

ARTICLE

Finer linkage mapping of a primary hip osteoarthritis susceptibility locus on chromosome 6

John Loughlin^{*1}, Zehra Mustafa¹, Barbara Dowling¹, Lorraine Southam¹, Lucy Marcelline¹, S Susanna Ränä², Leena Ala-Kokko^{2,3} and Kay Chapman¹

¹University of Oxford, Nuffield Department of Clinical Laboratory Sciences and Institute of Molecular Medicine, Oxford, UK; ²Collagen Research Unit, Biocenter and Department of Medical Biochemistry, University of Oulu, Oulu, Finland; ³Center for Gene Therapy and Department of Medicine, Tulane University Health Sciences Center, New Orleans, USA

Primary osteoarthritis (OA) is a common late-onset disease that exhibits complex genetic transmittance. A previous genome-wide linkage scan of OA affected sibling pair families (ascertained by total joint replacement surgery) identified a region of suggestive linkage on chromosome 6, with a maximum multipoint-LOD score (MLS) of 2.9 in 194 families containing sibling pairs concordant for total hip replacement (THR-families). However, up to 50 cM of the chromosome had a multipoint-LOD score > 2.0, indicating that the susceptibility locus was poorly mapped. We have now genotyped chromosome 6 to a higher density in an expanded cohort of 378 THR-families. We obtained an MLS of 2.8 to an 11.4 cM interval defined by markers D6S452 and 509-8B2, which map between 70.5 to 81.9 cM from the 6p-telomere. Stratification by gender revealed that this linkage was completely accounted for by female THR-families ($n=146$), with an MLS of 4.0 and with the highest two-point LOD score being 4.6 for marker D6S1573 (75.9 cM). The 11.4 cM interval just encompasses the candidate gene COL9A1 (81.9 cM). We identified and then genotyped twenty common single nucleotide polymorphisms (SNPs) from within COL9A1 in the 146 probands from our female THR-families and in 215 age-matched female controls. No SNP allele, genotype or haplotype demonstrated association to disease. Overall, we have narrowed the chromosome 6 OA susceptibility locus to a point at which linkage disequilibrium/association analysis is feasible, we have demonstrated that this locus is female specific, and found no evidence that COL9A1 encodes for the susceptibility.

European Journal of Human Genetics (2002) 10, 562–568. doi:10.1038/sj.ejhg.5200848

Keywords: osteoarthritis; linkage; association; chromosome 6; COL9A1

Introduction

Primary osteoarthritis (OA, MIM 165720) is a common late-onset arthritis characterised by degeneration of the hyaline cartilage of synovial joints. Twin-pair, sibling risk and segregation studies have demonstrated a significant genetic component to OA, which is transmitted as a multifactorial, oligogenic trait.¹ These studies also reported large differences in the level of OA heritability between different

joint sites (ie hip vs knee) and between the genders. They have therefore revealed a high degree of heterogeneity in the nature of the encoded susceptibility and highlighted the need for the judicious application of phenotypic stratification as a means of reducing heterogeneity in OA gene mapping studies.

We previously carried out a genome-wide linkage scan on OA affected sibling pair families collected in the UK. The affected individuals had all undergone one or more replacements of the total hip (THR), the total knee (TKR), or both, for primary OA. Stratification of the families by gender and by joint replaced (hip or knee) revealed suggestive susceptibility loci for hip OA on chromosomes 2q and 6, and for

*Correspondence: Dr J Loughlin, University of Oxford, Institute of Molecular Medicine, Oxford OX3 9DS, UK. Tel: 44 1865 220246; Fax: 44 1865 222498; E-mail: john.loughlin@ndcls.ox.ac.uk
 Received 7 February 2002; revised 24 April 2002; accepted 23 May 2002

Table 1 THR-family structures

| Family structure and type of affection | Total |
|--|-------|
| <i>Total no. of THR-families</i> | 378 |
| With affected sibling pairs | 328 |
| With affected sibling trios | 38 |
| With affected sibling quartets | 6 |
| With other ^a | 6 |
| <i>Total no. of female THR-families</i> | 146 |
| With affected sibling pairs | 136 |
| With affected sibling trios | 10 |
| <i>Total no. of male THR-families</i> | 83 |
| With affected sibling pairs | 77 |
| With affected sibling trios | 5 |
| With other ^a | 1 |
| <i>Total no. of THR affected individuals</i> | 820 |
| <i>No. of additional siblings</i> | 244 |

^aPairs of relatives, such as cousins, uncles and aunts

female-hip OA on chromosomes 4q, 11q and 16p.^{2–4} Since our genome scan employed only a low density of microsatellite markers (one every 15 cM), the intervals for these five linkages were broad. We have now finer linkage mapped these loci in an expanded cohort of OA families and report here the results for chromosome 6.

In the genome scan we had genotyped 15 chromosome 6 markers in 297 OA families.³ Multipoint linkage analysis gave an MLS of 1.0 within a large interval between markers D6S1610 and D6S262 (52.4–129.3 cM from the 6p-telomere). When we stratified this data, we obtained an MLS of 2.9 in the 194 families from the 297 that contained sibling pairs concordant for hip OA (THR-families). The linkage peak was centred between D6S462 and D6S262 (96.0–129.3 cM), with over 50 cM of the chromosome having a multipoint-LOD score >2.0. There was no evidence for linkage in families containing sibling pairs concordant for knee osteoarthritis. In the first step of the refined analysis reported here, we genotyped 20 chromosome 6 markers in an expanded cohort of 378 THR-families. We obtained an MLS of 2.2 between markers D6S1610 and D6S460 (52.4–86.8 cM), an interval of 34.4 cM. We subsequently genotyped six additional markers from within this interval and obtained an MLS of 2.8 between markers D6S452 and 509-8B2 (70.5–81.9 cM), an interval of 11.4 cM. Stratification by gender revealed that the linkage was completely accounted for by the female THR-families ($n=146$), with an MLS of 4.0 and with the highest two-point LOD score being 4.6 at D6S1573 (75.9 cM).

The COL9A1 gene resides just within the finer-linked interval, at 81.9 cM. This gene encodes for the $\alpha 1$ polypeptide chain of the heterotrimeric type IX collagen, a quantitatively minor cartilage collagen that is involved in skeletal development and which is required for maintaining

cartilage integrity.⁵ We identified and then genotyped 20 COL9A1 single nucleotide polymorphisms (SNPs) in the 146 probands from our female THR-families and in 215 age-matched female controls. No SNP allele, genotype or haplotype was associated with OA.

Overall, by genotyping a larger cohort of OA families for a higher density of microsatellite markers, and by applying gender stratification, we have finer linkage mapped the hip OA susceptibility locus on chromosome 6 to an 11.4 cM interval at 70.5–81.9 cM (6p12.3–q13). We have demonstrated that this locus is particularly relevant to female disease, whilst a comprehensive association analysis does not support COL9A1 as the chromosome 6 OA susceptibility gene.

Families and methods

Hip OA families

For the linkage analysis we examined 378 UK Caucasian families that each contained two or more siblings who had undergone one or more replacements of the total hip for primary OA (THR-families). Since OA is a late-onset disease, no parents to the siblings were available for study. We therefore collected additional siblings who had not undergone THR to assist in the determination of identical-by-descent allele transmittance (in the linkage analysis, these siblings were given an affected status of unknown). Table 1 lists the family structures. We excluded individuals who had undergone hip replacement due to prior joint trauma or other forms of arthritis by clinical examination and by reviewing X-rays and histological samples. Our ascertainment procedure has been described in detail elsewhere.² All families were genotyped for 50 unlinked microsatellite markers to confirm full-sib status using the RELATIVE program (<ftp://linkage.rockefeller.edu/software/relative/>). The average age of the affected individuals at the time of their first operation was 66 years (SD=8.9 years), with an average age of 66 years (SD=9.1 years) in affected women and an average age of 65 years (SD=8.6 years) in affected men. COL9A1 association analysis was performed on a case-control cohort comprising the 146 probands from our female THR-families (cases) and 215 age-matched, unrelated female controls. The controls were UK Caucasians who had not undergone any joint-replacement surgery or required clinical treatment for OA. The average age of the controls when recruited into the study was 73 years. Ethical approval for the study was obtained from the Central Oxford Research Ethics Committee and informed consent was obtained from all subjects.

Microsatellite markers and genotyping

Microsatellite markers were PCR amplified and the alleles sized using Genescan and Genotyper software (Applied Biosystems), as described previously.² We initially genotyped all 20 chromosome 6 markers from the Applied Biosystems LMS-MD10 panels 8–10 (<http://www.applied>

biosystems.com). These markers have an average spacing of 10 cM. We then genotyped four additional markers from the LMS-HD5 panels (D6S1549, D6S282, D6S452 and D6S1573) and the two COL9A1 microsatellite markers 509-8B2 (located in intron 5 of COL9A1) and 509-12B1 (located 14.3 kb upstream of exon 1 of COL9A1).⁶ These additional six markers map between LMS-MD10 markers D6S1610 and D6S460.

A quality control test was performed on the markers in which excess homozygosity, due to the preferential or missed amplification of one allele, was examined by RECODE v1.4 (<ftp://watson.hgen.pitt.edu/pub/>). All markers passed this check. Furthermore, Mendelian inconsistencies were examined in families containing at least three individuals by use of the PEDCHECK program (http://watson.hgen.pitt.edu/register/soft_doc.html). These tests, along with the RELATIVE analysis described above, enabled us to confirm the full-sib status of the siblings and to ensure that the chromosome 6 markers genotyped were reliable.

COL9A1 SNP discovery and genotyping

The COL9A1 gene structure was obtained from published sources.⁷ For SNP discovery two methodologies were employed: (1) heteroduplex analysis, with DNA variants then characterised by direct sequencing (Finnish group) and (2) by direct sequencing only (UK group). Nine hundred base pairs of promoter sequence, the 5' and 3' UTRs, all 38 COL9A1 exons (including acceptor and donor splice sites), differentially spliced exon 1*, and all of intron 1 were scanned in at least 35 unrelated individuals. All of intron 1 was screened since several collagen genes are known to contain *cis* elements in the first intron that can regulate gene expression.⁸ SNPs were genotyped by PCR-restriction enzyme analysis, with the digestion fragments separated by electrophoresis through 3% agarose. Further information regarding primer sequences, and the PCR and digestion conditions can be obtained from the authors.

Statistical analysis – microsatellites

In the linkage analysis, microsatellite allele frequencies were calculated from all scored genotypes by use of the GAS program (<http://users.ox.ac.uk/~ayoung/gas.html>) and subsequent two-point and multipoint linkage analysis were performed using ANALYZE (<ftp://linkage.cpmc.columbia.edu>) and ASPEX (<ftp://lahmed.stanford.edu/pub/aspex>), respectively. The two COL9A1 markers 509-8B2 and 509-12B1 were tested for evidence of transmission disequilibrium in our families using the TRANSMIT sibling-TDT program, version 2.5.2.⁹ This program is appropriate to our study as it incorporates data from families with missing parents.

Statistical analysis – SNPs

Allelic association with OA was evaluated for each COL9A1 SNP by χ^2 using standard contingency table analysis. We

further stratified our families using their individual non-parametric linkage (NPL) scores taken from the GENEHUNTER two-point linkage analysis (<http://waldo.wi.mit.edu/ftp/distribution/software/genehunter/gh2/>). Probands from the female THR-families which showed a positive contribution to the linkage for the two COL9A1 markers (positive NPL scores) were compared with female controls.

The pair-wise linkage disequilibrium coefficient D' values were calculated using the GOLD program (<http://www.well.ox.ac.uk/asthma/GOLD>) and displayed graphically on a physical map to reveal the SNP haplotype structure.¹⁰ The positions of the COL9A1 SNPs on the map were defined using published sources.⁷ Haplotype frequencies amongst blocks of SNPs in linkage disequilibrium were calculated using the estimate haplotype program EH-PLUS (<http://www.iop.kcl.ac.uk/IoP/Departments/PsychMed/GEpiBST/software.stm>). Haplotype frequency differences between the probands and the controls were then compared using χ^2 and contingency table analysis.

Results

Linkage analysis

Table 2a lists the two-point LOD scores and Figure 1a shows the multipoint linkage plot for the 20 LMS-MD10 markers genotyped in the 378 THR-families. The highest two-point LOD score was 1.6 for marker D6S257, whilst the maximum

Table 2 Two-point LOD score in the 378 THR-families

| Marker | Distance from 6p telomere (in cM) | Two-point LOD score |
|---|-----------------------------------|---------------------|
| <i>(a) For the chromosome 6 LMS-MD10 microsatellite markers</i> | | |
| D6S1574 | 8.2 | 0.0 |
| D6S309 | 12.9 | 0.6 |
| D6S470 | 17.7 | 0.0 |
| D6S289 | 29.5 | 0.0 |
| D6S422 | 34.8 | 0.0 |
| D6S276 | 43.7 | 0.0 |
| D6S1610 | 52.4 | 0.6 |
| D6S257 | 78.3 | 1.6 |
| D6S460 | 86.8 | 0.6 |
| D6S462 | 96.0 | 0.8 |
| D6S434 | 107.5 | 0.7 |
| D6S287 | 120.9 | 0.0 |
| D6S262 | 129.3 | 0.1 |
| D6S292 | 135.1 | 0.0 |
| D6S308 | 144.0 | 0.1 |
| D6S441 | 153.6 | 0.0 |
| D6S1581 | 163.1 | 0.0 |
| D6S264 | 177.6 | 0.0 |
| D6S446 | 184.9 | 0.4 |
| D6S281 | 197.5 | 0.3 |
| <i>(b) For the six additional chromosome 6 microsatellite markers</i> | | |
| D61549 | 59.6 | 0.0 |
| D6S282 | 64.8 | 1.4 |
| D6S452 | 70.5 | 1.5 |
| D6S1573 | 75.9 | 2.3 |
| COL9A1 509-8B2 | 81.9 | 1.6 |
| COL9A1 509-12B1 | 82.0 | 0.8 |

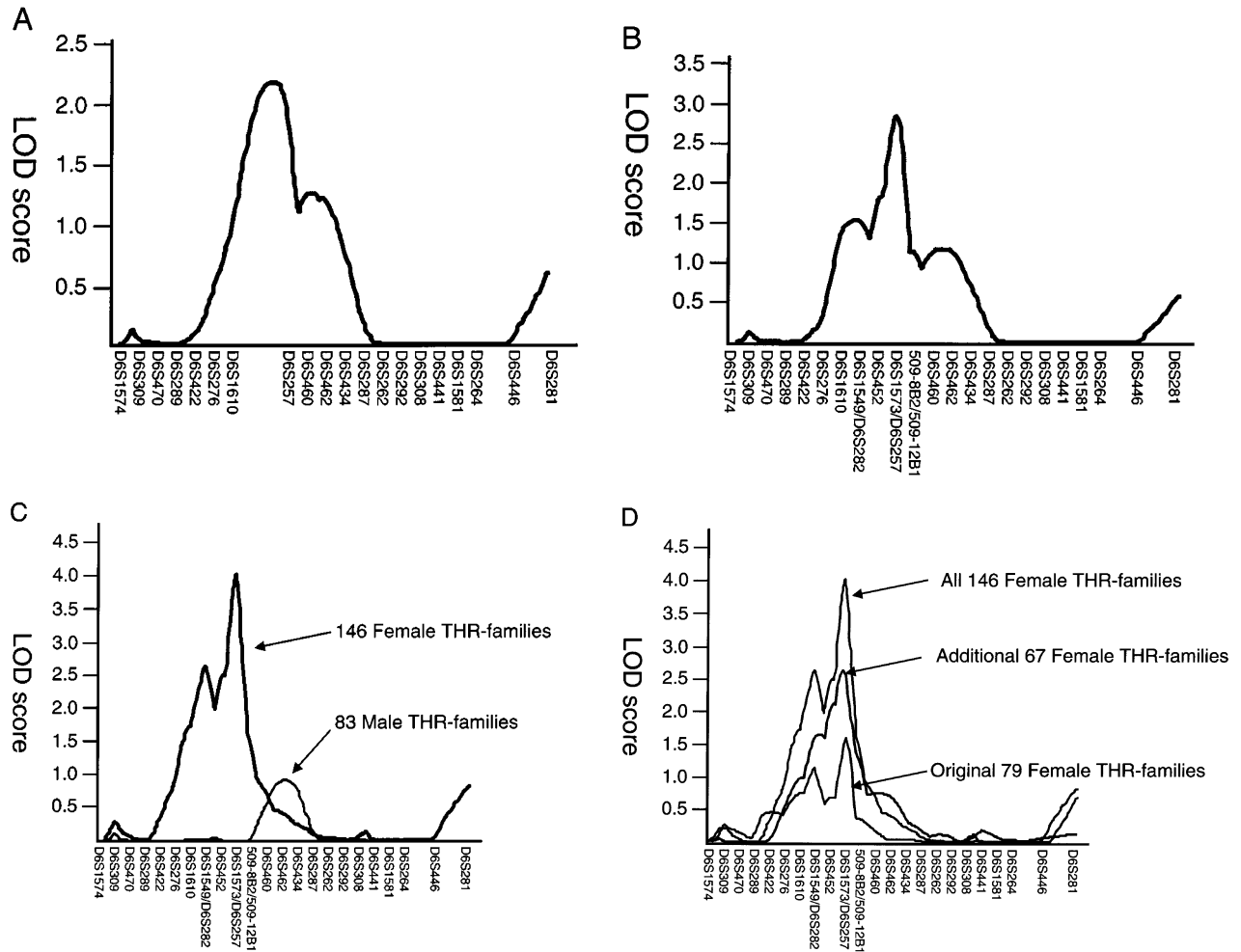


Figure 1 Multipoint linkage analysis of chromosome 6. (A) For the 20 chromosome 6 markers from the Applied Biosystems LMS-MD10 panels in the 378 THR-families. (B) With six additional markers between D6S1610 and D6S460. (C) For all 26 chromosome 6 markers with the 378 THR-families stratified by gender. (D) With the 146 female THR-families stratified into the 79 families from our original genome-wide linkage scan and the 67 families that were subsequently recruited.

multipoint-LOD score (MLS) was 2.2, located between markers D6S1610 and D6S460, a 34.4 cM interval. Although the LMS-MD10 markers have an average density of 10 cM, there is a relatively large gap between markers D6S1610 and D6S257 of 25.9 cM. Since our linkage peak was centred on D6S257 we genotyped four additional markers from within this gap. These were the LMS-HD5 markers D6S1549, D6S282, D6S452 and D6S1573. We also genotyped the two COL9A1 markers 509-8B2 and 509-12B1⁶, which map between D6S257 and D6S460. Table 2b lists the two-point LOD scores for these six additional markers. Four had LOD scores ≥ 1.4 , with the highest being 2.3 for marker D6S1573. Figure 1b shows the multipoint linkage plot for all 26 chromosome 6 markers. The MLS had increased to 2.8 in an 11.4 cM interval defined by markers D6S452 and 509-8B2.

As noted earlier, several studies have reported differences in the degree of heritability of OA between males and females.¹ We therefore stratified our linkage data by gender. The highest two-point LOD score was 4.6 for marker D6S1573 in the 146 female THR-families (Table 3). Four markers flanking D6S1573 had LOD scores ≥ 2.1 . In the 83 male THR-families, D6S1573 and its flanking markers showed no evidence for linkage. An MLS of 4.0 was obtained for the female THR-families in the 11.4 cM interval flanked by D6S452 and 509-8B2 (Figure 1c). The male THR-families were unlinked to this interval.

The 146 female THR-families are composed of 79 families used in our original genome-wide linkage scan³ and 67 families that were subsequently recruited. To demonstrate the contribution that each collection of families made to the linkage, we carried out the linkage analysis on each

cohort (Figure 1d). The MLS for the original 79 families was 1.6, the MLS for the additional 67 families was 2.6, whilst

Table 3 Two-point LOD scores for all 26 of the chromosome 6 markers genotyped, with the THR-families stratified by gender

| Marker | Distance from 6p telomere (in cM) | Two-point LOD score | |
|----------|-----------------------------------|-----------------------------|--------------------------|
| | | Female THR (n=146 families) | Male THR (n=83 families) |
| D6S1574 | 8.2 | 0.0 | 0.0 |
| D6S309 | 12.9 | 0.9 | 0.5 |
| D6S470 | 17.7 | 0.0 | 0.0 |
| D6S289 | 29.5 | 0.0 | 0.1 |
| D6S422 | 34.8 | 0.3 | 0.0 |
| D6S276 | 43.7 | 1.1 | 0.0 |
| D6S1610 | 52.4 | 0.6 | 0.0 |
| D6S1549 | 59.6 | 1.1 | 0.0 |
| D6S282 | 64.8 | 2.2 | 0.0 |
| D6S452 | 70.5 | 2.1 | 0.1 |
| D6S1573 | 75.9 | 4.6 | 0.0 |
| D6S257 | 78.3 | 2.5 | 0.0 |
| 509-8B2 | 81.9 | 2.2 | 0.0 |
| 509-12B1 | 82.0 | 1.0 | 0.0 |
| D6S460 | 86.8 | 0.8 | 0.3 |
| D6S462 | 96.0 | 0.3 | 0.9 |
| D6S434 | 107.5 | 0.1 | 0.5 |
| D6S287 | 120.9 | 0.0 | 0.0 |
| D6S262 | 129.3 | 0.2 | 0.0 |
| D6S292 | 135.1 | 0.0 | 0.0 |
| D6S308 | 144.0 | 0.3 | 0.1 |
| D6S441 | 153.6 | 0.0 | 0.0 |
| D6S1581 | 163.1 | 0.0 | 0.0 |
| D6S264 | 177.6 | 0.0 | 0.0 |
| D6S446 | 184.9 | 0.7 | 0.0 |
| D6S281 | 197.5 | 0.4 | 0.0 |

the combined MLS was 4.0. This analysis demonstrates that an increase in the sample size has increased the evidence for linkage, strongly suggesting that the chromosome 6 linkage is robust.

Association analysis of COL9A1

Twenty common SNPs (rare allele frequencies >0.05) were identified in COL9A1 (Table 4). One of these SNPs is located in the promoter, two are exonic (both non-synonymous), with the remaining 17 being intronic. None of the intronic SNPs were located in donor or acceptor splice sites, or affected the conserved (A) nucleotide of any putative branch consensus sites.

The 20 common SNPs were genotyped in the 146 probands from our female THR-families and in 215 age-matched female controls (all SNPs were in Hardy-Weinberg equilibrium in the control cohort). None of the SNP alleles or genotypes were associated with OA at $P \leq 0.05$. Three SNP alleles gave P -values ≤ 0.1 : intron 16 (+143) ($P=0.09$); intron 25 (-24) ($P=0.10$); and exon 28 (+44) ($P=0.10$) (Table 4). To determine whether any of these three SNPs were suggestive of an association, we stratified our female THR-families into those that showed a contribution to the linkage (positive GENEHUNTER NPL scores). We stratified against the COL9A1 markers 509-8B2 and 509-12B1. Seventy-four of the 146 female THR-families showed linkage to 509-8B2 and 67 showed linkage to 509-12B1. When the SNP allele and genotype frequencies were compared between the female controls and the probands from the families contributing to the linkage at each marker, none of the three SNPs tested showed a

Table 4 Single nucleotide polymorphisms in COL9A1 and association study of probands from the female THR-families and of controls

| Location ^a | M→m ^b | Change Amino acid | Allele frequency ^c | | χ^2 ^d | P |
|------------------------------|------------------|----------------------|-------------------------------|----------------|-----------------------|------|
| | | | Controls | OA probands | | |
| Promoter (-364) ^e | C→G | | 0.17 (75/355) | 0.16 (46/242) | 0.17 | 0.68 |
| Intron 1 (+39) | A→G | | 0.19 (82/348) | 0.17 (47/235) | 0.51 | 0.48 |
| Intron 1 (-197) | C→G | | 0.08 (34/394) | 0.06 (16/268) | 1.06 | 0.30 |
| Intron 1 (-54) | G→A | | 0.48 (205/221) | 0.50 (146/144) | 0.26 | 0.61 |
| Intron 2 (-12) | T→C | | 0.16 (69/353) | 0.18 (54/238) | 0.41 | 0.52 |
| Intron 4 (-69) | del G | | 0.09 (39/391) | 0.06 (18/268) | 1.45 | 0.23 |
| Intron 8 (-23) | C→G | | 0.08 (33/393) | 0.06 (18/268) | 0.35 | 0.56 |
| Intron 10 (+45) | G→A | | 0.17 (74/356) | 0.19 (56/236) | 0.33 | 0.57 |
| Exon 11 (+40) | T→C | Ser→Pro | 0.38 (164/266) | 0.37 (109/183) | 0.02 | 0.89 |
| Intron 16 (+12) | T→C | | 0.17 (73/357) | 0.17 (48/242) | 0.00 | 0.96 |
| Intron 16 (+143) | A→G | | 0.50 (203/205) | 0.57 (163/125) | 2.90 | 0.09 |
| Intron 17 (-65) | T→C | | 0.19 (82/348) | 0.18 (51/239) | 0.16 | 0.69 |
| Intron 18 (+117) | T→A | | 0.20 (85/345) | 0.15 (44/244) | 2.06 | 0.16 |
| Intron 19 (-88) | C→T | | 0.21 (88/340) | 0.17 (48/240) | 1.45 | 0.23 |
| Intron 23 (-26) | C→A | | 0.16 (68/354) | 0.16 (46/246) | 0.00 | 0.98 |
| Intron 25 (-24) | C→A | | 0.41 (173/253) | 0.34 (98/188) | 2.66 | 0.10 |
| Exon 28 (+44) | G→A | Arg→Gln | 0.41 (175/255) | 0.34 (98/188) | 2.75 | 0.10 |
| Intron 31 (+19) | T→C | | 0.18 (76/354) | 0.17 (49/243) | 0.04 | 0.84 |
| Intron 34 (+32) | C→T | | 0.49 (211/219) | 0.55 (162/130) | 2.60 | 0.11 |
| Intron 36 (+69) | C→T | | 0.08 (36/394) | 0.08 (24/266) | 0.00 | 0.93 |

^aNumber of nucleotides from nearest start of intron or exon; ^bM=major, common allele in controls; m=minor, less common allele in controls; ^cOf the allele that is less common in the controls; ^ddf=1; ^e364 nucleotides from the start of translation.

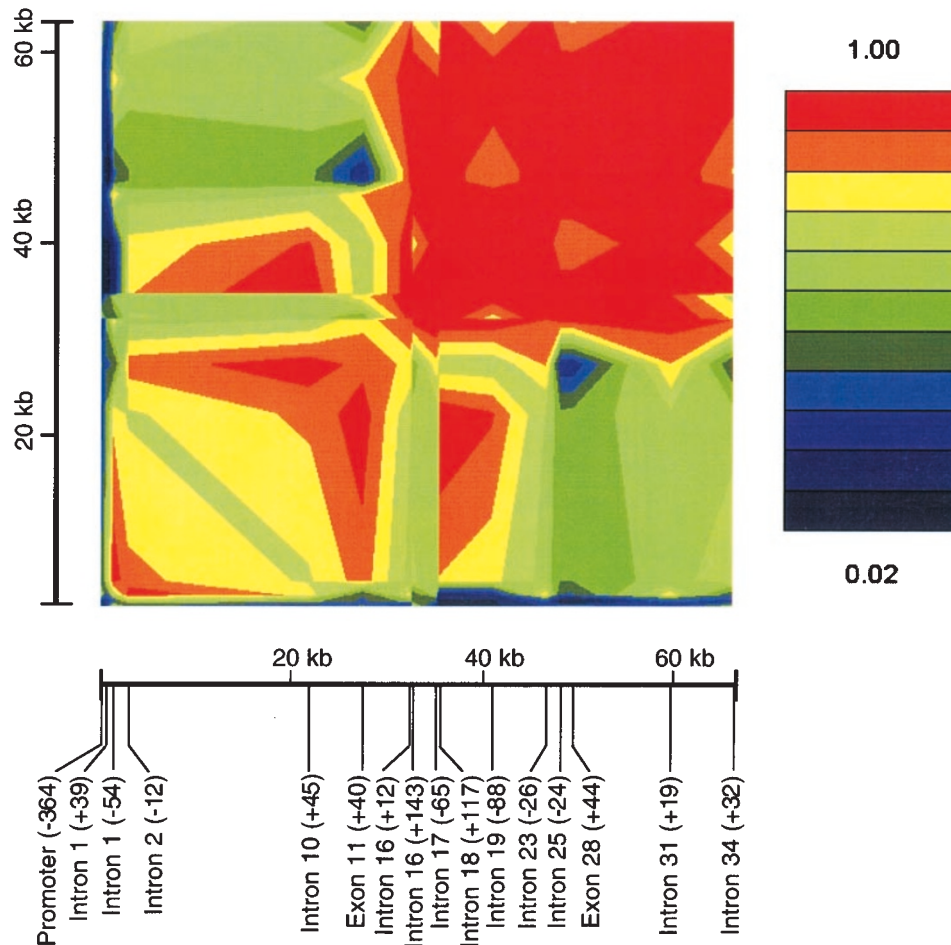


Figure 2 Graphical overview of pair-wise linkage disequilibrium (D') values across COL9A1 for the sixteen SNPs that have allele frequencies >0.1 . The axis of the gene runs along the diagonal from the bottom left to the top right. The D' values are colour coded from red (linkage disequilibrium) to blue (linkage equilibrium).

significant allele or genotype frequency difference at $P \leq 0.05$ (data not shown).

We next determined the COL9A1 SNP haplotype structure and its diversity. We focused on the 16 SNPs from the 20 that had minor allele frequencies >0.10 . Pair-wise linkage disequilibrium D' values were calculated using the GOLD program and displayed graphically against the physical distance between the SNPs (Figure 2). This revealed a high degree of linkage disequilibrium between the 14 SNPs encompassed by Intron 1 (-54) to Intron 34 (+32), inclusive, with D' never falling below 0.7. This represents a haplotype block of 65 kb. Haplotype blocks of this length have been reported on in other studies.^{11,12} The two remaining SNPs tested [Promoter (-364) and Intron 1 (+39)] approached linkage equilibrium with each other ($D'=0.16$), suggesting that a recombination hotspot resides immediately upstream of SNP Intron 1 (-54).¹³ Using EHPLUS, we constructed the haplotypes for the 14 SNPs in

linkage disequilibrium. Of the 16384 possible haplotypes, only 57 were observed confirming the high degree of linkage disequilibrium between the SNPs. Of the 57 haplotypes, only six had frequencies $>5\%$ and these accounted for 72% of the chromosomes in our probands and controls. There was no significant difference in the frequency of these common haplotypes between the probands and the controls ($P=0.60$).

Finally, we tested the two COL9A1 microsatellite markers 509-8B2 and 509-12B1 for association to OA by assessing the evidence of transmission disequilibrium in the 146 female THR-families using the TRANSMIT program. Neither marker demonstrated transmission distortion (data not shown).

Discussion

In this study, we have finer linkage mapped a susceptibility locus for severe hip OA in females to an 11.4 cM interval at

6p12.3-q13 (70.5 to 81.9 cM from the 6p-telomere). The COL9A1 gene, which encodes for the $\alpha 1$ polypeptide chain of type IX collagen, maps within this interval. COL9A1 is a logical candidate for OA susceptibility. A mutation in the gene was recently reported in the affected members of a multiple epiphyseal dysplasia pedigree, a rare chondrodysplasia that has early-onset (secondary) OA as one of its phenotypic components.¹⁴ However, our comprehensive association analysis did not produce any evidence supporting COL9A1 as the primary OA susceptibility locus that we have mapped to chromosome 6. We cannot exclude, however, the possibility that several rare alleles of COL9A1 encode for the chromosome 6 susceptibility, with our study lacking the necessary power to detect this association.

We have previously excluded a number of other genes that are mutated in the chondrodysplasias as candidates for primary OA susceptibility.¹⁵ The fact that the genes which encode for such secondary OA phenotypes are not necessarily those that are involved in primary OA susceptibility reinforces the view that these two forms of OA are distinct entities that do not share all causal genetic factors.

The finer linkage mapped region of chromosome 6 contains other potential candidates. Two other collagen genes map close to COL9A1; COL19A1, which encodes for the $\alpha 1$ polypeptide chain of the homotrimeric type XIX collagen, and COL12A1, which encodes for the $\alpha 1$ chain of the homotrimeric type XII collagen.¹⁶ Type XII collagen is present in cartilage, tendon and bone⁵, making COL12A1 a plausible candidate. However, this gene maps approximately 5 cM distal to COL9A1, placing it approximately 11 cM from the centre of the linked interval. A candidate that maps closer to the interval centre is the bone morphogenetic protein 5 (BMP5) gene. BMPs are a class of secreted signaling molecules that are involved in a number of different biological processes including skeletal patterning.¹⁷ The choice of candidates for subsequent linkage disequilibrium analysis should take into account the female specificity of the chromosome 6 linkage: genes that are under hormonal control should be studied first.

Overall, we have finer linkage mapped an OA susceptibility locus on chromosome 6. This was achieved by genotyping a large number of families for a high density of microsatellite markers, and by stratifying the data by gender. The dramatic effect of gender stratification highlights the heterogeneity of OA susceptibility that has been reported on in a number of epidemiological studies and which is now being confirmed by genetic analysis. The linked interval, at 11.4 cM, is sufficiently narrow for targeted and more systematic linkage disequilibrium/association analysis to be undertaken.

Acknowledgements

This work was funded by The Arthritis Research Campaign, The Wellcome Trust and the Nuffield Foundation (to J Loughlin), and

by the Arthritis Foundation, the Academy of Finland, Louisiana Gene Therapy Research Consortium (New Orleans, LA, USA) and HCA-The Health Care Company (Nashville, TN, USA) (to L Ala-Kokko). J Loughlin is an Arthritis Research Campaign Fellow. We thank Professor Andrew Carr, Dr Jai Chitnavis and Ms Kim Cliphsham who organised the collection of the patient and family samples used in this study. That collection was partly funded by The Wishbone Trust.

References

- Loughlin J: Genetic epidemiology of primary osteoarthritis. *Curr Opin Rheumatol* 2001; **13**: 111–116.
- Chapman K, Mustafa Z, Irven CM *et al*: Osteoarthritis-susceptibility locus on chromosome 11q, detected by linkage. *Am J Hum Genet* 1999; **65**: 167–174.
- Loughlin J, Mustafa Z, Irven C *et al*: Stratification analysis of an osteoarthritis genome screen – suggestive linkage to chromosomes 4, 6 and 16. *Am J Hum Genet* 1999; **65**: 1795–1798.
- Loughlin J, Mustafa Z, Smith A *et al*: Linkage analysis of chromosome 2q in osteoarthritis. *Rheumatol* 2000; **39**: 377–381.
- Ayad S, Boot-Handford RP, Humphries MJ, Kadler KE, Shuttleworth CA: *The extracellular matrix facts book*. 2nd edn. London: Academic Press, 1998; pp 84–89.
- Warman ML, Tiller GE, Polumbo PA *et al*: Physical and linkage mapping of the human and murine genes for the $\alpha 1$ chain of type IX collagen (COL9A1). *Genomics* 1993; **17**: 694–698.
- Pihlajamaa T, Vuoristo MM, Annunen S, Perälä M, Prockop DJ, Ala-Kokko L: Human COL9A1 and COL9A2 genes. Two genes of 90 and 15 kb code for similar polypeptides of the same collagen molecule. *Matrix Biol* 1998; **17**: 237–241.
- Horton WE, Wang L, Bradham D, Precht P, Balakir R: The control of expression of type II collagen: relevance to cartilage disease. *DNA Cell Biol* 1992; **11**: 193–198.
- Clayton D: A generalization of the transmission/disequilibrium test for uncertain haplotype transmission. *Am J Hum Genet* 1999; **65**: 1170–1177.
- Abecasis GR, Cookson WOC: GOLD – Graphical overview of linkage disequilibrium. *Bioinformatics* 2000; **16**: 182–183.
- Reich DE, Cargill M, Bolk S *et al*: Linkage disequilibrium in the human genome. *Nature* 2001; **411**: 199–204.
- Daly MJ, Rioux JD, Schaffner SF, Hudson TJ, Lander ES: High-resolution haplotype structure in the human genome. *Nature Genet* 2001; **29**: 229–232.
- Jeffreys AJ, Kauppi L, Neumann R: Intensely punctate meiotic recombination in the class II region of the major histocompatibility complex. *Nature Genet* 2001; **29**: 217–222.
- Czarny-Ratajczak M, Lohiniva J, Rogala P *et al*: A mutation in COL9A1 causes multiple epiphyseal dysplasia: further evidence for locus heterogeneity. *Am J Hum Genet* 2001; **69**: 969–980.
- Mustafa Z, Chapman K, Irven CM *et al*: Linkage analysis of candidate genes as susceptibility loci for osteoarthritis – suggestive linkage of COL9A1 to female hip osteoarthritis. *Rheumatol* 2000; **39**: 299–306.
- Gerecke DR, Olson PF, Koch M *et al*: Complete primary structure of two splice variants of collagen XII and assignment of $\alpha 1$ (XII) collagen (COL12A1), $\alpha 1$ (IX) collagen (COL9A1), and $\alpha 1$ (XIX) collagen (COL19A1) to human chromosome 6q12-q13. *Genomics* 1997; **41**: 236–242.
- Reddi AH: Interplay between bone morphogenetic proteins and cognate binding proteins in bone and cartilage development: noggin, chordin and DAN. *Arthritis Res* 2001; **3**: 1–5.