used as negative control. Experiments were performed in duplicate and repeated three times.

Cell proliferation assay

Cell proliferation was evaluated using the MTT assay as described previously (41). Briefly, after RNAi transfection, MC3T3-E1 pre-osteoblasts were incubated for 4 days in a 24-well plate. The MTT was added at a final concentration of 0.5 mg/ml for 3 h. Cells were then solubilized with isopropanol containing 0.1% sodium dodecyl sulfate and 0.04 N hydrochloric acid. Absorbance was measured at 570 nm. All experiments were performed in triplicates and repeated three times.

Semi-quantitative and quantitative real-time PCR

Total RNA was isolated from HBDCs and MC3T3-E1 cells at specific time points using Tri-reagent™. Real-time PCR was performed using the Taqman® one-step PCR master mix kit with gene-specific target primers and 3-FAM probes. Primers and PCR conditions are provided in the supplementary information. Each sample was simultaneously quantified using eukaryotic 18S primers and target probes. Amplification was performed in an ABI Prism 7900 HT sequence detector, using standard settings of 40 cycles with an annealing temperature of 60°C. All experiments were performed in triplicates and repeated three times.

Protein expression studies

For western blot analysis, whole-cell extracts of MC3T3-E1 were prepared by M-Per lysis buffer (Pierce Biotechnology) containing complete protease inhibitor mix (Roche Applied Science). Equal amounts of protein (10–20 μg) were resolved by 12% SDS–PAGE and electrotransferred to a nitrocellulose membrane. A mouse monoclonal primary antibody for *PBX1* (Abnova) with 1/1000 dilution was used. Protein expression was detected by ECL plus (Amersham Biosciences). For immunohistochemical analysis, HBDCs were fixed in ice-cold 10% formalin in PBS for 15 min. Cells were then incubated in a humidified chamber with 1:200 dilution of the primary antibody for *PBX1* (Abnova) and *col1a1* (Santa Cruz) and 1:1000 Alexa Fluro conjugated secondary antibody. In the negative control, primary antibody was replaced by mouse sera. All experiments were performed in triplicates and repeated three times.

Von-Kossa staining on mineralized nodules

For fixation and staining of mineralized bone nodules, cells cultured for 30 days were washed twice with ice-cold PBS and fixed in ice-cold fixative (10% formalin in PBS) for 15 min. Following three washes with de-ionized water, mineralized matrix was detected using Von-Kossa staining by treating fixed cells with 5% silver nitrate for 30 min, followed by washing with 5% sodium carbonate in 10% formalin for 1 min and 5% sodium thiosulfate for 5 min. The reaction was stopped by washing the cells with de-ionized water. The Von Kossa-stained nodules were viewed and counted under a microscope. All experiments were performed in triplicates and repeated three times.

SUPPLEMENTARY MATERIAL

Supplementary Material is available at *HMG* online.

REFERENCES

1. WHO (1994) Assessment of fracture risk and its application to screening for postmenopausal osteoporosis. World Health Organization, Switzerland.
2. Health NIo (2000) NIH consensus statement: osteoporosis prevention, diagnosis, and therapy. *NIH Consensus Statement*, **17**, 1–45.
3. Eisman, J.A. (1999) Genetics of osteoporosis. *Endocrinol. Rev.*, **20**, 788–804.
4. Flicker, L., Hopper, J.L., Rodgers, L., Kaymakci, B., Green, R.M. and Wark, J.D. (1995) Bone density determinants in elderly women: a twin study. *J. Bone Miner. Res.*, **10**, 1607–1613.
5. Pocock, N.A., Eisman, J.A., Hopper, J.L., Yeates, M.G., Sambrook, P.N. and Eberl, S. (1987) Genetic determinants of bone mass in adults. A twin study. *J. Clin. Invest.*, **80**, 706–710.
6. Smith, D.M., Nance, W.E., Kang, K.W., Christian, J.C. and Johnston, C.C. Jr (1973) Genetic factors in determining bone mass. *J. Clin. Invest.*, **52**, 2800–2808.
7. Ng, M.Y.M., Sham, P.C., Paterson, A.D., Chan, V. and Kung, A.W.C. (2005) Effect of environmental factors and gender on the heritability of bone mineral density and bone size. *Ann. Hum. Genet.*, **69**, 1–11.
8. Wilson, S.G., Reed, P.W., Andrew, T., Barber, M.J., Lindersson, M., Langdown, M., Thompson, D., Thompson, E., Bailey, M., Chiano, M. *et al.* (2004) A genome-screen of a large twin cohort reveals linkage

for quantitative ultrasound of the calcaneus to 2q33–37 and 4q12–21. *J. Bone Miner. Res.*, **19**, 270–277.

9. Devoto, M., Shimoya, K., Caminis, J., Ott, J., Tenenhouse, A., Whyte, M.P., Sereda, L., Hall, S., Considine, E., Williams, C.J. *et al.* (1998) First-stage autosomal genome screen in extended pedigrees suggests genes predisposing to low bone mineral density on chromosomes 1p, 2p and 4q. *Eur. J. Hum. Genet.*, **6**, 151–157.

10. Devoto, M., Spotila, L.D., Stabley, D.L., Wharton, G.N., Rydbeck, H., Korkko, J., Kosich, R., Prockop, D., Tenenhouse, A. and Sol-Church, K. (2005) Univariate and bivariate variance component linkage analysis of a whole-genome scan for loci contributing to bone mineral density. *Eur. J. Hum. Genet.*, **13**, 781–788.

11. Koller, D.L., Econs, M.J., Morin, P.A., Christian, J.C., Hui, S.L., Parry, P., Curran, M.E., Rodriguez, L.A., Conneally, P.M., Joslyn, G. *et al.* (2000) Genome screen for QTLs contributing to normal variation in bone mineral density and osteoporosis. *J. Clin. Endocrinol. Metab.*, **85**, 3116–3120.

12. Econs, M.J., Koller, D.L., Hui, S.L., Fishburn, T., Conneally, P.M., Johnston, C.C. Jr, Peacock, M. and Foroud, T.M. (2004) Confirmation of linkage to chromosome 1q for peak vertebral bone mineral density in premenopausal white women. *Am. J. Hum. Genet.*, **74**, 223–228.

13. Beamer, W.G., Shultz, K.L., Donahue, L.R., Churchill, G.A., Sen, S., Wergedal, J.R., Baylink, D.J. and Rosen, C.J. (2001) Quantitative trait loci for femoral and lumbar vertebral bone mineral density in C57BL/6J and C3H/HeJ inbred strains of mice. *J. Bone Miner. Res.*, **16**, 1195–1206.

14. Ioannidis, J.P., Ng, M.Y., Sham, P.C., Zintzaras, E., Lewis, C.M., Deng, H.W., Econs, M.J., Karasik, D., Devoto, M., Kammerer, C.M. *et al.* (2007) Meta-analysis of genome-wide scans provides evidence for sex- and site-specific regulation of bone mass. *J. Bone Miner. Res.*, **22**, 173–183.

15. Cheung, C.L., Huang, Q.Y., Ng, M.Y., Chan, V., Sham, P.C. and Kung, A.W.C. (2006) Confirmation of linkage to chromosome 1q for spine bone mineral density in southern Chinese. *Hum. Genet.*, **120**, 354–359.

16. Li, Y. and Abecasis, G.R. (2005) Mach 1.0: rapid haplotype reconstruction and missing genotype inference. *Am. J. Hum. Genet.*, **S79**, 2290.

17. Shang, Y. and Brown, M. (2002) Molecular determinants for the tissue specificity of SERMs. *Science*, **295**, 2465–2468.

18. Ritchie, M.D., Hahn, L.W., Roodi, N., Bailey, L.R., Dupont, W.D., Parl, F.F. and Moore, J.H. (2001) Multifactor-dimensionality reduction reveals high-order interactions among estrogen-metabolism genes in sporadic breast cancer. *Am. J. Hum. Genet.*, **69**, 138–147.

19. Bryne, J.C., Valen, E., Tang, M.H., Marstrand, T., Winther, O., da Piedade, I., Krogh, A., Lenhard, B. and Sandelin, A. (2008) JASPAR, the open access database of transcription factor-binding profiles: new content and tools in the 2008 update. *Nucleic Acids Res.*, **36**, D102–D106.

20. Lau, H.H.L., Ng, M.Y.M., Ho, A.Y.Y., Luk, K.D.K. and Kung, A.W.C. (2005) Genetic and environmental determinants of bone mineral density in Chinese women. *Bone*, **36**, 700.

21. Stranger, B.E., Nica, A.C., Forrest, M.S., Dimas, A., Bird, C.P., Beazley, C., Ingle, C.E., Dunning, M., Flicek, P., Koller, D. *et al.* (2007) Population genomics of human gene expression. *Nat. Genet.*, **39**, 1217–1224.

22. Styrkarsdottir, U., Halldorsson, B.V., Gretarsdottir, S., Gudbjartsson, D.F., Walters, G.B., Ingvarsson, T., Jonsdottir, T., Saemundsdottir, J., Center, J.R., Nguyen, T.V. *et al.* (2008) Multiple Genetic Loci for Bone Mineral Density and Fractures. *N. Engl. J. Med.*, **358**, 2355–2365.

23. Richards, J.B., Rivadeneira, F., Inouye, M., Pastinen, T.M., Soranzo, N., Wilson, S.G., Andrew, T., Falchi, M., Gwilliam, R., Ahmadi, K.R. *et al.* (2008) Bone mineral density, osteoporosis, and osteoporotic fractures: a genome-wide association study. *Lancet*, **371**, 1505–1512.

24. Komori, T., Yagi, H., Nomura, S., Yamaguchi, A., Sasaki, K., Deguchi, K., Shimizu, Y., Bronson, R.T., Gao, Y.H., Inada, M. *et al.* (1997) Targeted disruption of Cbfa1 results in a complete lack of bone formation owing to maturational arrest of osteoblasts. *Cell*, **89**, 755–764.

25. Otto, F., Thornell, A.P., Crompton, T., Denzel, A., Gilmour, K.C., Rosewell, I.R., Stamp, G.W., Beddington, R.S., Mundlos, S., Olsen, B.R. *et al.* (1997) Cbfa1, a candidate gene for cleidocranial dysplasia syndrome, is essential for osteoblast differentiation and bone development. *Cell*, **89**, 765–771.

26. Selleri, L., Depew, M.J., Jacobs, Y., Chanda, S.K., Tsang, K.Y., Cheah, K.S., Rubenstein, J.L., O'Gorman, S. and Cleary, M.L. (2001) Requirement for *PBX1* in skeletal patterning and programming chondrocyte proliferation and differentiation. *Development*, **128**, 3543–3557.

27. Ralston, S.H., Galwey, N., MacKay, I., Albagha, O.M., Cardon, L., Compston, J.E., Cooper, C., Duncan, E., Keen, R., Langdahl, B. *et al.* (2005) Loci for regulation of bone mineral density in men and women identified by genome wide linkage scan: the FAMOS study. *Hum. Mol. Genet.*, **14**, 943–951.

28. Huang, Q.Y., Ng, M.Y., Cheung, C.L., Chan, V., Sham, P.C. and Kung, A.W. (2006) Identification of two sex-specific quantitative trait Loci in chromosome 11q for hip bone mineral density in Chinese. *Hum. Hered.*, **61**, 237–243.

29. Cheung, C.L., Chan, V. and Kung, A.W. (2008) A differential association of ALOX15 polymorphisms with bone mineral density in pre- and post-menopausal women. *Hum. Hered.*, **65**, 1–8.

30. Kung, A.W.C., Lai, B.M.H., Ng, M.Y., Chan, V. and Sham, P.C. (2006) T-1213C polymorphism of estrogen receptor beta is associated with low bone mineral density and osteoporotic fractures. *Bone*, **39**, 1097–1106.

31. Orimo, H., Hayashi, Y., Fukunaga, M., Sone, T., Fujiwara, S., Shiraki, M., Kushida, K., Miyamoto, S., Soen, S., Nishimura, J. *et al.* (2001) Diagnostic criteria for primary osteoporosis: year 2000 revision. *J. Bone Miner. Metab.*, **19**, 331–337.

32. Hinds, D.A., Stuve, L.L., Nilsen, G.B., Halperin, E., Eskin, E., Ballinger, D.G., Frazer, K.A. and Cox, D.R. (2005) Whole-genome patterns of common DNA variation in three human populations. *Science*, **307**, 1072–1079.

33. International HapMap Consortium (2005) A haplotype map of the human genome. *Nature*, **437**, 1299–1320.

34. Ao, S.I., Yip, K., Ng, M., Cheung, D., Fong, P.Y., Melhado, I. and Sham, P.C. (2005) CLUSTAG: hierarchical clustering and graph methods for selecting tag SNPs. *Bioinformatics*, **21**, 1735–1736.

35. Sham, P.C., Ao, S.I., Kwan, J.S., Kao, P., Cheung, F., Fong, P.Y. and Ng, M.K. (2007) Combining functional and linkage disequilibrium information in the selection of tag SNPs. *Bioinformatics*, **23**, 129–131.

36. Abecasis, G.R. and Cookson, W.O. (2000) GOLD—graphical overview of linkage disequilibrium. *Bioinformatics*, **16**, 182–183.

37. Barrett, J.C., Fry, B., Maller, J. and Daly, M.J. (2005) Haploview: analysis and visualization of LD and haplotype maps. *Bioinformatics*, **21**, 263–265.

38. Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M.A., Bender, D., Maller, J., Sklar, P., de Bakker, P.I., Daly, M.J. and Sham, P.C. (2007) PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am. J. Hum. Genet.*, **81**, 559–575.

39. Bax, L., Yu, L.M., Ikeda, N., Tsuruta, H. and Moons, K.G. (2006) Development and validation of MIX: comprehensive free software for meta-analysis of causal research data. *BMC Med. Res. Methodol.*, **6**, 50.

40. Gartland, A., Buckley, K.A., Dillon, J.P., Curran, J.M., Hunt, J.A. and Gallagher, J.A. (2005) Isolation and culture of human osteoblasts. *BMC Med. Res. Methodol.*, **107**, 29–54.

41. Cheung, W.M.W., Ng, W.W. and Kung, A.W. (2006) Dimethyl sulfoxide as an inducer of differentiation in preosteoblast MC3T3-E1 cells. *FEBS Lett.*, **580**, 121–126.