ARD Online First, published on September 6, 2012 as 10.1136/annrheumdis-2012-202081

EXTENDED REPORT

Evaluation of the genetic overlap between osteoarthritis with body mass index and height using genome-wide association scan data

Katherine S Elliott,[1*] Kay Chapman,[1,2,*] Aaron Day-Williams,[3] Kalliope Panoutsopoulou,[3] Lorraine Southam,[1,3] Cecilia M Lindgren,[1] the GIANT consortium[4*] Nigel Arden,[5,6] Nadim Aslam,[7] Fraser Birrell,[8,9] Ian Carluke,[9] Andrew Carr,[2,4] Panos Deloukas,[3] Michael Doherty,[10] John Loughlin,[11] Andrew McCaskie,[11,12] William E R Ollier,[13] Ashok Rai,[14] Stuart Ralston,[15] Mike R Reed,[9] Timothy D Spector,[16] Ana M Valdes,[16] Gillian A Wallis,[17] Mark Wilkinson,[18,19] the arcOGEN consortium[4] Eleftheria Zeggini[3]

For numbered affiliations see end of article.

Correspondence to
Eleftheria Zeggini, Wellcome Trust Sanger Institute, Hinxton, Cambridge, CB10 1SA, UK; eleftheria@sanger.ac.uk

Additional supplementary materials are published online only. To view these files please visit the journal online (http://dx.doi.org/10.1136/annrheumdis-2012-202081).

*For GIANT and arcOGEN consortium members see supplementary material (available online only).

Received 28 May 2011
Accepted 19 July 2012

ABSTRACT

Objectives Obesity as measured by body mass index (BMI) is one of the major risk factors for osteoarthritis. In addition, genetic overlap has been reported between osteoarthritis and normal adult height variation. We investigated whether this relationship is due to a shared genetic aetiology on a genome-wide scale.

Methods We compared genetic association summary statistics (effect size, p value) for BMI and height from the GIANT consortium genome-wide association study (GWAS) with genetic association summary statistics from the arcOGEN consortium osteoarthritis GWAS. Significance was evaluated by permutation. Replication of osteoarthritis association of the highlighted signals was investigated in an independent dataset. Phenotypic information of height and BMI was accounted for in a separate analysis using osteoarthritis-free controls.

Results We found significant overlap between osteoarthritis and height ($p=3.3\times10^{-5}$ for signals with $p\leq0.05$) when the GIANT and arcOGEN GWAS were compared. For signals with $p\leq0.001$ we found 17 shared signals between osteoarthritis and height and four between osteoarthritis and BMI. However, only one of the height or BMI signals that had shown evidence of association with osteoarthritis in the arcOGEN GWAS was also associated with osteoarthritis in the independent dataset: rs12149832, within the *FTO* gene (combined $p=2.3\times10^{-5}$). As expected, this signal was attenuated when we adjusted for BMI.

Conclusions We found a significant excess of shared signals between both osteoarthritis and height and osteoarthritis and BMI, suggestive of a common genetic aetiology. However, only one signal showed association with osteoarthritis when followed up in a new dataset.

INTRODUCTION

Osteoarthritis is a common complex disease of synovial joints characterised by degeneration of hyaline cartilage and bone remodelling, usually affecting middle-aged to elderly individuals. It is a leading cause of pain and chronic disability worldwide.[1] Comorbidities such as obesity are frequently observed with osteoarthritis and epidemiological studies have noted a link between osteoarthritis and obesity as measured by body mass index (BMI). In particular, reports show a consistent relationship between overweight measures and knee osteoarthritis.[2] Some population studies have demonstrated that the weight of individuals at age 37 years (median) could predict the onset of knee osteoarthritis 36 years later.[1] In addition, a decrease in BMI of two units over the 10 years preceding diagnosis can reduce the odds of knee osteoarthritis.[3] A Norwegian population-based study of approximately 265 000 individuals concluded that the risk of developing hip osteoarthritis was dependent on the age at which weight gain was most dramatic. Younger adults (<20 years) are at greater risk compared with older individuals (>30 years).[4] In a large prospective population-based cohort from Iceland the incidence of clinically severe osteoarthritis (as indicated by arthroplasty), in relation to measures of overweight, found that 36% and 50% of those with hip and knee osteoarthritis, respectively, had a BMI greater than 30.[5] This is compared with a national prevalence of 17% of the adult population (the International Obesity Task force, http://www.iotf.org/). Furthermore, the Chingford Study has demonstrated that in middle-aged women a one unit increase in BMI is associated with a 10% increased risk of total knee replacement in the following 19 years.[6] There have been a number of large-scale genome-wide association studies (GWAS) for obesity and/or BMI, establishing several genetic loci at genome-wide significance association levels.[7–13] Based on the well-established epidemiological link, we hypothesise here that osteoarthritis and obesity may have a shared genetic background.[14]

Variation in human adult height is also highly heritable. Numerous studies have identified over 180 loci to be associated with the trait.[15–18] There is weak or conflicting evidence for shared genetic determinants between osteoarthritis and height. However, for example, the functional single nucleotide polymorphism (SNP) (rs143383, T/C) in the 5′ untranslated region of the *GDF5* gene, previously observed to be significantly associated

with osteoarthritis in Asian and European cohorts, is also significantly associated with normal variation in human height.[19–23] In addition, height itself has been reported to be a risk factor for non-generalised severe hip osteoarthritis even after adjusting for age, gender and BMI.[24]

The aim of this study was to carry out an investigation of the genetic overlap between osteoarthritis and the two traits of BMI and height by examining the overlap of SNPs association signals across the genome. This may uncover possible common mechanistic pathways.

MATERIALS AND METHODS

Description of datasets

Genome-wide summary statistics (effect size, p values) for BMI and height from the Genetic Investigation of Anthropometric Traits (GIANT) consortium GWAS were compared with genome-wide osteoarthritis data from the arcOGEN consortium. The GIANT consortium has brought together GWAS data from 46 studies.[25 26] Overlap analysis with osteoarthritis utilised 2 400 344 SNPs and 32 387 individuals from the BMI dataset and 2 834 208 SNPs and 133 653 individuals from the height dataset.

The arcOGEN GWAS was carried out in two stages and includes a total of 7567 osteoarthritis cases from the UK (ascertained by radiographic evidence of disease, Kellgren–Lawrence score ≥2, or clinical evidence of the disease to a level requiring total joint replacement) genotyped on the Illumina HumanHap 610-Quad panel. Stage 1 of the arcOGEN GWAS was employed in the main overlap analysis and included 3177 osteoarthritis cases and 4894 population-based controls from the UK (WTCCC2).[27] Genotypes of 17 SNPs that were imputed in arcOGEN stage 1 were validated by direct typing using Sequenom in the stage 1 cases (n=2949) and examining concordance. Replication of association with osteoarthritis for the signals highlighted from the overlap analysis (tables 1 and 2) was carried out using 4324 stage 2 cases from the arcOGEN Consortium and 6518 population-based controls (from the WTCCC2, T1DGC, ALSPAC study and PoBI studies) (see supplementary methods, available online only). SNPs that were not genotyped in the stage 2 arcOGEN GWAS or did not pass quality control were genotyped with Sequenom in 5165 cases and 6115 controls (see supplementary methods, available online only).

Analysis accounting for phenotypic information of height and BMI was carried out using 1671 unrelated female individuals from the osteoarthritis-free TwinsUK cohort as a control set. This cohort is ascertained to study the heritability of age-related diseases and contains full phenotypic information for osteoarthritis status as well as height and BMI.[24] Additional quality control that was performed for this study is described in the supplementary methods (available online only). arcOGEN stage 1 female cases (n=1009 for height; n=1358 for BMI) were utilised for these analyses.

Osteoarthritis replication genotyping

Osteoarthritis association signals at directly typed variants, highlighted in tables 3 and 4, were followed up through in-silico replication using stage 2 arcOGEN GWAS data (see supplementary methods, available online only). Association signals at imputed variants were followed up by carrying out de-novo genotyping in 5165 arcOGEN stage 2 cases and 6115 WTCCC2 controls using the Sequenom MassArray iPLex Gold assay at the Wellcome Trust Sanger Institute. Genotypes were assigned using the MassArray TyperAnalyser software V.4.0 (Sequenom). All genotypes were confirmed manually and passed standard quality control checks (see supplementary methods, available online only).

Analysis strategy

We carried out pairwise comparisons between osteoarthritis and height and between osteoarthritis and BMI genome-wide summary statistics. For each comparison, we focused on the intersection of SNPs for which summary statistics were present in both GWAS. We then sorted these SNPs based on association p value for osteoarthritis. This list of SNPs was then thinned to an independent unlinked set using an r^2 threshold of 0.05 based on HapMap CEU release #27. Starting with the first SNP in the list, any subsequent SNP with $r^2>0.05$ was removed and then the next available SNP was taken. This

Table 1 Shared genetic determinants ($p \leq 1.0\times10^{-3}$) between osteoarthritis and height

SNP	Chromosome	Allele	Osteoarthritis p value	Osteoarthritis OR	95% CI	Height p value	Height OR	95% CI	Nearest gene
rs2744718	1	T	8.9×10^{-5}	1.20	1.10 to 1.31	3.5×10^{-5}	0.98	0.97 to 0.99	*WNT4*
rs6670486	1	T	9.6×10^{-5}	1.16	1.07 to 1.24	6.5×10^{-7}	0.98	0.98 to 0.98	*COL11A1*
rs4833772	4	G	2.5×10^{-4}	1.12	1.05 to 1.19	1.3×10^{-5}	1.02	1.01 to 1.02	*TMEM155*
rs572004	6	G	2.4×10^{-4}	1.14	1.06 to 1.22	7.5×10^{-5}	0.98	0.97 to 0.99	*EYA4*
rs3822856	6	A	9.2×10^{-4}	1.12	1.05 to 1.19	6.5×10^{-4}	0.99	0.98 to 0.99	*NT5DC1*
rs1635853	7	T	2.0×10^{-5}	1.15	1.08 to 1.22	8.8×10^{-21}	1.04	1.03 to 1.04	*JAZF1*
rs10094727	8	A	3.7×10^{-4}	1.23	1.10 to 1.39	6.0×10^{-5}	0.97	0.06 to 0.98	*MSR1*
rs9657371	8	A	1.3×10^{-4}	1.13	1.06 to 1.21	2.7×10^{-5}	0.98	0.97 to 0.99	*CSMD1*
rs11991139	8	C	6.2×10^{-5}	1.14	1.06 to 1.22	4.3×10^{-4}	0.98	0.98 to 0.99	*BLK*
rs3808880	9	G	7.5×10^{-4}	1.13	1.05 to 1.21	6.8×10^{-4}	1.01	1.01 to 1.02	*ROD1*
rs11198893	10	A	3.8×10^{-4}	1.22	1.09 to 1.36	5.7×10^{-4}	1.02	1.01 to 1.03	*GRK5*
rs7932272	11	A	1.1×10^{-4}	1.30	1.12 to 1.52	2.4×10^{-7}	1.04	1.03 to 1.06	*PACS1*
rs7297051	12	T	3.3×10^{-4}	1.14	1.05 to 1.23	6.0×10^{-5}	1.02	1.01 to 1.03	*PTHLH*
rs10506474	12	C	5.2×10^{-4}	1.15	1.06 to 1.27	4.4×10^{-5}	0.99	0.97 to 0.99	*HMGA2*
rs4793927	17	C	1.4×10^{-4}	1.14	1.06 to 1.22	3.0×10^{-6}	1.02	1.01 to 1.03	*HOXB3*
rs2864419	19	G	8.4×10^{-5}	1.14	1.07 to 1.21	1.1×10^{-12}	1.03	1.02 to 1.04	*DOT1L*
rs8105885	19	T	7.9×10^{-4}	1.20	1.08 to 1.37	4.3×10^{-4}	0.98	0.96 to 0.98	*ZNF98*

SNP, single nucleotide polymorphism.

Table 2 Shared genetic determinants ($p \leq 1.0 \times 10^{-3}$) between osteoarthritis and BMI

SNP	Chromosome	Allele	Osteoarthritis p value	Osteoarthritis OR	95% CI	BMI p value	BMI z score	Nearest gene
rs4856346	3	T	6.1×10^{-4}	1.14	1.05 to 1.22	7.1×10^{-4}	3.48	*GBE1*
rs7828042	8	G	1.1×10^{-4}	1.14	1.07 to 1.22	2.1×10^{-4}	4.05	*SLURP1*
rs7203219	16	T	3.5×10^{-4}	1.19	1.08 to 1.31	3.4×10^{-4}	3.87	*GPR139*
rs12149832	16	A	2.8×10^{-4}	1.12	1.06 to 1.20	1.9×10^{-16}	8.47	*FTO*

BMI, body mass index; SNP, single nucleotide polymorphism.

continued until a set of independent SNPs was obtained (osteoarthritis–BMI n=62 280, osteoarthritis–height n=64 702).

We investigated the distribution of p values above and below given thresholds (0.5, 0.1, 0.05, 0.04, 0.03, 0.02, 0.01, 0.005, 0.001, 0.0005) for each trait. The distribution of counts in the resulting 2×2 contingency tables was analysed using the χ^2 test. A significant excess of signals with p values less than the given threshold for both phenotypes was taken to indicate a concurrence of signals.

In addition, we examined the SNPs for each comparative analysis of osteoarthritis–height and osteoarthritis–BMI to see if there was an overabundance of discordant or concordant risk alleles between the datasets (see supplementary methods, available online only).

Permutations

Based on the results obtained for the analysis of the signal overlap between the osteoarthritis–height and osteoarthritis–BMI comparisons (table 5), we selected p value thresholds of 0.001 (osteoarthritis–BMI) and 0.05 (osteoarthritis–height) for further follow-up. We permuted the p value signals for the entire datasets as well as for linkage disequilibrium (LD)-thinned data of $r^2 = 0.2$ and 0.05. We generated 100 000 permutations of the arcOGEN case–control data using PLINK[28] (make-perm-pheno command) and performed a GWAS for each permutation under the log-additive model. We carried out the overlap analysis for each permuted case–control dataset considering only SNPs that were directly typed in arcOGEN and present in the GIANT data. We thus generated a null distribution of p values. From this we calculated the probability of seeing an overlap p value equal to or less than the original p value for directly typed SNPs.

In addition, we sought to get a more precise empirical p value for the osteoarthritis–height comparison as this gave the most compelling results for the signal overlap analysis (table 5). Using the LD-thinned data ($r^2 = 0.05$) and p value threshold of 0.05, we generated 500 000 000 permutations of the arcOGEN and GIANT height datasets by permuting which p value was associated with which SNP. We performed 500 000 000 overlap analyses by randomly choosing without replacement a permutation from each dataset to generate the null distribution of overlap p values, given a specific distribution of original p values. From this constructed null distribution of p values we calculated the probability of seeing an overlap p value equal to or less than the original overlap p value for the entire dataset.

Replication of osteoarthritis association for overlapping signals

Case–control association analysis under the log-additive model was carried out using PLINK for directly typed SNPs and SNPTEST for imputed SNPs.[28 30] Combined estimates of OR and p values for stages 1 and 2 of arcOGEN were obtained

Table 3 Replication of osteoarthritis association at shared genetic determinants between osteoarthritis and height

		Stage 1			Stage 2			Combined		
SNP	Chromosome	p Value	OR	95% CI	P Value	OR	95% CI	p Value	OR	95% CI
rs2744718†‡	1	8.9×10^{-5}	1.20	1.10 to 1.31	Failed QC	Failed QC				
rs6670486	1	9.6×10^{-5}	1.16	1.07 to 1.24	0.58	1.02	0.96 to 1.08	3.5×10^{-3}	1.07	1.02 to 1.12
rs4833772*	4	2.5×10^{-4}	1.12	1.05 to 1.19	0.24	0.97	0.92 to 1.02	1.6×10^{-1}	1.03	0.99 to 1.07
rs572004	6	2.4×10^{-4}	1.14	1.06 to 1.22	0.29	1.03	0.97 to 1.09	9.0×10^{-4}	1.07	1.03 to 1.14
rs3822856	6	9.2×10^{-4}	1.12	1.05 to 1.19	0.29	1.02	0.97 to 1.09	2.5×10^{-3}	1.06	1.02 to 1.11
rs1635853	7	2.0×10^{-5}	1.15	1.08 to 1.22	0.56	1.02	0.96 to 1.07	4.4×10^{-4}	1.08	1.03 to 1.12
rs10094727‡	8	3.7×10^{-4}	1.23	1.10 to 1.39	0.80	0.99	0.89 to 1.09	3.0×10^{-2}	1.09	1.01 to 1.17
rs9657371‡	8	1.3×10^{-4}	1.13	1.06 to 1.21	0.60	1.02	0.96 to 1.07	3.7×10^{-3}	1.06	1.02 to 1.11
rs11991139*	8	6.2×10^{-5}	1.14	1.06 to 1.22	0.37	1.02	1.0 to 1.09	3.1×10^{-3}	1.05	1.02 to 1.10
rs3808880*	9	7.5×10^{-4}	1.13	1.05 to 1.21	0.03	0.94	0.88 to 0.99	5.6×10^{-1}	1.01	0.97 to 1.06
rs11198893†‡	10	3.8×10^{-4}	1.22	1.09 to 1.36	Failed QC					
rs7932272	11	1.1×10^{-4}	1.30	1.12 to 1.52	0.08	1.10	0.99 to 1.20	5.2×10^{-4}	1.16	1.06 to 1.26
rs7297051	12	3.3×10^{-4}	1.14	1.05 to 1.23	0.15	0.95	0.89 to 1.01	4.1×10^{-1}	1.02	0.97 to 1.07
rs10506474	12	5.2×10^{-4}	1.15	1.06 to 1.27	0.37	0.97	0.90 to 1.04	1.5×10^{-1}	1.04	0.98 to 1.09
rs4793927	17	1.4×10^{-4}	1.14	1.06 to 1.22	0.29	0.97	0.92 to 1.03	1.2×10^{-1}	1.03	0.99 to 1.07
rs2864419‡	19	8.4×10^{-5}	1.14	1.07 to 1.21	0.72	1.01	0.96 to 1.07	4.6×10^{-3}	1.06	1.02 to 1.11
rs8105885	19	7.9×10^{-4}	1.20	1.08 to 1.37	0.15	0.93	0.85 to 1.03	3.2×10^{-1}	1.04	0.96 to 1.12

*Proxies used for analysis due to failure of SNP in stage 2 replication. Proxy for rs4833772 is rs4833233 ($r^2 = 1$), for rs11991139 is rs13280813 ($r^2 = 0.94$) and for rs3808880 is rs13293285 ($r^2 = 0.89$).
†No proxies found for $r^2 > 0.3$ (rs2744718); $r^2 > 0.43$ (rs11198893).
‡Directly typed SNP analysed from arcOGEN genome-wide association scan.
QC, qualtiy control; SNP, single nucleotide polymorphism.

Table 4 Replication of osteoarthritis association at shared genetic determinants between osteoarthritis and body mass index

		Stage 1			Stage 2			Combined		
SNP	Chromosome	p Value	OR	95% CI	p Value	OR	95% CI	p Value	OR	95% CI
rs4856346*	3	6.1×10^{-4}	1.14	1.06 to 1.23	0.87	0.99	0.94 to 1.06	0.041	1.05	1.00 to 1.10
rs7828042†	8	1.1×10^{-4}	1.14	1.07 to 1.22	0.45	0.98	0.92 to 1.04	0.050	1.04	1.00 to 1.08
rs7203219†	16	3.5×10^{-4}	1.19	1.08 to 1.31	0.92	1.01	0.93 to 1.09	0.016	1.08	1.01 to 1.15
rs12149832*	16	2.8×10^{-4}	1.12	1.06 to 1.20	0.009	1.07	1.02 to 1.14	2.3×10^{-5}	1.10	1.05 to 1.15

*Proxies ($r^2 > 0.93$) were used for analysis due to failure of SNP in stage 2 replication. Proxy for rs4856346 is rs898763 ($r^2 = 1$), for rs12149832 is rs8050136 ($r^2 = 0.93$).
†Directly typed SNP from arcOGEN genome-wide association scan.
SNP, single nucleotide polymorphism.

using fixed-effect meta-analyses in GWAMA.[29] Heterogeneity was checked using the Cochran's Q and I^2 statistics.

Analysis with adjustment for BMI and height

In order to adjust for height and BMI as covariates, we repeated the case–control analysis for stage 1 of the arcOGEN dataset using the TwinsUK cohort as controls. Analysis was carried out using PLINK when the SNPs were directly typed or SNPTEST[30] when they were imputed. The analysis was performed twice; with and without an adjustment for height and BMI.

RESULTS

Overlapping signals and permutations

Our findings suggest an excess of shared signals both between osteoarthritis and height and osteoarthritis and BMI. A comparison of signals indicates an excess of sharing at p value thresholds of 0.05, 0.04, 0.03, 0.02 and 0.01 for osteoarthritis and height; there is evidence for overlap between osteoarthritis and BMI at the p value thresholds of 0.005, 0.001 and 0.0005 (table 5).

To test the strength of the observed overlap we ran a series of permutations (table 6). The LD-thinned datasets provide the most robust results because the probability of seeing an overlap p value equal to or less than the original analysis p value by chance is unlikely. Both comparisons of osteoarthritis–height and osteoarthritis–BMI showed an excess of overlapping signals ($p=1.4\times10^{-3}$ for the 0.05 p value threshold and $p=2.28\times10^{-2}$ for the 0.001 p value threshold). Based on these results we performed 500 000 000 permutations of the entire LD-thinned ($r^2 = 0.05$) height dataset. This showed that the probability of seeing a p value less than or equal to 3.04×10^{-5} was 3.3×10^{-5}. A total of 17 SNPs with p values of 1.0×10^{-3} or less were shared between osteoarthritis and height (table 1), while four SNPs were shared between osteoarthritis and BMI (table 2). These SNPs were distributed throughout the genome (11 chromosomes in the osteoarthritis–height comparison and three chromosomes in the osteoarthritis–BMI comparison). Some of these signals, such as the ones near *COL11A1*, *PTHLH* and *FTO* are well-known loci with established associations with bone development, bone mineral density and obesity, respectively.[8 19 31]

Table 5 Analysis of shared excess signals between osteoarthritis and normal height variation and osteoarthritis and BMI

	Osteoarthritis–normal adult height comparison		Osteoarthritis–BMI comparison	
Total no of SNPs	62280		64702	
Signal definition (p value)	Overlapping SNPs (n)	p Value for overlap	Overlapping SNPs (n)	p Value for overlap
0.5	28194	0.5189	23526	0.4299
0.1	3571	0.0026	1867	0.5814
0.05	1491	3.04×10^{-5}	584	0.7483
0.04	1160	1.00×10^{-6}	412	0.6315
0.03	814	1.00×10^{-6}	267	0.1715
0.02	511	1.00×10^{-16}	143	0.0368
0.01	213	6.20×10^{-4}	50	0.0356
0.005	92	0.0173	20	0.0057
0.001	17	0.5999	4	1.2×10^{-5}
0.0005	8	0.1278	3	4.8×10^{-17}

BMI, body mass index; SNP, single nucleotide polymorphism.

Replication of osteoarthritis association

To evaluate the observed overlap further we attempted to replicate osteoarthritis association of these 21 signals employing stage 2 of the arcOGEN dataset (tables 3 and 4). Seven of the SNP failed quality control in stage 2 and proxies ($r^2>0.85$) were sought. No proxies were found for two of the seven SNPs, rs2744718 and rs11198893. Of the 19 SNPs successfully taken forward for validation, rs12149832 on chromosome 16 within the *FTO* gene was the only one found to be associated ($p<0.01$) with osteoarthritis in the replication dataset ($p=0.009$, in the same direction). The combined p value of both stages increased in significance for this SNP relative to stage 1 alone ($p=2.8\times10^{-4}$ for stage 1 vs $p=2.3\times10^{-5}$ for stages 1 and 2 combined, table 4).

Adjustment for BMI and height

Adjustment for height and BMI (tables 7 and 8) only affected the signal at the *FTO* SNP rs12149832. Here we found an eight-fold increase in the p value after adjustment for BMI ($p=0.22576$) compared with the unadjusted result ($p=0.029219$).

DISCUSSION

Identification of the genetic loci contributing to variation in quantitative traits such as height and BMI, and risk of osteoarthritis could help elucidate possible mechanistic pathways. There is an established genetic link between height and osteoarthritis. The pleiotropic action of *GDF5* on human height is an example that may shed light on shared signalling functions and pathways affecting the two traits.[19] Epidemiological evidence has also suggested a link between osteoarthritis and BMI.[32] It is plausible that these traits also share genetic associations and we carried out a SNP-by-SNP pairwise comparison of GWAS data to investigate their genetic overlap.

We obtained evidence for overlap of association signals between osteoarthritis and height and between osteoarthritis and BMI at different definition thresholds, corroborated by permutation analyses to obtain empirical p values. We investigated specific signals

Table 6 Permutation results for osteoarthritis–height (p value threshold 0.05) and osteoarthritis–BMI overlap (p value threshold 0.001)

	Osteoarthritis–adult height comparison			Osteoarthritis–BMI comparison		
	All data	$r^2 = 0.2$	$r^2 = 0.05$	All data	$r^2 = 0.2$	$r^2 = 0.05$
No of SNPs†	489098	63802	27728	461707	55275	24599
p Value*	0	2.62×10^{-7}	3.04×10^{-5}	0	6.41×10^{-5}	1.24×10^{-5}
Tested p value**	3.30×10^{-6}	5.33×10^{-6}	1.10×10^{-3}	6.19×10^{-11}	2.91×10^{-4}	2.25×10^{-4}
No of permutations	100000	100000	100000	100000	100000	100000
Permutation p value	0.0128	$<1\times10^{-5}$	0.0014	0.00374	0.04716	0.02288

†Directly typed only analysed.
*p Value of original analysis.
**p Value of the overlap for directly typed SNPs only. This is the p value that the permutation analysis is tested against.
BMI, body mass index; SNPs, single nucleotide polymorphism.

that may be representative of these findings and looked at all SNPs with $p \leq 1.0\times10^{-3}$ for both comparisons. Some signals reside in the vicinity of genes, such as the structural protein collagen gene *COL11A1* and the parathyroid hormone-related protein *PTHLH* that regulates endochondral bone development, which have previously been identified as possible candidates for osteoarthritis susceptibility.[33–36] For the osteoarthritis–BMI comparison the *FTO* gene for obesity was highlighted.

Using a second dataset we attempted to replicate the osteoarthritis association of overlapping signals. The fact that the *FTO* locus was the only one to replicate in our second osteoarthritis dataset suggests that the other signals may have been false positive signals for osteoarthritis, or low power in the replication cohort. Adjustment for BMI attenuated this osteoarthritis signal, indicating that the primary association is with BMI.

The established osteoarthritis and height overlapping signal rs143383 located in *GDF5* was not identified in this analysis. We found it to be strongly associated with height ($p=1.94\times10^{-50}$), but not associated with osteoarthritis in the arcOGEN dataset (p=0.602). Although association between the *GDF5* locus and hip and knee osteoarthritis was first reported in a study of Japanese and Chinese individuals in 2007,[20] it took several years and large-scale meta-analysis efforts to replicate the association robustly in European populations.[21–23] In addition to allele frequency disparities between ethnic groups, this observation also highlights the limited power (<10% at $\alpha=5\times10^{-8}$) of a dataset such as arcOGEN (comprising 3177 cases and 4894 controls) to detect a signal with modest effect (OR 1.15) and common risk allele frequency (~0.60 for the *GDF5* signal).[27] Our results should be interpreted within the power constraints of our study. First, osteoarthritis is a heterogeneous disease and the definition of the cases here was primarily based on painful rather than structural osteoarthritis. Second, we examined GWAS platform SNP content rather than known causal variants. Finally, the osteoarthritis GWAS used population-based controls, which can dilute power due to misclassifications of cases as controls in a common disease such as osteoarthritis.

In conclusion, our genome-wide comparison of GIANT and arcOGEN generated evidence for an overall excess of overlapping signals between osteoarthritis and the two quantitative traits of BMI and height. The *FTO* signal was robustly associated with BMI and osteoarthritis, and showed evidence of association in the replication osteoarthritis dataset. This signal underpins the known epidemiological link between BMI and osteoarthritis, and represents the single largest genetic effect for BMI, which may have facilitated its identification as a shared locus. Better-powered GWAS datasets, along with large-scale replication samples, will help unveil additional shared loci and highlight common biological pathways.

Table 7 Results of osteoarthritis association analysis adjusting for height

			Unadjusted			Adjusted		
SNP	Chromosome	Allele	OR	95% CI	p Value	OR	95% CI	p Value
rs2744718	1	T	0.94	0.80 to 1.10	0.437	0.94	0.80 to 1.10	0.442
rs6670486	1	T	1.19	1.05 to 1.36	0.007	1.19	1.05 to 1.36	0.006
rs4833772	4	G	1.01	0.90 to 1.13	0.877	1.01	0.90 to 1.13	0.945
rs572004	6	G	1.08	0.96 to 1.22	0.201	1.08	0.96 to 1.22	0.210
rs3822656	6	A	1.02	0.91 to 1.14	0.694	1.02	0.91 to 1.14	0.678
rs1635853	7	T	1.17	1.04 to 1.31	0.008	1.17	1.04 to 1.31	0.007
rs10094727	8	A	0.87	0.71 to 1.10	0.161	0.87	0.71 to 1.06	0.186
rs9657371	8	A	0.92	0.82 to 1.03	0.168	0.93	0.83 to 1.04	0.158
rs11991139	8	C	1.03	0.92 to 1.15	0.601	1.03	0.92 to 1.15	0.491
rs3808880	9	G	0.99	0.88 to 1.12	0.946	0.99	0.88 to 1.12	0.980
rs11198893	10	A	1.22	1.02 to 1.45	0.025	1.25	1.05 to 1.49	0.013
rs7932272	11	A	1.12	0.85 to 1.37	0.195	1.12	0.85 to 1.37	0.174
rs7297051	12	T			Failed QC			
rs10506474	12	C	1.22	1.06 to 1.398	0.006	1.22	1.06 to 1.40	0.009
rs4793927	17	C	1.13	1.01 to 1.26	0.037	1.13	1.01 to 1.26	0.028
rs2864419	19	G			Failed QC			
rs8105885	19	T	1.16	0.95 to 1.42	0.149	1.16	0.95 to 1.42	0.154

QC, qualtiy control; SNP, single nucleotide polymorphism.

Table 8 Results of osteoarthritis association analysis adjusting for BMI

			Unadjusted			Adjusted		
SNP	Chromosome	Allele	OR	95% CI	p Value	OR	95% CI	p Value
rs4856346	3	T	1.06	0.94 to 1.19	0.380	1.06	0.94 to 1.19	0.614
rs7828042	8	G	0.85	0.77 to 0.95	3.5×10^{-3}	0.87	0.77 to 0.97	1.0×10^{-2}
rs7203219	16	T	1.03	0.88 to 1.21	0.667	1.04	0.89 to 1.23	0.591
rs12149832	16	A	1.12	1.01 to 1.24	2.9×10^{-2}	1.12	1.01 to 1.24	0.226

BMI, body mass index; SNP, single nucleotide polymorphism.

Author affiliations
[1]Wellcome Trust Centre for Human Genetics, University of Oxford, Oxford, UK
[2]Nuffield Department of Orthopaedics, Rheumatology and Musculoskeletal Sciences, University of Oxford, Oxford, UK
[3]Wellcome Trust Sanger Institute, Hinxton, Cambridge, UK
[4]NIHR Biomedical Research Unit, University of Oxford, Oxford, UK
[5]MRC Epidemiology Resource Centre, University of Southampton, Southampton, UK
[6]Orthopaedic Department, Worcestershire Acute Hospitals NHS Trust, Worcester, UK
[7]Musculoskeletal Research Group, Newcastle University, Institute of Cellular Medicine, The Medical School, Newcastle upon Tyne, UK
[8]Northumbria Healthcare NHS Foundation Trust, Wansbeck General Hospital, Ashington, UK
[9]Academic Rheumatology, University of Nottingham, Nottingham, UK
[10]Institute of Cellular Medicine, Musculoskeletal Research Group, Newcastle University, Newcastle upon Tyne, UK
[11]The Newcastle upon Tyne Hospitals NHS Trust Foundation Trust, The Freeman Hospital, Newcastle upon Tyne, UK
[12]Centre for Integrated Genomic Medical Research, University of Manchester, Manchester, UK
[13]Rheumatology Department, Worcestershire Acute Hospitals NHS Trust, Worcester, UK
[14]Institute of Genetics and Molecular Medicine, University of Edinburgh, Edinburgh, UK
[15]Department of Twin Research and Genetic Epidemiology, King's College London, London, UK
[16]Wellcome Trust Centre for Cell Matrix Research, University of Manchester, Manchester, UK
[17]Academic Unit of Bone Metabolism, Department of Human Metabolism, University of Sheffield, Sheffield, UK
[18]Sheffield NIHR Bone Biomedical Research Unit, Centre for Biomedical Research, Northern General Hospital, Sheffield, UK
[19]xxxxx

Funding arcOGEN (http://www.arcogen.org.uk/) is funded by a special purpose grant from Arthritis Research UK (grant 18030). This study was funded by The Wellcome Trust (098051 and WT079557MA), Arthritis Research UK, The Wishbone Trust, The Collisson Foundation, The Botnar Foundation, The Lord Nuffield Orthopaedic Trust, The Jean Shanks foundation, NIHR Musculoskeletal Biomedical Research Unit and EU Fp7 Treat-OA. For Newcastle we acknowledge the support of the UK NIHR BRC for Ageing and Age-related disease award to the Newcastle upon Tyne Hospitals NHS Foundation Trust and the support of the Northumberland, Tyne and Wear CLRN. KC is a Botnar Research Fellow.

Contributors Data generation: KSE, KC, ADW, KP, LS, CML; GIANT: NKA, AC, PD, MD, JL, AM, WERO, SR, TDS, AMV, GAW, JMW, AR, NA, FB, MRR, IC; arcOGEN: EZ. Data analysis: KSE, KC, ADW, KP, EZ. Manuscript preparation: KSE, KC, KP, ADW, EZ. Critical reading and approval of manuscript: KSE, KC, ADW, KP, LS, CML, NKA, AC, PD, MD, JL, AM, WERO, SR, TDS, AMV, GAW, JMW, AR, NA, FB, MRR, IC, EZ. The first two authors contributed equally.

Competing interests None.

Ethics approval Ethics approval for this study was obtained from National Research Ethics Service (REC ref# 07/H0606/150).

Provenance and peer review Not commissioned; externally peer reviewed.

REFERENCES

1. **Creamer P,** Hochberg MC. Osteoarthritis. *Lancet* 1997;**350**:503–8.
2. **Felson DT,** Anderson JJ, Naimark A, *et al*. Obesity and knee osteoarthritis: the Framingham Study. *Ann Intern Med* 1988;**109**:18–24.
3. **Felson DT,** Zhang Y, Anthony JM, *et al*. Weight loss reduces the risk for symptomatic knee osteoarthritis in women: the Framingham study. *Ann Intern Med* 1992;**116**:535–9.
4. **Apold H,** Meyer HE, Espehaug B, *et al*. Weight gain and the risk of total hip replacement a population-based prospective cohort study of 265,725 individuals. *Osteoarthritis Cartilage* 2011;**19**:809–15.
5. **Lohmander LS,** Gerhardsson M, Rollof J, *et al*. Incidence of severe knee and hip osteoarthritis in relation to different measures of body mass. A population-based prospective cohort study. *Ann Rheum Dis* 2008;**68**:490–6.
6. **Goulston LM,** Kiran A, Javaid MK, *et al*. Does obesity predict knee pain over fourteen years in women, independently of radiographic changes? *Athritis Car Res* 2011;**63**:1398–406.
7. **Loos RJF,** Lindgren CM, Li S, *et al*. Common variants near *MC4R* are associated with fat mass, weight and risk of obesity. *Nat Genet* 2008;**40**:768–75.
8. **Frayling TM,** Timpson NJ, Weedon MN, *et al*. A common variant in the *FTO* gene is associated with body mass index and predisposes to childhood and adult obesity. *Science* 2007;**316**:889–94.
9. **Scuteri A,** Sanna S, Chen WM, *et al*. Genome-wide association scans shows genetic variants in the *FTO* gene are associated with obesity-related traits. *PLOS Genet* 2007;**3**:e115.
10. **Liu YJ,** Liu XG, Wang L, *et al*. Genome-wide association scans identified CTNNBL1 as a novel gene for obesity. *Hum Mol Gen* 2008;**17**:1803–13.
11. **Willer CJ,** Speliotes EK, Loos RJ, *et al*. Six new loci associated with body mass index highlight a neuronal influence on body weight regulation. *Nat Gen* 2009;**41**:25–34.
12. **Thorleifsson G,** Walters GB, Gudbjartsson DF, *et al*. Genome-wide association yields newq sequence variants at seven loci that associate with measures of obesity. *Nat Gen* 2009;**41**:18–24.
13. **Speliotes E,** Willer CJ, Berndt SI, *et al*. Association analyses of 249,796 individuals reveal 18 new loci associated with body mass index. *Nat Gen* 2010;**42**:937–48.
14. **Spector TD,** Reneland RH, Mah S, *et al*. Association between a variation in LRCH1 and knee osteoarthritis: a genome-wide single-nucleotide polymorphism association study using DNA pooling. *Arthritis Rheum* 2006;**54**:524–32.
15. **Lango Allen H,** Estrada K, Lettre G, *et al*. Hundreds of variants clustered in genomic loci and biological pathways affect human height. *Nature* 2010;**467**:832–8.
16. **Weedon MN,** Lettre G, Freathy RM, *et al*. A common variant of HMGA2 is associated with adult and childhood height in the general population. *Nat Gen* 2007;**39**:1245–50.
17. **Lettre G,** Jackson AU, Gieger C, *et al*. Identification of ten loci associated with height highlights new biological pathways in human growth. *Nat Genet* 2008;**40**:584–91.
18. **Weedon MN,** Lango H, Lindgren CM, *et al*. Genome-wide association analysis identifies 20 loci that influence height. *Nat Genet* 2008;**40**:575–83.
19. **Sanna S,** Jackson AU, Nagaraja R, *et al*. Common variants in the GDF5-UQCC region are associated with variation in human height. *Nat Genet* 2008;**40**:198–203.
20. **Miyamoto Y,** Mabuchi A, Shi D, *et al*. Functional polymorphism in the 5′ UTR of *GDF5* is associated with susceptibility to osteoarthritis. *Nat Genet* 2007;**39**:529–33.
21. **Chapman K,** Takahashi A, Meulenbelt I, *et al*. A meta-analysis of European and Asian cohorts reveals a global role of a functional SNP in the 5′ UTR of *GDF5* with osteoarthritis susceptibility. *Hum Mol Genet* 2008;**17**:1497–504.
22. **Evangelou E,** Chapman K, Meulenbelt I, *et al*. Large-scale analysis of association between GDF5 and FRZB variants and osteoarthritis of the hip, knee, and hand. *Arthritis Rheum* 2009;**60**:1710–21.
23. **Valdes AM,** Evangelou E, Kerkhof HJ, *et al*. The GDF5 rs143383 polymorphism is associated with osteoarthritis of the knee with genome-wide statistical significance. *Ann Rheum Dis* 2011;**70**:873–5.
24. **Valdes AM,** McWilliams D, Arden NK, *et al*. Involvement of different risk factors in clinically severe large joint osteoarthritis according to the presence of hand interphalangeal nodes. *Arthritis Rheum* 2010;**62**:2688–95.
25. **GIANT consortium.** Membership of the GIANT consortium. http://hmg.oxfordjournals.org/content/suppl/2009/01/21/ddp041.DC1/ddp041supp.pdf (accessed 1 May 2012).
26. **Consortium.** Membership of the GIANT consortium. http://www.helmholtz-muenchen.de/epi/beitraege-zu-netzwerken/giant-genomewide-investigation-of-anthropometric-measures/index.html (accessed 1 May 2012).

27. **Panoutsopoulou K,** Southam L, Elliott KS, *et al*. Insights into the genetic architecture of osteoarthritis from stage 1 of the arcOGEN study. *Ann Rheum Dis*. 2011;**70**:864–7.
28. **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.
29. **Magi R,** Morris AP. GWAMA: software for genome-wide association meta-analysis. *BMC Bioinformatics* 2010;**11**:288.
30. **Marchini J,** Howie B, Myers S, *et al*. A new multipoint method for genome-wide association studies by imputation of genotypes. *Nat Genet* 2007;**39**:906–13.
31. **Estrada K,** Styrkarsdottir U, Evangelou E, *et al*. Genome-wide meta-analysis identifies 56 bone mineral density loci and reveals 14 loci associated with risk of fracture. *Nat Genet* 2012;**44**:491–501.
32. **Hart DJ,** Spector TD. The relationship of obesity, fat distribution and osteoarthritis in women in the general population: the Chingford Study. *J Rheumatol* 1993;**20**:331–5.
33. **Xu L,** Flahiff CM, Waldman BA, *et al*. Osteoarthritis-like changes and decreased mechanical function of articular cartilage in the joints of mice with the chondrodysplasia gene (cho). *Arthritis Rheum* 2003;**48**:2509–18.
34. **Okano K,** Tsukazaki T, Ohtsuru A, *et al*. Expression of parathyroid hormone-related peptide in human osteoarthritis. *J Orthop Res* 1997;**15**:175–80.
35. **Tompson SW,** Bacino CA, Safina NP, *et al*. Fibrochondrogenesis results from mutations in the COL11A1 type XI collagen gene. *Am J Hum Genet* 2010;**87**:708–12.
36. **MacGregor AJ,** Andrew T, Sambrook PN, *et al*. Structural, psychological, and genetic influences on low back and neck pain: a study of adult female twins. *Arthritis Rheum* 2004;**51**:160–7.

Evaluation of the genetic overlap between osteoarthritis with body mass index and height using genome-wide association scan data

Katherine S Elliott, Kay Chapman, Aaron Day-Williams, et al.

Ann Rheum Dis published online September 6, 2012
doi: 10.1136/annrheumdis-2012-202081

Notes