Bone mineral density, osteoporosis, and osteoporotic fractures: a genome-wide association study

Supplementary methods

Correspondence to:
Prof T D Spector, Twin Research and Genetic Epidemiology Unit, St Thomas' Hospital, London SE1 7EH, UK
tim.spector@kcl.ac.uk

Study samples and clinical characteristics

The TwinsUK discovery cohort consisted of members of the TwinsUK cohort, who were previously shown to be representative of singleton populations and the UK population in general.[1] Participants in the Rotterdam cohort were derived from the Rotterdam study (n=7983), a single-centre prospective population-based study of determinants of chronic disabling diseases in elderly people (aged 55 years and older).[2,3] The Chingford study was a population-based study of 1003 individuals who were recruited in 1989 for prospective follow-up for osteoarthritis and osteoporosis. It is listed by the NIH as an important epidemiological resource and one of the few such cohorts with wide-ranging musculoskeletal data.

Measurement of bone mineral density

The bone mineral den sity of all participants in the TwinsUK study was measured at lumbar spine (L1–L4) and femoral neck using dual energy x-ray absorptiometry (QDR 2000W, Hologic, Bedford, MA, USA).[4] For monozygotic twins in TwinsUK, the genotypic information of one individual per sibling pair was included in the analysis. For these monozygotic twins, the mean bone mineral density for the twin pair was used as the phenotype, since monozygotic twins share identical genetic information. If a single twin had missing data, or was excluded, the remaining sibling was regarded as a singleton in the statistical analysis. In the Rotterdam study, bone mineral density measurements of the lumbar spine and right proximal femur used dual energy x-ray absorptiometry (DXA) with a Lunar DPX-L densitometer (Lunar Radiation Corporation, Madison, WI, USA) and were analysed with DPX-IQ version 4.7d software.[5] Methods, quality assurance, accuracy, and precision issues for these DXA measurements have been described previously.[5] In the Chingford study, bone mineral density at the L1–L4 spinal region and femoral neck was measured with DXA on a Hologic QDR 1000 densitometer.

We did not attempt to standardise DXA measurements because we sought replication of genotype dependent differences rather than of absolute bone mineral density values by genotype. However, all bone mineral density measurements were converted to standardised, cohort-specific *Z* scores.

Genotyping and quality control

For the TwinsUK discovery cohort, all samples were typed with the Infinium assay (Illumina, San Diego, USA) with fully compatible SNP arrays, the Hap300 Duo, Hap300, and Hap550. We pooled the normalised intensity data[6] for 2820 Twins UK samples typed at Centre National de Génotypage, Duke University, NC, USA; Helsinki University, Finland; and the Wellcome Trust Sanger Institute, Cambridge, UK, and called genotypes on the basis of the Illluminus algorithm.[7] No calls were assigned if the most likely call was less than a posterior probability of 0·95. Validation of pooling was done by visual inspection of 100 random, shared SNPs for overt batch effects; none were observed.

We excluded 127 participants for whom the sample call rate was more than 95%; 162 for whom autosomal heterozygosity was more than 37% or less than 33%; and 341 for whom genotype concordance with another sample was more than 97% and the sample was of lesser call rate. We also removed misclassified monozygotic twins. Therefore, we removed 508 samples in total.

We excluded 2704 SNPs because $p \leq 1{\cdot}0 \times 10^{-4}$ in test for deviation from Hardy–Weinberg equilibrium; 725 because the minor allele frequency was 1% or less; and 733 because the call rate was 90% or less. We retained 314 075, 98·7% of all available SNPs for analysis, with a resultant call rate of 99·3%. We also visually inspected all intensity cluster plots of SNPs that showed either an association for overdispersion of the clusters, biased no calling, or erroneous genotype assignment. We discarded SNPs with any of these characteristics.

The Rotterdam study samples were assayed with the same Infinium protocol. Intensity files were analysed with the Beadstudio Genotyping Module software v.3.1.14. A no-call threshold of 0·15 was applied to a custom-generated cluster file derived from the Illumina-provided cluster file. In the custom-cluster file 2308 SNPs with Genecall scores of less than 0·90 were visually checked by two observers and manually reclustered or zeroed accordingly. Samples with a low call rate and 10th percentile Genecal score were excluded before we called genotypes.

We excluded 209 samples with a call rate below 98%. 21 had heterozygosity rates above 37% or below 33% across all autosomal SNPs; six had ambiguous estimates of X chromosome inbreeding (homozygosity) ($0{\cdot}2<F<0{\cdot}8$); 36 had mismatch between called and phenotypic gender; 102 had outliers (3 SD) identified by the clustering analysis of identity by state; and 129 had outliers identified by identity-by-state probability of greater than 97%. In total, 706 samples were removed. The SNP quality control applied to the TwinsUK cohort was also applied to the Rotterdam cohort. After exclusions, 535 188 (95·3%) of all available SNPs were available for the replication analysis. Replicated SNPs were tested for Hardy–Weinberg equilibrium. We compared genotype accuracy against 22 Taqman SNPs, and recorded less than 0·3% discrepancy across genotyping methods.

Population stratification

In the TwinsUK study, although population stratification has not been evident in other population-based studies of self-reported white Britons, we assessed for possible stratification with STRUCTURE software.[8] We selected 517 SNPs that were informative for ancestry from the Perlegen database (F_{ST} greater than 0·2 and physical distance greater than 5 Mb). Only one randomly selected twin from each twin pair was analysed with STRUCTURE. Genotype data for unrelated HapMap population (60 from the CEPH group from Utah, 60 Yoruba, and 89 Han Chinese or Japanese) were also included to better estimate ancestral origin for non-white participants. The model allows possible admixture with allele frequencies correlated in three populations. We set so-called burn-ins as 30 000 steps, followed by 100 000 Markov chain Monte Carlo steps. We excluded 14 individuals who lay outside the CEPH cluster with ancestry coefficiency of less than 0·6. We therefore removed all people with self-described ethnicity as "not white British" from the sample. Despite controlling for overt population structure, cryptic relatedness and systematic genotyping errors could produce spuriously significant results.[9]

Although TwinsUK is a sibship-based sample, and therefore could be generally robust to population stratification, some singletons and only one monozygote per monozygote pair were genotyped. Therefore we selected all singletons and randomly selected one individual per sibship and controlled for this possible bias by applying genomic control, and found the genomic inflation factor to be 1·02 or less, based on the median χ^2 statistic for both phenotypes of bone mineral density. Therefore, our sample had little evidence of cryptic relatedness. QQ plots for the TwinsUK genome-wide scan are presented in the figure.

The Rotterdam study data was examined for potential population stratification after excluding outliers detected by the IBS clustering analysis.[10] The genomic control inflation factor[9] for the distribution of test statistic of bone mineral density in the Rotterdam study was 1·002 and 1·049 across all analyses.

Statistical methods

To calculate the false-positive report probability we used results from previous genome-wide association studies, we estimated that one in 5000 SNPs is a true positive.[11,12] SNPs with a false-positive report probability of less than 0·5 were considered for replication in the subsample of the Rotterdam cohort, as suggested by Wacholder and colleagues.[13] The computational power required by the PLINK program (version 1.01) to perform more than 1013 permutations, to control for family structure, was prohibitive. We used the Merlin software package to assess the association between the six SNPs tested for replication in stage two.[14]

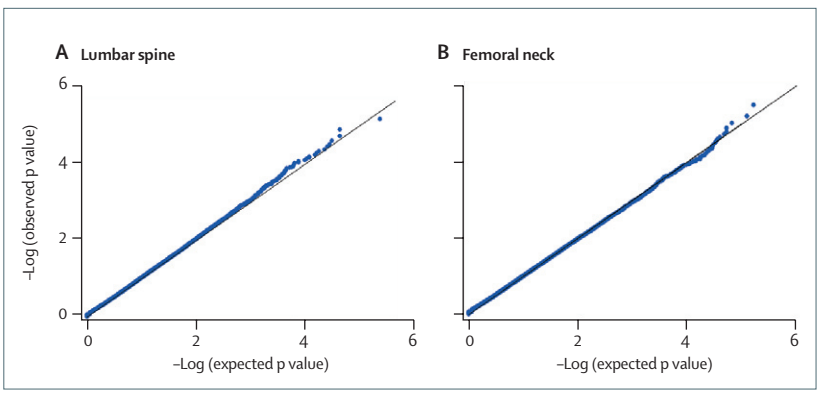

Figure: QQ plots for the TwinsUK discovery cohort

Allelic expression analysis

We used lymphoblastoid cell lines derived from populations of white people, which included 55 unrelated individuals genotyped by the International HapMap Project and 44 lymphoblastoid cell lines derived from an independent panel of unrelated individuals of white European ancestry (both from Coriell Cell Repositories and independent sources). Cells were grown at 37°C and 5% CO_2 in RPMI 1640 medium (Invitrogen, Burlington, Canada) supplemented with 15% heat-inactivated fetal bovine serum (Sigma-Aldricht, Oakville, Canda), 2 mmol/L L-glutamine (Invitrogen, Burlington, Canada) and penicillin/streptomycin (Invitrogen, Burlington, Canada). The cell growth was monitored with a haemocytometer and the cells were harvested when the density reached $0·8\times10^6$–$1·1\times10^6$ cells per mL. Cells were then resuspended and lysed in TRIzol reagent (Invitrogen, Burlington, Canada).

Quantitative sequencing of RT-PCR products for determination of allelic expression

We assessed the extent of allelic imbalance by quantitative sequencing.[15] We assessed RNA quality with an Agilent 2100 Bioanalyzer (Agilent, Palo Alto, USA) before synthesising first strand cDNA with random hexamers (Invitrogen, Burlington, Canada) and Superscript II reverse transcriptase (Invitrogen, Burlington, Canada). For each locus, we designed locus-specific primers to target heteronuclear (unspliced pre-mRNA) at least 50 bp away from the SNP studied. We used PeakPeaker v.2.1[15] with the default settings to quantify the relative amount of the two alleles measured from the chromatogram after peak intensity normalisation.

The normalised heterozygote ratios of genomic DNA samples were used to calculate mean and SD for each SNP. If both heterozygote ratios in two independent RNA samples showed concordant deviation greater than two SDs from the genomic DNA heterozygote ratio mean the sample was called to have allelic imbalance. If one of the two RNA replicates was within two SDs or if the RNA replicates deviated to opposite directions, the sample was

For the **phase information from HapMap** see http://www.hapmap.org

defined as "undetermined". If two or more marker SNPs had available allelic imbalance data for the same sample, the phase information from HapMap allowed assessment of concordance of allelic expression calls: if a site showed relative overexpression of the transcript derived from one chromosome, the alleles on this chromosome were assigned a "+" sign and the alternate alleles were assigned a "–" sign. If discordant calls were made at independent SNPs (n=2), the sample was set to be informative for qualitative allelic expression mapping. The chromosomal alleles neighbouring a test gene that had an assignment of "+" or "–" were used in the association tests; allele counts for each unequivocally phased site were tabulated in a two-by-two table. If the distribution of alleles deviated significantly (with a two-tailed Fisher's exact test) between the "+" and "–" chromosomes, a putative association for allelic expression was determined.[16] The quantitative allelic expression assessment was based on average allele ratio for each sample that was heterozygous for rs435580,[1] and even inconsistent allelic expression calls were included in this analysis. We correlated genotypes from HapMap against expression data by linear expression using expression data from two published studies in HapMap CEU lymphoblastoid cell lines[17,18] to corroborate allelic expression results.

References

1 Andrew T, Hart DJ, Snieder H, de Lange M, Spector TD, MacGregor AJ. Are twins and singletons comparable? A study of disease-related and lifestyle characteristics in adult women. *Twin Res* 2001; **4:** 464–77.

2 Hofman A, Grobbee DE, de Jong PT, van den Ouweland FA. Determinants of disease and disability in the elderly: the Rotterdam Elderly Study. *Eur J Epidemiol* 1991; **7:** 403–22.

3 Hofman A, Breteler MM, van Duijn CM, et al. The Rotterdam Study: objectives and design update. *Eur J Epidemiol* 2007; **22:** 819–29.

4 Hunter DJ, De Lange M, Andrew T, Snieder H, MacGregor AJ, Spector TD. Genetic variation in bone mineral density and calcaneal ultrasound: a study of the influence of menopause using female twins. *Osteoporos Int* 2001; **12:** 406–11.

5 Burger H, van Daele PL, Algra D, et al. The association between age and bone mineral density in men and women aged 55 years and over: the Rotterdam Study. *Bone Miner* 1994; **25:** 1–13.

6 Kermani BG. Artificial intelligence and global normalization methods for genotyping. US Patent 7035740. Washington, DC, USA: US Patent and Trademark Office, 2006.

7 Teo YY, Inouye M, Small KS, et al. A genotype calling algorithm for the Illumina BeadArray platform. *Bioinformatics* 2007; **23:** 2741–46.

8 Pritchard JK, Stephens M, Donnelly P. Inference of population structure using multilocus genotype data. *Genetics* 2000; **155:** 945–59.

9 Devlin B, Roeder K. Genomic control for association studies. *Biometrics* 1999; **55:** 997–1004.

10 Purcell S, Neale B, Todd-Brown K, et al. PLINK: A tool set for whole-genome association and population-based linkage analyses. *Am J Hum Genet* 2007; **81:** 559–75.

11 Mayer B, Erdmann J, Schunkert H. Genetics and heritability of coronary artery disease and myocardial infarction. *Clin Res Cardiol* 2007; **96:** 1–7.

12 Samani NJ, Erdmann J, Hall AS, et al. Genomewide association analysis of coronary artery disease. *N Engl J Med* 2007; **357:** 443–53.

13 Wacholder S, Chanock S, Garcia-Closas M, El Ghormli L, Rothman N. Assessing the probability that a positive report is false: an approach for molecular epidemiology studies. *J Natl Cancer Inst* 2004; **96:** 434–42.

14 Abecasis GR, Cherny SS, Cookson WO, Cardon LR. Merlin-rapid analysis of dense genetic maps using sparse gene flow trees. *Nat Genet* 2002; **30:** 97–101.

15 Ge B, Gurd S, Gaudin T, et al. Survey of allelic expression using EST mining. *Genome Res* 2005; **15:** 1584–91.

16 Pastinen T, Ge B, Gurd S, et al. Mapping common regulatory variants to human haplotypes. *Hum Mol Genet* 2005; **14:** 3963–71.

17 Stranger BE, Nica AC, Forrest MS, et al. Population genomics of human gene expression. *Nat Genet* 2007; **39:** 1217–24.

18 Kwan T, Benovoy D, Dias C, et al. Genome-wide analysis of transcript isoform variation in humans. *Nat Genet* 2008; **40:** 225–31.