

Identification of QTL genes for BMD variation using both linkage and gene-based association approaches

Gloria Hoi-Yee Li · Ching-Lung Cheung · Su-Mei Xiao · Kam-Shing Lau · Yi Gao · Cora H. Bow · Qing-Yang Huang · Pak-Chung Sham · Annie Wai-Chee Kung

Received: 30 November 2010 / Accepted: 28 February 2011 / Published online: 19 March 2011
© The Author(s) 2011. This article is published with open access at Springerlink.com

Abstract Low bone mineral density (BMD) is a risk factor for osteoporotic fracture with a high heritability. Previous large scale linkage study in Northern Chinese has identified four significant quantitative trait loci (QTL) for BMD variation on chromosome 2q24, 5q21, 7p21 and 13q21. We performed a replication study of these four QTL in 1,459 Southern Chinese from 306 pedigrees. Successful replication was observed on chromosome 5q21 for femoral neck BMD with a LOD score of 1.38 (nominal p value = 0.006). We have previously identified this locus in a genome scan meta-analysis of BMD variation in a white population. Subsequent QTL-wide gene-based association analysis in 800 subjects with extreme BMD identified *CAST* and *ERAP1* as novel BMD candidate genes (empirical p value of 0.032 and 0.014, respectively). The associations were independently replicated in a Northern European population (empirical p value of 0.01 and 0.004 for *CAST* and *ERAP1*, respectively). These findings

provide further evidence that 5q21 is a BMD QTL, and *CAST* and *ERAP1* may be associated with femoral neck BMD variation.

Introduction

Family-based linkage analysis has successfully identified genes that contribute to relatively rare disorders with monogenic patterns of inheritance. In contrast, efforts to extend this family-based approach to common disorders and quantitative traits have been less successful. It is often particularly difficult to identify the causal variant under the quantitative trait loci (QTL). Linkage analysis is based on the principle of identification by descent (Cheung et al. 2010) to test the co-segregation of a genetic marker (e.g., microsatellite marker) and the trait-of-interest. The co-segregated genomic loci can also be very large, and multiple disease genes may be present under the QTL. In addition, although there may be only one disease gene in the QTL, the linkage signal may be elicited by multiple causal variants in the disease gene and in different families. The presence of allelic heterogeneity thus limits the ability to identify the causal variant using a SNP association approach.

To date, more than fifteen genome-wide linkage scans for bone mineral density (BMD) have been performed in study populations of different ethnicity, of which two were Chinese. Niu et al. (1999) determined that 2p21 and 13q34 are modestly linked to forearm BMD in a genome-wide linkage scan conducted in 218 individuals from Anqing district. Hsu et al. (2007) observed significant evidence of linkage at 2q24 for total hip, 7p21 for femoral neck and 5q21 for combined BMD in 3,093 siblings from Anhui, and 13q21 for lumbar spine BMD in women in the same study

G. H.-Y. Li, C.-L. Cheung contributed equally to this work.

Electronic supplementary material The online version of this article (doi:10.1007/s00439-011-0972-2) contains supplementary material, which is available to authorized users.

G. H.-Y. Li · C.-L. Cheung (✉) · S.-M. Xiao · K.-S. Lau · Y. Gao · C. H. Bow · Q.-Y. Huang · A. W.-C. Kung
Department of Medicine, The University of Hong Kong,
Lab Block, 21 Sassoon Road, Hong Kong, Hong Kong
e-mail: lung1212@hku.hk

S.-M. Xiao · A. W.-C. Kung
Research Centre of Heart, Brain, Hormone and Healthy Aging,
The University of Hong Kong, Hong Kong, Hong Kong

P.-C. Sham
Department of Psychiatry, The University of Hong Kong,
Hong Kong, Hong Kong

population. With reported genetic heterogeneity between Northern and Southern Chinese, it is not known if these quantitative trait loci (QTL) are associated with BMD variation in Southern Chinese. In addition, the QTL genes have not been identified.

Independent replication of linkage is a fundamental prerequisite for the commencement of positional cloning studies, in order to reduce the possibility of false positive initial finding. It is believed that some genes associated with BMD are likely to be site-specific and sex-specific; we therefore performed linkage analysis using high density microsatellite markers to investigate the linkage between four QTLs reported by Hsu et al. (2007) in a sample of 1,459 individuals from 306 Southern Chinese families of probands with low BMD. Subsequent gene-based association study was performed to identify potential QTL genes underlying the linkage peak.

Methods and materials

Subjects for linkage and association analyses

Subjects for the current study were drawn from an expanding database of more than 8,000 subjects at the Osteoporosis Centre, Queen Mary Hospital, Hong Kong. Subject recruitment has been ongoing since 1998 with individuals drawn from road shows and health talks on osteoporosis. Study subjects gave informed written consent and were invited to the Centre for bone mineral density measurement and physical examination. They were also interviewed by a trained research assistant using a structured questionnaire to obtain data on ethnicity, social, medical and reproductive history, dietary and lifestyle factors, and family history of osteoporosis. Families with a proband with BMD Z-score ≤ -1.28 (which is equivalent to the lowest 10% of the population) at either the lumbar spine or hip were identified. Individuals with a previous or family history of disease known to affect bone metabolism, premature menopause (age <40), bilateral oophorectomy or drug use that could affect bone turnover and BMD were excluded from study. All study subjects were of Southern Chinese Han descent. Based on the previous estimates of the heritability of BMD (Ng et al. 2006) and the relative informativeness of the 1,021 pedigrees in our cohort (Cheung et al. 2008), 306 families were selected for the current study.

In the association study of QTL genes, 800 unrelated Hong Kong Southern Chinese (HKSC) women with extremely high ($n = 376$) or low BMD ($n = 424$) were selected from the same database as mentioned above. This study aimed to identify genetic factors that affect BMD variation in women. We adopted a threshold defined case-

control study design. Subjects with BMD Z-score ≤ -1.28 at either the lumbar spine (LS) or femoral neck (FN) (the lowest 10% of the total population) were categorized as low-BMD subjects. High-BMD subjects comprised individuals with BMD Z-score $\geq +1.0$ at either site (which is equivalent to the highest 15% of the total population). These selection criteria captured the extreme 25% BMD information of the total population. Subjects who reported the presence of disease or environmental factors that may affect BMD and bone metabolism were excluded from study. The recruitment procedure and exclusion criteria have been detailed elsewhere (Cheung et al. 2008) but the latter included the following: a history of chronic medical illness, premature menopause age below 40 years, malabsorption, previous major gastrointestinal surgery, metabolic bone disease, endocrine disorders including hyper- and hypothyroidism, or prescription of medication that may affect bone and calcium metabolism, hormone replacement therapy, anti-osteoporosis medication, and active vitamin D3 metabolites. The 800 studied subjects selected for the association study did not overlap with the pedigrees for linkage analysis.

Measurement

BMD (g/cm^2) at the lumbar spine (L1–L4), femoral neck (FN), trochanter and total hip was measured using dual-energy X-ray absorptiometry (DEXA, Hologic QDR 2000 plus, Hologic, Waltham, MA, USA). The in vivo precision of the machine for lumbar spine, FN and total hip region was 1.2, 1.5 and 1.5%, respectively.

Microsatellite and SNP genotyping

Genomic DNA was extracted from peripheral blood leucocytes of the subjects using a standard phenol/chloroform method. Nineteen selected microsatellite markers (Table 3) in close proximity to the four loci reported by Hsu et al. (2007) were selected using the marker order and map positions obtained from the Marshfield electronic database (<http://www.research.marshfieldclinic.org/genetics/MarkerSearch/searchMarkers.asp>). Fluorescently labeled PCR primers (ABI) were used to amplify the selected microsatellite loci. Electrophoresis and size determination of the amplified DNA fragments were performed using a Genetic Analyzer 3700 (ABI) and GENESCAN (ABI), respectively. Control DNA CEPH 1347-02 was used to monitor PCR amplification efficiency and control for gel-to-gel variation.

For the gene-based association study, subjects were genotyped via the Infinium assay (Illumina, San Diego, CA, USA) with Human610-quad chip including 564,214 SNPs. PLINK (version 1.04) used for data management

and quality control statistics. After exclusion of individuals based on strict quality-control criteria, 785 individuals and 488,853 SNPs were retained for analysis. Subjects were excluded according to the following criteria: (1) genotyping call rate less than 95% ($n = 5$); (2) autosomal heterozygosity <27 or $>31\%$ (the same five subjects with low genotyping call rate); (3) related or identical to other individuals in the sample ($n = 7$); and (4) discordance of observed gender and estimated sex ($n = 3$). SNPs were excluded if: (1) genotyping call rate was $\leq 95\%$ (1,158 SNPs), (2) Hardy–Weinberg equilibrium (HWE) p value $<1.0 \times 10^{-4}$ (904 SNPs), and (3) minor allele frequency (MAF) <0.01 (73,589 SNPs). The average genotyping call rate of retained SNPs was 99.91%.

Linkage analysis and gene-based association study

Genotype inconsistencies were identified using Pedstats (Wigginton and Abecasis 2005) and linkage analysis conducted using Merlin-regress (Sham et al. 2002). To control for gender and age effects, a multi-point analysis was performed at 0.5 cM grid of equal-space locations for BMD Z-scores at the lumbar spine, femoral neck, trochanter and total hip in the whole population and all sub-groups. Sub-group analysis was performed by setting trait values for the uninterested subgroups as missing. A sex-averaged genetic map was adopted for the whole-sample analysis and a sex-specific map was applied for sub-group analysis. A heritability estimate of 0.7 was employed based on our previous heritability study of BMD at the spine and hip (Ng et al. 2006). A nominal p value of 0.01 is required to declare successful replication of a previously reported linkage signal (Lander and Kruglyak 1995).

For the gene-based association study, QTL with a nominal p value <0.01 were selected for regional gene-based association study. Candidate genes located within the QTL were included in the analysis. We first obtained a standardized residual of BMD with adjustment of age and weight, since height was not significantly associated with BMD in our cohort. This standardized residual of BMD was normally distributed (Kolmogorov–Smirnov test p value >0.05). The association p value was obtained using PLINK (version 1.04) with the linear regression model testing the association between standardized residual of BMD with each SNP. For each SNP, the asymptotic p value for the relationship between the number of minor alleles and BMD was derived from a two-sided t statistic assuming the minor allele had an additive effect. The association p values were used to calculate gene-based test statistics of each gene ($n = 18$). The empirical p value was calculated as the proportion of simulated test statistics (using Monte Carlo approach) that exceeded the observed gene-based test statistic. In silico replication in 5,858

European subjects was performed using the p value obtained from a recent meta-analysis of GWAS of BMD (Styrkarsdottir et al. 2008). Gene-based analysis in each population was performed using VEGAS (Liu et al. 2010). The brief description of VEGAS is provided in supplementary information.

In silico gene expression study

In order to study whether the candidate genes are regulated by osteogenic molecule (bmp2) in murine MC3T3-1b pre-osteoblast cell line, we performed in silico gene expression study by retrieving the normalized microarray data [Accession number: GDS679] (Zamurovic et al. 2004) from GEO database. Two datasets were compared, one was grown under the stimulation by bmp2 for 1 day, and the control was grown without stimulation of bmp2. The differences in total RNA expression in MC3T3-1b cells with and without stimulation by bmp2 and ascorbic acid were calculated using an independent t test.

Results

Clinical characteristics of the subjects for linkage and association studies are described in Tables 1 and 2, respectively. Briefly, there were 306 Mendelian-consistent pedigrees, with 1,166 females and 293 males. The average size and number of generations of these pedigrees was 6.46 and 2.33, respectively. The overall drop-out rate of microsatellite markers was 3.65%. These genotypes were discarded due to either low intensity of the amplified PCR product or Mendelian inconsistency. The average heterozygosity of these markers was 0.777. Details of the 19 markers included in this study are shown in Table 3.

In the whole-sample analysis, a multi-point maximum LOD score (MLS) of 1.383 (nominal $p = 0.006$) was detected at 107.26 cM on 5q21–22 for femoral neck BMD (Fig. 1). Sub-group analysis was also performed, but no LOD score exceeded one. For other QTL, no region had LOD score exceeding one.

QTL on chromosome 5q with a nominal p value ≤ 0.01 were defined between 104.76 and 112.26 cM, equivalent to physical position 95,738,450–104,090,852 (hg 18) approximately. A total of 18 genes were annotated in this region: *PCSK1*, *CAST*, *ERAP1*, *ERAP2*, *LNPEP*, *LIX1*, *RIOK2*, *RGMB*, *CHD1*, *FAM174A*, *ST8SIA4*, *SLCO4C1*, *SLCO6A1*, *PAM*, *GIN1*, *HISPPD1*, *C5orf30* and *NUDT12*. To identify a BMD gene(s) responsible for the QTL on chromosome 5p21–22 in Hong Kong Southern Chinese, we conducted a gene-based statistical analysis for femoral neck BMD on 800 unrelated individuals with extreme BMD using genotypes from 910 heterozygous SNPs (as defined by MAF

Table 1 Demographic information of 1,459 individuals from 306 multi-generation pedigrees

Subject	Total (<i>n</i> = 1,459)	Male (<i>n</i> = 293)	Female (<i>n</i> = 1,166)
Age (years)	47.4 ± 15.3	49 ± 16	47 ± 15.1
Height (cm)	157.8 ± 0.08	167.8 ± 0.06	155.2 ± 0.64
Weight (kg)	56.2 ± 10.4	66.3 ± 9.9	53.6 ± 8.9
Post-menopausal women (<i>N</i>)	485 (33.2%)	NA	485 (41.6%)
Spine BMD (g/cm ²)	0.902 ± 0.168	0.921 ± 0.149	0.897 ± 0.172
Femoral neck BMD (g/cm ²)	0.697 ± 0.129	0.720 ± 0.125	0.691 ± 0.130
Trochanter BMD (g/cm ²)	0.604 ± 0.121	0.644 ± 0.113	0.594 ± 0.121
Total hip BMD (g/cm ²)	0.797 ± 0.143	0.856 ± 0.136	0.782 ± 0.140
Spine Z-score	−0.33 ± 1.18	−0.33 ± 1.18	−0.33 ± 1.18
Femoral neck Z-score	−0.32 ± 1.04	−0.33 ± 0.80	−0.32 ± 1.09
Trochanter Z-score	−0.26 ± 1.11	−0.22 ± 0.93	−0.27 ± 1.15
Total Hip Z-score	−0.30 ± 1.15	−0.19 ± 0.95	−0.32 ± 1.20
Spine T-score	−0.80 ± 1.33	−0.55 ± 1.17	−0.86 ± 1.36
Femoral neck T-score	−0.85 ± 1.19	−0.76 ± 0.86	−0.86 ± 1.26
Trochanter T-score	−0.71 ± 1.25	−0.55 ± 0.95	−0.75 ± 1.32
Total hip T-score	−0.72 ± 1.30	−0.61 ± 0.96	−0.75 ± 1.37

Table 2 Demographic information of 800 unrelated individuals with extreme BMD

Group	High BMD	Low BMD
Number	376	424
Age (years)	46.6 ± 14.8	51.1 ± 15.9
Height (m)	1.58 ± 0.06	1.53 ± 0.07
Weight (kg)	61.1 ± 9.9	49.1 ± 6.7
Post-menopausal women (<i>N</i>)	131 (34.8%)	270 (63.7%)
BMD Z-score		
Lumbar spine	1.05 ± 0.78	−1.59 ± 0.53
Femoral neck	1.15 ± 0.79	−1.36 ± 0.60

Data are expressed as mean ± standard deviation. The 800 subjects with extreme BMD phenotype were selected from a growing database of Hong Kong Southern Chinese (more than 7,000 volunteers). The low BMD subjects are defined as an individual having BMD Z-score ≤ −1.28 at either lumbar spine (LS) or femoral neck (FN), which belongs to the lowest 10% of the total cohort, while high BMD subjects are individuals with BMD Z-score ≥ +1.0 at either of the two skeletal sites

>0.01) that were available in the Illumina Human 610 quad chip. Association analysis was restricted to femoral neck BMD, because the significant replication of linkage was observed only in femoral neck BMD. A linear regression model was used to calculate the SNP-based *p* value. Fifty-nine SNPs showed a higher than expected nominal *p* value of 0.05. Table 4 shows the results of gene-based association test of 18 genes with femoral neck BMD. Among the 18 genes, 2 genes *CAST* and *ERAPI* showed empirical *p* value <0.05 after 100,000 rounds of simulation.

We then performed in silico replication using the data from 5,858 Northern European subjects to independently

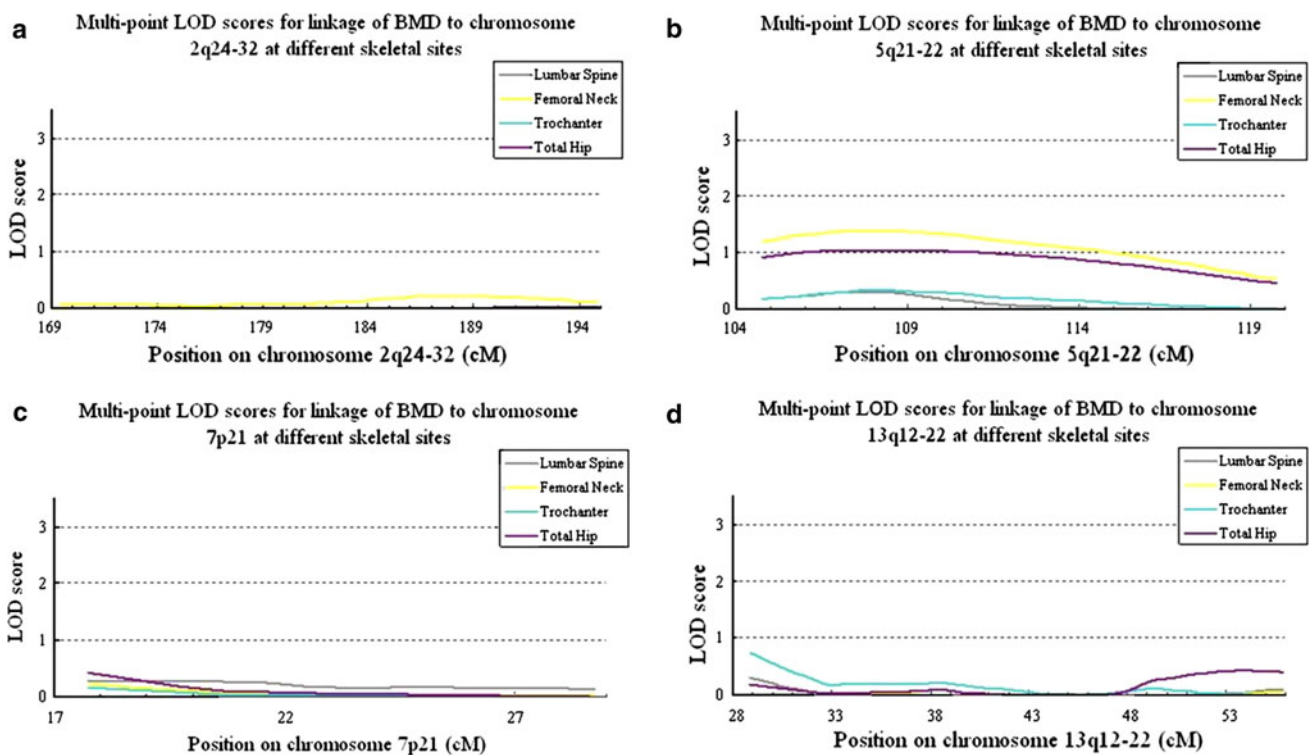
substantiate these gene-based association results in a separate ethnic population. Interestingly, these two genes also showed significant associations with femoral neck BMD with empirical *p* values <0.05. Meta-analysis using a weighted *z*-transformed test revealed a more significant *p* value of 0.019 and 0.007 for *CAST* and *ERAPI*, respectively (Table 5). To provide preliminary insight of the roles of *CAST* and *ERAPI* in osteoblast, we retrieved the normalized expression data of *CAST* from the GEO dataset record GDS679. We observed that in the presence of BMP2 and ascorbic acid, *CAST* expression was significantly decreased by 37% when compared with control (without stimulation of ascorbic acid and bmp2) with a *p* value of 0.001. Since *ERAPI* was not included in the microarray, no analysis was done.

Discussion

In this study, we replicated a previously reported linkage peak on 5q21-22 for femoral neck BMD (*p* = 0.006) in Hong Kong Southern Chinese and subsequently identified two BMD genes, *CAST* and *ERAPI*, using a gene-based association approach in HKSC and Northern European populations. Our bioinformatics analysis suggested that *CAST* is regulated by BMP in osteoblasts. In our previous meta-analysis of genome-wide scans in Caucasians, a broad region on chromosome 5 (5q14.3-q23.2) was identified as a femoral neck BMD QTL (Ioannidis et al. 2007). This bin nonetheless flanks more than 4 Mb in the genome and poses problems in identification of the QTL gene. Thus, this study is vital since it solved many of the

Table 3 The chromosomal position, heterozygosity and average inter-marker distance of the 19 selected microsatellite markers

Chromosome	Region	Marker	Sex-averaged position (cM)	Heterozygosity	Average inter-marker distance (cM)
2	2q24.3	D2S2330	169.41	0.82	6.18
	2q31.1	D2S335	175.91	0.80	
	2q31.1	D2S2188	180.79	0.64	
	2q32.1	D2S364	186.21	0.80	
	2q32.3	D2S117	194.45	0.82	
5	5q21.1	D5S644	104.76	0.85	4.91
	5q21.1	D5S495	108.07	0.80	
	5q21.3	D5S433	111.97	0.85	
	5q22.2	D5S2027	119.5	0.78	
7	7p21.3	D7S2464	17.74	0.68	3.67
	7p21.2	D7S664	20.84	0.71	
	7p21.1	D7S2557	23.29	0.73	
	7p21.1	D7S507	28.74	0.89	
13	13q12.3	D13S219	28.87	0.66	5.40
	13q13.1	D13S218	32.9	0.65	
	13q13.2	D13S263	38.32	0.85	
	13q14.2	D13S153	45.55	0.82	
	13q21.31	D13S1320	49.38	0.76	
	13q22.1	D13S156	55.85	0.80	

**Fig. 1** Multi-point LOD scores for linkage of BMD at lumbar spine, femoral neck, trochanter and total hip at (a) 2q24-32, (b) 5q21-22, (c) 7p21 and (d) 13q12-22

problems: first, we successfully replicated this QTL in HKSC, suggesting that this BMD QTL is important in determining BMD in multiple ethnicity groups. Second, the

linkage signal was only observed for femoral neck BMD, not for other skeletal sites, providing further evidence that this QTL is specific for femoral neck BMD determination.

Table 4 Gene-based association study of 18 genes located with QTL on 5q21-22 in HKSC

Gene	Start	Stop	No. of SNPs	No. of simulations	Test	Gene <i>P</i>	Best SNP	SNP <i>P</i>
PCSK1	95,751,874	95,794,708	5	1,000	1.463201	0.823	rs6860611	0.490
CAST	96,023,696	96,136,141	67	100,000	141.1646	0.034		0.0002
ERAP1	96,122,269	96,169,648	42	100,000	114.8847	0.024	rs13160562	0.0002
ERAP2	96,237,399	96,281,162	21	1,000	13.2657	0.624	rs2927612	0.040
LNPEP	96,297,101	96,390,871	18	1,000	1.771527	0.996	rs2303138	0.611
LIX1	96,453,329	96,504,276	29	1,000	10.34403	0.926	rs4869151	0.148
RIOK2	96,524,396	96,544,700	24	1,000	12.31268	0.823	rs4869151	0.148
RGMB	98,132,899	98,160,098	16	1,000	10.62398	0.608	rs12652843	0.047
CHD1	98,218,807	98,290,138	16	1,000	12.15287	0.525	rs9327919	0.036
FAM174A	99,899,022	99,950,339	15	1,000	6.197312	0.808	rs247938	0.150
ST8SIA4	100,170,802	100,266,869	17	1,000	12.09712	0.521	rs4703116	0.179
SLCO4C1	101,597,590	101,660,152	37	1,000	45.63731	0.259	rs10479200	0.004
SLCO6A1	101,735,550	101,862,619	37	1,000	26.08828	0.542	rs10479200	0.004
PAM	102,229,425	102,393,316	29	1,000	17.7014	0.649	rs467693	0.094
GIN1	102,449,602	102,483,741	4	1,000	3.431025	0.430	rs26521	0.073
HISPPD1	102,493,155	102,566,808	8	1,000	3.539095	0.573	rs26521	0.073
C5orf30	102,622,340	102,642,260	9	1,000	1.094518	0.972	rs187580	0.608
NUDT12	102,912,454	102,926,389	29	1,000	16.2771	0.831	rs861085	0.130

Gene-based *p* value <0.05 are bolded

Table 5 Replication and meta-analysis of CAST and ERAP1 in 5,858 Northern European subjects

Gene	No. of SNPs	No. of simulations	Test	Gene <i>P</i>	Best SNP	SNP <i>P</i>	Meta- <i>P</i> ^a	<i>Q</i> stats	<i>P</i> _{hetero}
CAST	48	100,000	95.53	0.032	rs42398	0.010	0.019	417.42	<0.001
ERAP1	31	100,000	87.42	0.014	rs845028	0.004	0.007	756.00	<0.001

^a Meta-analysis was done using weighted *Z*-transformed test

Third, we fine mapped the QTL to less than 1 Mb region using high density microsatellite markers. Fourth, the two BMD genes initially identified in HKSC were subsequently in silico replicated in Northern Europeans suggesting these genes affect BMD in both populations. This is in accordance with our observations in the previous meta-analysis and the current linkage result.

Among the four QTL studied, we observed significant replication only on chromosome 5, and were unable to replicate other findings by Hsu et al. (2007) of the significant QTL on chromosome 2 (2q24), 7 (7p21) and 13 (13q21). Indeed, at these chromosomal locations, our peak LOD scores were less <1 (Fig. 1). This lack of replication in genetic loci between the two studies may be due to several factors. As with other complex disease genetic studies, power to localize QTL may be limited and result in between-study discrepancies. In addition, genetic factors vary across different ethnicities. The study by Hsu et al. (2007) focused on a Northern Chinese population, whereas our study examined Southern Chinese as the genetic backgrounds of different populations in different Chinese geographical areas that are known to differ (Xu et al.

2009). If relatively rare variants are involved in the determination of BMD variation, we may expect considerable differences in the localization of the most important genetic loci across different populations or within the same population at different geographic locations. Linkage studies of such complex traits cannot exclude the possibility of the presence of genetic regions for important QTLs. Thus, the lack of concordance cannot be interpreted as evidence to dispute the hypothesis that a QTL exists in a particular genomic region. Although linkage of BMD to 7p21 was not detected in the current study, linkage to 7p15-13, a well-replicated QTL in Caucasians, was confirmed in the same Southern Chinese family cohort. In the whole-sample analysis, evidence suggestive of linkage of total hip BMD to 7p14 (MLS = 2.75, nominal *p* = 0.0002) was observed (Li et al. 2010). In fact, the loci 7p15-13 is located just beside 7p21. The differences observed between our Southern Chinese population and the Anhui sample (Hsu et al. 2007) may be attributed to the population substructure of Han Chinese (Xu et al. 2009).

Results from the linkage of femoral neck BMD detected a significant replication of QTL on chromosome 5q22-23

in HKSC participants. Further investigation of 18 genes within this region with p value ≤ 0.01 demonstrated genetic associations with two genes *CAST* and *ERAPI*. Notably, rs13160562 that showed the most significant p value was associated with these two genes as rs13160562 was located within 50 kb upstream and downstream region of these two gene loci. The femoral neck BMD Z-score for each genotype of rs13160562 is provided in Supplementary Figure 1. rs13160562 is located upstream of the gene *CAST*, but within the gene locus of *ERAPI*. Nonetheless, functional study is required to confirm whether this SNP affects *CAST* or *ERAPI* or both. In silico replication study in Northern Europeans and the meta-analysis suggested that both genes may be important for BMD variation, although the p values of heterogeneity indicate the need for cautious interpretation. Nevertheless, to further support this observation, we used RNA expression data (Zamurovic et al. 2004) from the MC3T3-1B cell and found that there was a significant decrease in *CAST* mRNA expression upon stimulation of BMP2. These data suggest that *CAST* may be a negative modulator of BMD and a downstream target of BMP2 in modulating *CAST* mRNA expression of osteoblasts. This observation is in line with a study that suggested that calpain–calpastatin system is essential for osteoblast proliferation and differentiation although the system is regulated by BMP (Murray et al. 1997). *CAST* encodes for calpastatin, which is an endogenous calpain inhibitor. In MC3T3 preosteoblast cell line, calpastatin inhibits calpains and mediates the PTH-stimulated cAMP accumulation, suggesting calpastatin is involved in PTH signaling pathway and bone metabolism (Shimada et al. 2005). *ERAPI* encodes for endoplasmic reticulum aminopeptidase 1, which is a member of the M1 family of zinc metallopeptidases. It regulates MHC class I antigen presentation (York et al. 2002) and associated with autoimmune diseases, such as psoriasis (Sun et al. 2010) and ankylosing spondylitis (Chen et al. 2011), in the recent GWA meta-analyses. However, *ERAPI* transcript was not noticed in the microarray data, thus it remains to be determined whether *ERAPI* is functional in osteoblast proliferation and differentiation.

Since 2007, a number of genome wide association studies for BMD variation have been published (Rivadeneira et al. 2009; Styrkarsdottir et al. 2008, 2009) with a number of well-known candidate genes and novel candidate genes confirmed or identified. Nevertheless, none of the genes suggested by the current study reached genome-wide significant level in these studies. This is not surprising as the ultraconservative Bonferroni correction inflates the false-negative rate (Cheung et al. 2010). Statistical power may also be further reduced by allelic heterogeneity in different ethnic groups although gene-based association study is an approach little influenced by allelic

heterogeneity (Cheung et al. 2008). As such, this study had a number of advantages: first, we used family-based genetic data and successfully replicated and localized a common BMD QTL that has been implicated in Northern Chinese and meta-analysis in a relatively small genetic region. This is particularly important as the number of genes tested was reduced by the linkage analysis with high density microsatellite markers. Additionally, QTL-wide gene-based association study and meta-analysis identified two novel potential BMD genes in Southern Chinese and Northern Europeans. The function of one of these genes, *CAST*, in bone metabolism was further supported by an mRNA expression study suggesting this gene is a negative modulator of osteoblast differentiation. The current study involved the combined use of linkage, genetic association analysis and RNA expression data to account for potential variation within a given phenotype. Successful genetic association studies result in the localization of a gene or genetic region involved with the phenotype being investigated. Thus, various sources of independent data can be used to augment evidence for the involvement of a particular gene and improve assumptions regarding causal associations. There were nonetheless some limitations in this study. Our linkage cohort had sufficient power to detect QTL with large effect size, but not QTL with small effect size (Cheung et al. 2006; Huang et al. 2006). We only have replication data from the European population, despite the fact that the sample size is large. Moreover, a significant genetic heterogeneity was detected between two studies. The genetic heterogeneity could be contributed by the difference in study design, such as inclusion of both men and women in the Northern European study, but only women in the Southern Chinese study; random cohort design in the Northern European, but threshold defined case–control design in the Southern Chinese study. The p values of *CAST* and *ERAPI* obtained in HKSC study did not pass Bonferroni correction, and the association signals are relatively marginal in all analyses. Successful in silico replication in Northern European and in silico expression analysis suggested that these two genes are less likely to be spurious findings although cautious interpretation is needed. Future replications and/or functional analysis will be required to examine the effect of these two genes on bone metabolism. In addition, this study could not confirm whether *CAST* or *ERAPI* are the QTL genes in the pedigrees, since the association study was performed in a non-family based unrelated cohort. Nevertheless, our comprehensive analysis suggested that they are potentially important in bone metabolism.

In summary, we replicated a QTL on chromosome 5q21–22 that influences femoral neck BMD in HKSC. Through association analysis of 18 genes in this region, we identified significant association between *CAST* and *ERAPI* with

femoral neck BMD variation. Future replication of these candidate genes will be required to validate our findings, while resequencing will be useful to better determine whether common and/or rare genetic variation within these two genes is associated with BMD variation.

Acknowledgments The project is supported by NSFC/GRC Joint Research Scheme N-HKU-715/07; Osteoporosis Research Fund and Matching Grant, The University of Hong Kong. KC Wong Education Foundation.

Open Access This article is distributed under the terms of the Creative Commons Attribution Noncommercial License which permits any noncommercial use, distribution, and reproduction in any medium, provided the original author(s) and source are credited.

References

- Chen R, Yao L, Meng T, Xu W (2011) The association between seven ERAP1 polymorphisms and ankylosing spondylitis susceptibility: a meta-analysis involving 8,530 cases and 12,449 controls. *Rheumatol Int* 1–6
- Cheung CL, Huang QY, Ng MY, Chan V, Sham PC, Kung AW (2006) Confirmation of linkage to chromosome 1q for spine bone mineral density in southern Chinese. *Hum Genet* 120:354–359
- Cheung CL, Sham PC, Chan V, Paterson AD, Luk KD, Kung AW (2008) Identification of LTBP2 on chromosome 14q as a novel candidate gene for bone mineral density variation and fracture risk association. *J Clin Endocrinol Metab* 93:4448–4455
- Cheung CL, Xiao SM, Kung AW (2010) Genetic epidemiology of age-related osteoporosis and its clinical applications. *Nat Rev Rheumatol* 6:507–517
- Hsu YH, Xu X, Terwedow HA, Niu T, Hong X, Wu D, Wang L, Brain JD, Bouxsein ML, Cummings SR, Rosen CJ (2007) Large-scale genome-wide linkage analysis for loci linked to BMD at different skeletal sites in extreme selected sibships. *J Bone Miner Res* 22:184–194
- Huang QY, Ng MY, Cheung CL, Chan V, Sham PC, Kung AW (2006) Identification of two sex-specific quantitative trait loci in chromosome 11q for hip bone mineral density in Chinese. *Hum Hered* 61:237–243
- Ioannidis JP, Ng MY, Sham PC, Zintzaras E, Lewis CM, Deng HW, Econs MJ, Karasik D, Devoto M, Kammerer CM 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
- Lander E, Kruglyak L (1995) Genetic dissection of complex traits: guidelines for interpreting and reporting linkage results. *Nat Genet* 11:241–247
- Li HY, Kung WC, Huang QY (2010). Bone mineral density is linked to 1p36 and 7p15-13 in a southern Chinese population. *J Bone Miner Metab*
- Liu JZ, McRae AF, Nyholt DR, Medland SE, Wray NR, Brown KM, Hayward NK, Montgomery GW, Visscher PM, Martin NG, Macgregor S (2010) A versatile gene-based test for genome-wide association studies. *Am J Hum Genet* 87:139–145
- Murray SS, Grisanti MS, Bentley GV, Kahn AJ, Urist MR, Murray EJ (1997) The calpain-calpastatin system and cellular proliferation and differentiation in rodent osteoblastic cells. *Exp Cell Res* 233:297–309
- Ng MY, Sham PC, Paterson AD, Chan V, Kung AW (2006) Effect of environmental factors and gender on the heritability of bone mineral density and bone size. *Ann Hum Genet* 70:428–438
- Niu T, Chen C, Cordell H, Yang J, Wang B, Wang Z, Fang Z, Schork NJ, Rosen CJ, Xu X (1999) A genome-wide scan for loci linked to forearm bone mineral density. *Hum Genet* 104:226–233
- Rivadeneira F, Styrkarsdottir U, Estrada K, Halldorsson BV, Hsu YH, Richards JB, Zillikens MC, Kavvoura FK, Amin N, Aulchenko YS et al (2009) Twenty bone-mineral-density loci identified by large-scale meta-analysis of genome-wide association studies. *Nat Genet* 41:1199–1206
- Sham PC, Purcell S, Cherny SS, Abecasis GR (2002) Powerful regression-based quantitative-trait linkage analysis of general pedigrees. *Am J Hum Genet* 71:238–253
- Shimada M, Mahon MJ, Greer PA, Segre GV (2005) The receptor for parathyroid hormone and parathyroid hormone-related peptide is hydrolyzed and its signaling properties are altered by directly binding the calpain small subunit. *Endocrinology* 146:2336–2344
- Styrkarsdottir U, Halldorsson BV, Gretarsdottir S, Gudbjartsson DF, Walters GB, Ingvarsson T, Jonsdottir T, Saemundsdottir J, Center JR, Nguyen TV et al (2008) Multiple genetic loci for bone mineral density and fractures. *N Engl J Med* 358:2355–2365
- Styrkarsdottir U, Halldorsson BV, Gretarsdottir S, Gudbjartsson DF, Walters GB, Ingvarsson T, Jonsdottir T, Saemundsdottir J, Snorrardottir S, Center JR et al (2009) New sequence variants associated with bone mineral density. *Nat Genet* 41:15–17
- Sun L-D, Cheng H, Wang Z-X, Zhang A-P, Wang P-G, Xu J-H, Zhu Q-X, Zhou H-S, Ellinghaus E, Zhang F-R et al (2010) Association analyses identify six new psoriasis susceptibility loci in the Chinese population. *Nat Genet* 42:1005–1009
- Wigginton JE, Abecasis GR (2005) PEDSTATS: descriptive statistics, graphics and quality assessment for gene mapping data. *Bioinformatics* 21:3445–3447
- Xu S, Yin X, Li S, Jin W, Lou H, Yang L, Gong X, Wang H, Shen Y, Pan X et al (2009) Genomic dissection of population substructure of Han Chinese and its implication in association studies. *Am J Hum Genet* 85:762–774
- York IA, Chang S-C, Saric T, Keys JA, Favreau JM, Goldberg AL, Rock KL (2002) The ER aminopeptidase ERAP1 enhances or limits antigen presentation by trimming epitopes to 8–9 residues. *Nat Immunol* 3:1177–1184
- Zamurovic N, Cappellen D, Rohner D, Susa M (2004) Coordinated activation of notch, Wnt, and transforming growth factor-beta signaling pathways in bone morphogenic protein 2-induced osteogenesis. Notch target gene *Hey1* inhibits mineralization and *Runx2* transcriptional activity. *J Biol Chem* 279:37704–37715