Genome-wide Association Study Identified TIMP2 Genetic Variant with Susceptibility to Osteoarthritis

Bhumsuk Keam[2¶], Joo-Yeon Hwang[1¶], Min Jin Go[1], Jee Yeon Heo[1], Mi Sun Park[1], Ji Young Lee[1], Nam Hee Kim[1], Miey Park[1], Ji Hee Oh[1], Dong-Hyun Kim[3], Jin-Young Jeong[4], Jong-Young Lee[1], Bok-Ghee Han[1] and Juyoung Lee[1*]

[1]Center for Genome Science, National Institute of Health, Osong Health Technology Administration Complex, Chung-cheongbuk- do 363-951, Korea, [2]Department of Internal Medicine, Seoul National University Hospital, 101 Daehak-ro, Seoul 110-744, Korea, [3]Department of Social and Preventive Medicine, College of Medicine, Hallym University, 1, Ok-cheon, Chuncheon, Gangwon-do 200-702, Korea, [4]Institute of Aging, Hallym University, 1, Ok-cheon, Chuncheon, Gangwon-do 200-702, Korea

Abstract

Osteoarthritis (OA) is the most common degenerative joint disorder in the elderly population. To identify OA-associated genetic variants and candidate genes, we conducted a genome-wide association study (GWAS). A total 3,793 samples (476 cases: wrist + knee and 3317 controls) from a community-based epidemiological study were genotyped using the Affymetrix SNP 5.0. An intronic SNP (rs4789934) in the TIMP2 (tissue inhibitor of metalloproteinase-2) showed the most significance with OA (odd ratio [OR] = 2.06, 95% confidence interval [CI] = 1.52-2.81, p = 4.01 $\times$ 10^{-6}). Furthermore, a polymorphism (rs1352677) in the NKAIN2 (Na$^+$/K$^+$ transporting ATPase interacting 2) was suggestively associated with OA (OR = 1.43, CI = 1.22-1.66, p = 7.01 $\times$ 10^{-6}). The present study provides new insights into the identification of genetic predisposing factors for OA.

Keywords: genome-wide association study, osteoarthritis, polymorphism, TIMP2

Introduction

Osteoarthritis (OA) is the most common joint disorder characterized by progressive cartilage loss, and con-

sequently increases serious social/economic problems in the elderly population worldwide (Du *et al.*, 2005; Ikegawa, 2007). As with any joint affected by osteoarthritis, the primary symptom for wrist and knee OA as predominant weight-bearing joints is intractable pain with increased stiffness. Specifically, wrist OA can occurs secondary to an intra-articular fracture of the distal radius or an extra-articular fracture resulting in malunion (Weiss and Rodner, 2007). Several risk factors such as age, gender, obesity, and trauma have been confirmed as environmental confounding factors (Haq *et al.*, 2003).

So far, genetic factors contribute to the risk of OA have been identified through candidate gene approaches (Loughlin, 2002; MacGregor and Spector, 1999). However, candidate gene-based studies (Miyamoto *et al.*, 2007; Mototani *et al.*, 2005; Rodriguez-Lopez *et al.*, 2008) have only a small fraction of genetic interactions involved in the pathogenesis, and thereby provide little chance of discovering novel candidate genes.

As a solution to the problem, OA susceptibility genes by recent GWAS have been reported across different groups (Abel *et al.*, 2006; Dai and Ikegawa, 2010; Loughlin, 2002; Spector *et al.*, 2006). However, most of the previously reported findings have produced conflicting results with differential genetic effects and limited replication between ethnic populations. (Jiang *et al.*, 2008; Kerkhof *et al.*, 2008; Shi *et al.*, 2008; Snelling *et al.*, 2007; Valdes *et al.*, 2007). The ethnic-specific variants remain to be identified in East Asian population to overcome ethnic difference and population structure. We therefore undertook a genome-wide association study to identify genetic polymorphisms influencing OA.

Methods

Study population

The study samples were drawn from the Korean Genome and Epidemiology Study (KoGES) which is an ongoing prospective community-based epidemiological study in the communities of Ansung (rural) and Ansan (urban) (Table 1). Details of the KoGES and the methods are described in our previous report (Cho *et al.*, 2006; Cho *et al.*, 2009; Kim *et al.*, 2005; Lim *et al.*, 2006). In brief, eligible subjects (age 40-69 years) were examined in 2001-2002 for demography and epidemiology and then follow up biannually. A total of 3,588 men and

3,927 women agreed to participate in a baseline examination which included an interview, blood tests, X-rays, and measurement of the body. Informed consents were obtained from all of the participants, and Declaration of Helsinki for biomedical research involving human subjects was also followed. The study protocol was approved by the Institutional Review Board of the Korea Centers for Disease Control and Prevention.

Diagnostic criteria

X-rays of both wrists and knees of subjects were taken. Radiographic OA was assessed using the Kellgren/Lawrence (K/L) grading system (Kellgren and Lawrence, 1957). Patients with K/L grades of 2 or higher in wrists or knees were diagnosed with OA. Among the 7,515 subjects of the study, 3,793 subjects consented to the X-ray examination. Of these, 476 patients were diagnosed with OA (387 women, 89 men, median age = 68 years [range 44-74]), and 3,317 subjects were classified as normal controls (women=1,773, men=1,544, median age = 57 years [range 43-74]).

Genome-wide association

Genomic DNA from peripheral blood mononuclear cells was used for this study. We performed a pilot GWAS, typing cases and controls on a single platform using the Affymetrix Genome-Wide Human SNP 5.0 (Affymetrix, Inc., Santa Clara, CA, USA). Genotype calls were determined by Bayesian Robust Linear Modeling using the Mahalanobis Distance algorithm (Rabbee and Speed, 2006). Details of the genotyping process are presented in our previous report (Cho *et al.*, 2009). Of the SNPs assayed on the chip, 179,626 SNPs were excluded because they showed: 1) a call rate lower than 96.0% in cases or controls; or 2) a minor allele frequency (MAF) $< 1\%$ in the population; or 3) a significant distortion from Hardy-Weinberg equilibrium (p < 0.05). A total of 320,942 SNPs passed all the quality control filters (mean call rate 98.0%). To identify the effect of population stratification, the genomic inflation factor (λ) (Clayton *et al.*, 2005) was calculated for all individuals. The two study cohorts (Ansung and Ansan) have shown very
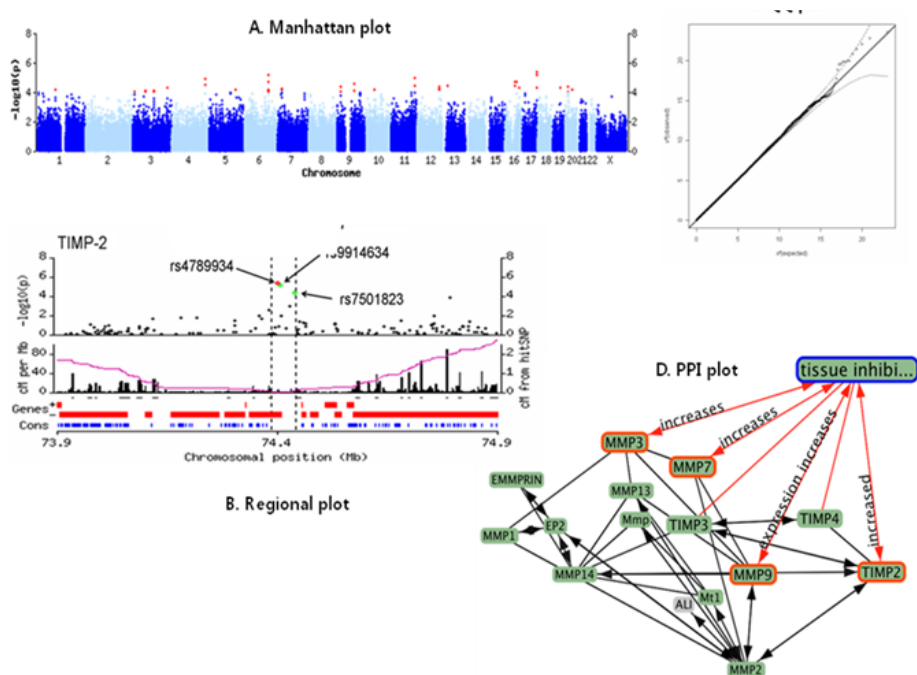

Fig. 1. Results of the genome-wide association study (GWAS) for OA. A. $-\log_{10}$ (p) values for associations (Red dots represent the 42 SNPs that passed the inclusion criteria of a P-value $<10^{-4}$). B. Signal intensity plots showing the lead SNPs associated with OA (rs4789934 in TIMP2 with the most significant association, adjusted P-values = 4.01 $\times$ 10^{-6}). C. Quantile-Quantile plots of test statistics for the associations of SNPs with OA (Y axis is the observed $-\log_{10}$ (p) values. X axis is the expected $-\log_{10}$ (p) values under the null distribution of no association for studied genome-wide SNPs, which would be expected to follow the solid black line) (the genomic inflation factor λ = 1.009). D. TIMP2-centered protein-protein interaction (PPI) network resulted from meaningful relativity underlying literature (A path length between associated nodes represents correlation by confidence scoring).

similar MAFs for the SNPs on the array, and quantile-quantile (Q-Q) analysis for the comparison of genotype frequencies in both cohorts confirmed the genetic homogeneity of these two components of the KoGES study population (Cho *et al.*, 2009).

Statistical analysis

The allele frequencies in OA cases were compared to those in controls using the Chi-square test and logistic regression analysis. The associations between the case-control status and each individual SNP were measured by the odds ratio (OR). Covariates used for multivariable-adjustment were age, sex, and body mass index (BMI), because these variables were significantly associated with OA in our cohort.

Pair-wise linkage disequilibrium (LD) among the polymorphisms was determined by the expectation maximization algorithm. When determining LD between SNPs, we used either of two measures of LD: an r^2 value $\geq$ 0.8 or a disequilibrium coefficient (D') $\geq$ 0.95. PLINK software program (http://pngu.mgh.harvard.edu/~purcell/plink/), Haploview program (http://www.broad.mit.edu/mpg/haploview), and SAS program (version 9.1; SAS Institute Inc., Cary, NC, USA) were used for statistical analysis.

Results

The Q-Q plot indicated lack of inflation due to stratification, and no differences between case and control populations were detected. The estimated genomic control inflation factor (λ) was 1.009, indicating limited evidence of population stratification in the KARE study samples - are described in Fig. 1.

Table 1. Basic characteristics of study subjects

	Case	Control	P-value
Number	476	3317	
Sex (M/F)	89/387	1554/1773	< 0.0001
Age (years)	66.5 ± 5.8	57.5 ± 8.6	< 0.0001
BMI	25.5 ± 3.5	24.4 ± 3.2	< 0.0001

Data are expressed as mean ± standard deviation.

Table 2 lists the most significant genetic variants from GWAS analysis. We identified two intronic SNPs with P-values $< 1 \times 10^{-5}$. A SNP (rs4789934) in the TIMP2 (tissue inhibitor of metalloproteinase-2) gene showed the most significant association with OA. The adjusted ORs were 2.06 (95% confidence interval (CI) = 1.52-2.81), and the adjusted P-values were 4.01×10^{-6}. In addition, our study also identified suggestive association in the NKAIN2 (Na$^+$/K$^+$ transporting ATPase interacting 2) (OR = 1.43, CI = 1.22-1.66, p = 7.01×10^{-6}).

To support the biological relevance in silico, we provide graphically summarized information on preliminary association by confidence values of textual evidence (Fig. 1. D). Nodes represent biological entities with different colors for different classes. Edges represent strong associations between two entities (i.e., TIMP2 effects to tissue inhibition, and functional interactions with TIMP4).

Discussion

In this study, we identified two novel polymorphisms associated with OA. Intronic SNPs, rs4789934 in TIMP2 and rs1352677 in NKAIN2, showed significant associations with OA. To date, OA susceptibility genes from genome-wide association studies have been reported primarily in samples of European descent.

The function of TIMP2 is inhibition of MMP. Osteoarthritic cartilage is characterized by the imbalance between MMPs and TIMPs. The level of lytic enzymes such as MMP-1, MMP-3, and MMP-13 was increased, and the level of TIMP-1 was decreased in OA cartilage (Krane and Inada, 2008; Martel-Pelletier *et al.*, 1994). With the decreasing level of TIMP, MMP activity increases and leads to further destruction of cartilage, the key process in OA (Cawston, 1998; Krane and Inada, 2008). Estrogen may regulate the balance between MMP and TIMP activity (Lee *et al.*, 2003), which would explain the gender and menopausal correlations with OA. Although prior reports focused primarily on TIMP-1, levels of TIMP2, a member of the same gene family, have also been shown to increase in osteoarthritic cartilage (Martel-Pelletier *et al.*, 1994). Lee *et al.* reported that TIMP-4 polymorphism (rs17035945) was associated with the risk of OA in the Korean population as a result

Table 2. Summary of association result from GWAS study

SNP rs#	Chr	Position	Gene	Risk / Other allele	MAF case	MAF control	OR (95% CI)a	Adjusted P-value
rs4789934	17	74423131	TIMP2	G/A	0.08	0.04	2.06 (1.52-2.81)	4.01×10^{-6}
rs1352677	6	124244288	NKAIN2	C/G	0.49	0.43	1.43 (1.22-1.66)	7.01×10^{-6}

Chr, chromosome; MAF, minor allele frequency; OR, odds ratio; CI, confidence interval.
[a]OR was adjusted by age, sex, and body mass index.

of altering the folding of the mRNA (Lee *et al.*, 2008).

The Na$^+$/K$^+$ ATPase is an important regulator of ion homeostasis in many cell types (Mobasheri *et al.*, 2000), but its role in OA is not clear. In osteoblasts, which synthesize collagen and other extracellular matrix (ECM) components, the Na$^+$/K$^+$ ATPase appears to maintain the transmembrane gradient of Na$+$ that drives intracellular calcium homeostasis and ECM calcification (Francis *et al.*, 2002). Hence, the role of NKAIN2 in the pathogenesis of OA may be related to control of mineralization and maintenance of the ECM. However, the Na$^+$/K$^+$ ATPase is expressed in diverse cell types in the human body and is not specific to OA. The detailed contributions of the Na$^+$/K$^+$ ATPase to the pathogenesis of OA should be clarified in further investigations.

The genetic susceptibility to OA is likely to vary for different joint sites (Peach *et al.*, 2005). The criteria used to diagnose OA also differ among studies and may explain conflicting information. Some studies used clinical diagnosis while other studies using X-ray and K/L grades (Kerkhof *et al.*, 2008; Miyamoto *et al.*, 2007; Spector *et al.*, 2006).

The present study contains some limitations. No SNPs met conventional genome-wide significance level of $< 5 \times 10^{-8}$, and no replication. However it would be interesting to use the data available to perform meta-analysis or SNP lookup from a large international consortium in future studies to evaluate the significance of our results.

In an attempt to address function prediction of SNP allele, we conducted in silico functional investigation to an uncharacterized sequence variant (Yuan *et al.*, 2006). The lead SNP (rs4789934) was a predicted functional cofactor, affecting transcription factor binding sites in the intronic enhancer region. This result was further associated with functional connectivity derived from bio-informative approaches (Lee and Shatkay, 2009; Szklarczyk *et al.*). Therefore, this concordance of the functional evidence with the previous studies suggests an encouraging sign of the results in our GWAS. However, the biological relevance of literature-derived targets needs to be fully validated with experimental tests.

In conclusion, we suggest that a novel polymorphism (rs4789934) in TIMP2 was associated with OA. Follow-up studies performed with multiple ethnic populations or larger sample are needed before the effect of the variant can be fully and accurately evaluated. We anticipate that identification of the causal variants through deep sequencing and their functional consequences will lead to a better understanding of the pathogenesis of OA.

List of abbreviations

BMI = body mass index; ECM = extracellular matrix; GWAS = genome-wide association study; KoGES = Korean Genome and Epidemiology Study; LD = linkage disequilibrium; MAF = minor allele frequency; MMP = metalloproteinase; OA = osteoarthritis; OR = odds ratio; Q-Q plot = quantile-quantile plot; SNP= single nucleotide polymorphism; TIMP2 = tissue inhibitor of metalloproteinase 2.

Acknowledgements

References

Abel, K., Reneland, R., Kammerer, S., Mah, S., Hoyal, C., Cantor, C.R., Nelson, M.R., and Braun, A. (2006). Genome-wide SNP association: identification of susceptibility alleles for osteoarthritis. *Autoimmun. Rev.* 5, 258-263.

Cawston, T. (1998). Matrix metalloproteinases and TIMPs: properties and implications for the rheumatic diseases. *Mol. Med. Today* 4, 130-137.

Cho, N., Joo, S., Kim, J., Abbott, R.D., Kimm, K., and Shin, C. (2006). Relation of habitual snoring with components of metabolic syndrome in Korean adults. *Diabetes Res. Clin Pract.* 71, 256-263.

Cho, Y.S., Go, M.J., Kim, Y.J., Heo, J.Y., Oh, J.H., Ban, H.J., Yoon, D., Lee, M.H., Kim, D.J., Park, M., Cha, S.H., Kim, J.W., Han, B.G., Min, H., Ahn, Y., Park, M.S., Han, H.R., Jang, H.Y., Cho, E.Y., Lee, J.E., Cho, N.H., Shin, C., Park, T., Park, J.W., Lee, J.K., Cardon, L., Clarke, G., McCarthy, M.I., Lee, J.Y., Oh, B., and Kim, H.L. (2009). A large-scale genome-wide association study of Asian populations uncovers genetic factors influencing eight quantitative traits. *Nat. Genet.* 41, 527-534.

Clayton, D.G., Walker, N.M., Smyth, D.J., Pask, R., Cooper, J.D., Maier, L.M., Smink, L.J., Lam, A.C., Ovington, N.R., Stevens, H.E., Nutland, S., Howson, J.M., Faham, M., Moorhead, M., Jones, H.B., Falkowski, M., Hardenbol, P., Willis, T.D., and Todd, J.A. (2005). Population structure, differential bias and genomic control in a large-scale, case-control association study. *Nat. Genet.* 37, 1243-1246.

Dai, J., and Ikegawa, S. (2010). Recent advances in association studies of osteoarthritis susceptibility genes. *J. Hum. Genet.* 55, 77-80.

Du, H., Chen, S.L., Bao, C.D., Wang, X.D., Lu, Y., Gu, Y.Y., Xu, J.R., Chai, W.M., Chen, J., Nakamura, H., and Nishioka, K. (2005). Prevalence and risk factors of knee osteoarthritis in Huang-Pu District, Shanghai, China.

Rheumatol. Int. 25, 585-590.

Francis, M.J., Lees, R.L., Trujillo, E., Martin-Vasallo, P., Heersche, J.N., and Mobasheri, A. (2002). ATPase pumps in osteoclasts and osteoblasts. *Int. J. Biochem. Cell. Biol.* 34, 459-476.

Haq, I., Murphy, E., and Dacre, J. (2003). Osteoarthritis. *Postgrad. Med. J.* 79, 377-383.

Ikegawa, S. (2007). New gene associations in osteoarthritis: what do they provide, and where are we going? *Curr. Opin. Rheumatol.* 19, 429-434.

Jiang, Q., Shi, D., Nakajima, M., Dai, J., Wei, J., Malizos, K.N., Qin, J., Miyamoto, Y., Kamatani, N., Liu, B., Tsezou, A., Nakamura, T., and Ikegawa, S. (2008). Lack of association of single nucleotide polymorphism in LRCH1 with knee osteoarthritis susceptibility. *J. Hum. Genet.* 53, 42-47.

Kellgren, J.H., and Lawrence, J.S. (1957). Radiological assessment of osteo-arthrosis. *Ann. Rheum. Dis.* 16, 494-502.

Kerkhof, J.M., Uitterlinden, A.G., Valdes, A.M., Hart, D.J., Rivadeneira, F., Jhamai, M., Hofman, A., Pols, H.A., Bierma-Zeinstra, S.M., Spector, T.D., and van Meurs, J.B. (2008). Radiographic osteoarthritis at three joint sites and FRZB, LRP5, and LRP6 polymorphisms in two population-based cohorts. *Osteoarthritis. Cartilage.* 16, 1141-1149.

Kim, B.G., Park, J.T., Ahn, Y., Kimm, K., and Shin, C. (2005). Geographical difference in the prevalence of isolated systolic hypertension in middle-aged men and women in Korea: the Korean Health and Genome Study. *J. Hum. Hypertens.* 19, 877-883.

Krane, S.M., and Inada, M. (2008). Matrix metalloproteinases and bone. *Bone.* 43, 7-18.

Lee, H.J., Lee, G.H., Nah, S., Lee, K.H., Yang, H., Kim, Y.M., Chun, W., Hong, S., and Kim, S. (2008). Association of TIMP-4 gene polymorphism with the risk of osteoarthritis in the Korean population. *Rheumatol. Int.* 28, 845-850.

Lee, P.H., and Shatkay, H. (2009). An integrative scoring system for ranking SNPs by their potential deleterious effects. *Bioinformatics* 25, 1048-1055.

Lee, Y.J., Lee, E.B., Kwon, Y.E., Lee, J.J., Cho, W.S., Kim, H.A., and Song, Y.W. (2003). Effect of estrogen on the expression of matrix metalloproteinase (MMP)-1, MMP-3, and MMP-13 and tissue inhibitor of metalloproternase-1 in osteoarthritis chondrocytes. *Rheumatol. Int.* 23, 282-288.

Lim, S., Jang, H.C., Lee, H.K., Kimm, KC., Park, C., and Cho, N.H. (2006). A rural-urban comparison of the characteristics of the metabolic syndrome by gender in Korea: the Korean Health and Genome Study (KHGS). *J. Endocrinol. Invest.* 29, 313-319.

Loughlin, J. (2002). Genome studies and linkage in primary osteoarthritis. *Rheum. Dis. Clin. North. Am.* 28, 95-109.

MacGregor, A.J., and Spector, T.D. (1999). Twins and the genetic architecture of osteoarthritis. *Rheumatology.* 38, 583-588.

Martel-Pelletier, J., McCollum, R., Fujimoto, N., Obata, K., Cloutier, J.M., and Pelletier, J.P. (1994). Excess of metal-

loproteases over tissue inhibitor of metalloprotease may contribute to cartilage degradation in osteoarthritis and rheumatoid arthritis. *Lab. Invest.* 70, 807-815.

Miyamoto, Y., Mabuchi, A., Shi, D., Kubo, T., Takatori, Y., Saito, S., Fujioka, M., Sudo, A., Uchida, A., Yamamoto, S., Ozaki, K., Takigawa, M., Tanaka, T., Nakamura, Y., Jiang, Q., and Ikegawa, S. (2007). A functional polymorphism in the 5' UTR of GDF5 is associated with susceptibility to osteoarthritis. *Nat. Genet.* 39, 529-533.

Mobasheri, A., Avila, J., Cozar-Castellano, I., Brownleader, M.D., Trevan, M., Francis, M.J., Lamb, J.F., and Martin-Vasallo, P. (2000). Na+, K+-ATPase isozyme diversity; comparative biochemistry and physiological implications of novel functional interactions. *Biosci. Rep.* 20, 51-91.

Mototani, H., Mabuchi, A., Saito, S., Fujioka, M., Iida, A., Takatori, Y., Kotani, A., Kubo, T., Nakamura, K., Sekine, A., Murakami, Y., Tsunoda, T., Notoya, K., Nakamura, Y., and Ikegawa, S. (2005). A functional single nucleotide polymorphism in the core promoter region of CALM1 is associated with hip osteoarthritis in Japanese. *Hum. Mol. Genet.* 14, 1009-1017.

Peach, C.A., Carr, A.J., and Loughlin, J. (2005). Recent advances in the genetic investigation of osteoarthritis. *Trends. Mol. Med.* 11, 186-191.

Rabbee, N., and Speed, T.P. (2006). A genotype calling algorithm for affymetrix SNP arrays. *Bioinformatics* 22, 7-12.

Rodriguez-Lopez, J., Mustafa, Z., Pombo-Suarez, M., Malizos, K.N., Rego, I., Blanco, F.J., Tsezou, A., Loughlin, J., Gomez-Reino, J.J., and Gonzalez, A. (2008). Genetic variation including nonsynonymous polymorphisms of a major aggrecanase, ADAMTS-5, in susceptibility to osteoarthritis. *Arthritis. Rheum.* 58, 435-441.

Shi, D., Ni, H., Dai, J., Qin, J., Xu, Y., Zhu, L., Yao, C., Shao, Z., Chen, D., Xu, Z., Yi, L., Ikegawa, S., and Jiang, Q. (2008). Lack of association between the CALM1 core promoter polymorphism (-16C/T) and susceptibility to knee osteoarthritis in a Chinese Han population. *BMC Med. Genet.* 9, 91.

Snelling, S., Sinsheimer, J.S., Carr, A., and Loughlin, J. (2007). Genetic association analysis of LRCH1 as an osteoarthritis susceptibility locus. *Rheumatology* 46, 250-252.

Spector, T.D., Reneland, R.H., Mah, S., Valdes, A.M., Hart, D.J., Kammerer, S., Langdown, M., Hoyal, C.R., Atienza, J., Doherty, M., Rahman, P., Nelson, M.R., and Braun, A. (2006). Association between a variation in LRCH1 and knee osteoarthritis: a genome-wide single-nucleotide polymorphism association study using DNA pooling. *Arthritis. Rheum.* 54, 524-532.

Szklarczyk, D., Franceschini, A., Kuhn, M., Simonovic, M., Roth, A., Minguez, P., Doerks, T., Stark, M., Muller, J., Bork, P., Jensen, L.J., and von Mering, C. (2011). The STRING database in 2011: functional interaction networks of proteins, globally integrated and scored. *Nucleic. Acids. Res.* 39, 561-568.

Valdes, A.M., Loughlin, J., Oene, M.V., Chapman, K., Surdulescu, G.L., Doherty, M., and Spector, T.D. (2007).

Sex and ethnic differences in the association of ASPN, CALM1, COL2A1, COMP, and FRZB with genetic susceptibility to osteoarthritis of the knee. *Arthritis. Rheum.* 56, 137-146.

Weiss, K.E., and Rodner, C.M. (2007). Osteoarthritis of the wrist. *J. Hand. Surg. Am.* 32, 725-746.

Yuan, H.Y., Chiou, J.J., Tseng, W.H., Liu, C.H., Liu, C.K., Lin, Y.J., Wang, H.H., Yao, A., Chen, Y.T., and Hsu, C.N. (2006). FASTSNP: an always up-to-date and extendable service for SNP function analysis and prioritization. *Nucleic. Acids. Res.* 34, 635-641.