SciMedCentral

JSM Arthritis

Editorial

Can Genetic Studies Identify the Target Molecules of Osteoarthritis?

Hiroshi Kawaguchi*

The Chief of the Spine Center, Japan Community Health Care Organization (JCHO), Japan

***Corresponding author**
Hiroshi Kawaguchi, The Chief of the Spine Center, Japan Community Health Care Organization (JCHO), Tsukudo 5-1, Tokyo 162-8543, Japan, Tel: 03-3269-8111; Fax: 03-3260-7840; E-mail: kawaguchi0126@gmail.com

Submitted: 16 May 2016

Accepted: 17 May 2016

Published: 19 May 2016

Despite high prevalence and social impact, osteoarthritis (OA) is far behind other skeletal diseases like osteoporosis in the development of disease-modifying treatments. This is mainly because little is known about the underlying molecular mechanism that could be the therapeutic target. Since OA is a multifactorial disease caused by complex interplay between environmental and genetic factors with estimates of around 50% heritability depending on the site [1], numerous efforts and great expense have been spent on human genetic studies on OA worldwide. Although linkage studies have shown large areas of chromosomes associated with the disease, they have failed to detect the susceptible genes. Candidate gene studies have proposed over 100 genes as being responsible; however, most of them have not later been reproduced in larger meta-analysis studies. Recently, while genome-wide association studies (GWAS) have led to the discovery of over 600 gene loci in over 50 common multifactorial diseases, most of the gene variants are of only minimal individual effect. Even though the identified genes with such small effect sizes could possibly be therapeutic targets or at least prognostic markers, it is questionable whether or not these conventional OA genetic studies are worthy of such enormous investment. Aiming at a well-powered approach for this highly polygenic disease with multiple risk loci conferring small effects, consortium studies like Treat-OA and arcOGEN have been developed to enlarge the sample size. Considering the disease characteristics and prevalence, however, it is our opinion that not only the quantity but also the quality of studies is critical for identification of the genetic architecture. In this sense, the conventional OA genetic studies do not seem to us who are clinicians, although not genetic experts, to have been performed with sufficient scientific strictness, even as compared to those on other common diseases.

Several studies indicate that inconsistent and ambiguous definition of OA is a critical limitation of conventional genetic studies [2]. In addition to the stringency of disease definition raised by them, here we propose two other capital issues in the conventional studies: selection of appropriate controls and adjustment for environmental/clinical factors, from a clinician's point of view.

Stringency of disease definition

Although most conventional genetic studies determine OA on radiographs as Kellgren-Lawrence (KL) score =2 or higher (Table 1) [3-7], the KL grading is limited in reproducibility and sensitivity due to the subjective judgment of observers and the categorical classification into only a five-grade scale [8]. In the ROAD (Research on Osteoarthritis against Disability) study with a high-quality population-based cohort database of detailed environmental and genetic information of more than 3,000 participants [9], we delete the intermediate and ambiguous KL=2 subgroup for the case-control analysis to increase the detection power. For example, our association analysis of the *EPAS1* gene which was identified to be crucial for OA development in mice was able to detect a significant difference of the minor allelic frequency (mAF) of a SNP in the gene between KL=3 & 4 (case; mAF=11.1%) and KL=0 & 1 (control; mAF=15.2%) [10]. The mAF of the omitted KL=2 subgroup was 12.3%, confirming an inverse relationship between mAF of the SNP and KL scores. This clearly indicates that inclusion of the KL=2 subjects in the case group had caused a decrease in the detection power. In fact, this association was not reproduced by conventional Japanese and Chinese studies that include KL=2 in the case group [11]. Considering that prevalence of the KL=2 subgroup is shown to be fairly high in representative epidemiologic studies (17.3-41.3%; difference between KL≥2 and KL≥3 in Table 2), removal of this subgroup may inevitably cause a decrease in the total sample size.

Generally, a lack of objective and quantitative measure for the disease definition remains a fatal limitation of clinical OA studies. The ROAD study has recently established the fully automatic program KOACAD (knee OA computer-aided diagnosis) to quantify the major OA parameters (joint space, osteophyte, etc.) on plain radiographs [8]. We believe that the KOACAD system as well as magnetic resonance image systems [12] will serve as optimal measures for the definition of OA in the near future, just as bone mineral density does in osteoporosis.

Selection of appropriate controls

In genetic studies on common diseases with a high prevalence, selection of disease-free controls is essential to avoid the potential bias due to contamination of affected subjects in the control. In representative epidemiologic studies worldwide, the prevalence of radiographic knee OA (KL≥2) in the elderly was ≥30% in all populations and >60% in Asian populations like Japan

Table 1: Source of subjects, definition, and adjustment for confounders in representative knee OA genetic studies.

Gene Reference	GDF5 3	PTGS2 4	DVWA 5	Chromosome 7q22 6	HLA Class II/III 7
	Candidate	GWAS	GWAS	GWAS	GWAS
Discovery population (Source (N), definition, mean age, %female)	**Case** PBC+HP(718) KL≥2, CL 72y, 83% **Control** PBC+HP(861) KL≤1, OR 49y, 54%	**Case** HP(243), PBC(114) KL≥2, CL NA, 100% **Control** PBC(196), HS(89) KL≤1 NA, 100%; NA, 87%	**Case** HP(740) CL 72y, 90%; 72y, 82% **Control** HP(1,289) OR 49y, 44%; 54y, 46%	**Case** PBC(698) KL≥2 NA, NA **Control** PBC(1,893) KL≤1 NA, NA	**Case** HP(899) CL 72y, 84% **Control** HP+HS(3,396) OD+NA 53y, 44%
Replication population (Source (N), definition, mean age, %female)	**Case** HP(313) CL 59y, 66% **Control** HS(485) NS 57y, 65%	**Case** PBC(647), HP(530) KL≥2, CL NA, 100% **Control** PBC(1,712), HS(660) KL≤1, ND NA, 100%	**Case** HP(417), PBC(242) KL≥2, CL 71y, 75%; 60y, 70% **Control** PBC(485), HS(413) KL≤1, NS 68y, 63%; 56y, 74%	**Case** HP(3,142), PBC(741) KL≥2, CL, TKR NA, NA **Control** HS(33,825), PBC(2,718), HP(294) KL≤1, KL=0, NS, NA NA, NA	**Case** HP(813), PBC(167) KL≥2, TKR 74y, 74%; 68y, 81% 66y, 82% **Control** HS(1,071), PBC(347) KL≤1, KL=0, NS 66y, 64%; 68y, 39% 60y, 65%
Adjustment	Population	Gender, population	Population	Gender, population	Population

GDF5: growth differentiation factor 5. PTGS2: prostaglandin- endoperoxide synthase 2. DVWA: double von Willebrand factor A domains. HLA: human leukocyte antigen.
HP: hospital patients. PBC: population-based cohort. HS: healthy subjects. CL: clinical diagnosis. TKR: total knee replacement. OR: orthopaedic disease or injury. OD: other disease than OA. ND: not diagnosed for OA. NS: no sign of OA. NA: not available.

(ROAD study) and China (Shanghai) (Table 2) [13]. Furthermore, the prevalence of asymptomatic knee OA was 24-36% in all populations. Hence, if so-called healthy subjects without knee symptoms were collected as controls, a considerable number of OA subjects would be included in the control group. Even in a series of genetic studies in Japan with a high OA prevalence [13], the control subjects are miscellaneous mixtures of various populations including considerable numbers of so-called healthy subjects and other disease patients without radiographic diagnosis (Table 1) [3,5,7], indicating that a substantial percentage in the control groups are affected subjects. A recent analysis of the effect of controls selected with different levels of stringency on the association of known knee OA susceptibility genes demonstrates that a control with poor selection or without selection cannot be compensated by increase of the sample size [14]. Hence, selection of appropriate controls confirmed to be disease-free may be crucial to achieve a high detection power.

Adjustment for confounding environmental/clinical factors

Lastly, we should again note that OA is a multifactorial disease with environmental and genetic backgrounds and that the genetic contribution is less than half in knee OA [1]. A report by Takahashi et al. constructed knee OA prediction models based on genotype (combination of three risk alleles of ASPN, GDF5 and DVWA) and environmental/clinical information (age, gender and BMI), and evaluated the predictive power by area under the curve (AUC; range, 0.5 [worst] to 1 [best]) on a receiver operating characteristic (ROC) curve in a case-control association study [15]. The result was that the power by the genotype information was very small (AUC=0.554), implicating uselessness of the three famous genotypes as a prognostic marker. Contrarily, the environmental/clinical information was a much better predictor (AUC=0.678), but was little improved by the combination with the genotype information (AUC=0.685), again confirming its uselessness. Hence, to achieve a high detection power for the susceptibility gene, all efforts should be made to exclude the influence of environmental/clinical factors. Surprisingly, however, there are big differences in age and gender between case and control groups in previous representative studies (Table 1). Even a sole difference in age of about 20 years between case and control groups that is seen in the Japanese studies [3,5,7] is calculated to cause an increase of odds ratio for OA to 2.65 ($=1.05^{20}$), according to the authors' own estimation (1.05 / year) [15]. Indeed, we are not opposed to recent activities of OA consortiums to pool subjects worldwide; however, we should note that the pooled subjects are miscellaneous mixtures of various populations with different backgrounds. Selection of case and control subjects with similar backgrounds is essential to minimize selection bias which strongly influences the results in genetic studies with small effect sizes of the risk alleles. Hence, at least for the initial screening, case and control groups should be selected from a single population-based cohort to adjust the living environment and stratified by confounding environmental/clinical factors which have been identified in preparatory epidemiologic analysis in the cohort. The reproducibility may then be examined in other replication cohorts of the worldwide consortiums, after adjustment for the specific confounding factors in the respective cohorts.

Taken together, conventional OA genetic studies appear to

Table 2: Prevalence of radiographic knee OA in representative population-based cohorts.

Cohort	ROAD	Framingham	Zoetermeer	Johnston County	Beijing	Shanghai	NHANES III
Ethnicity	Japan	White in USA	Netherlands	Black & whites in USA	China	China	Black & whites in USA
Age	≥ 60	≥ 63	≥ 60	≥ 65	≥ 60	60-69	≥ 60
Total number	2,282	1,420	1,123	1,175	1,781	700	2,415
Radiographic knee OA (%)							
KL ≥ 2	61.9	33	30	40.6	38.8	64.1	37.4
KL ≥ 2 (symptomatic)	26.1	9.5		13.6	12		12.1
KL ≥ 2 (asymptomatic)	35.8	23.5		27.0	26.8		25.3
KL ≥ 3	20.6	15.7	10.2	13.6			10.2

KL≥ 2 (asymptomatic) was defined as KL≥ 2 (radiographic) but KL≥ 2 (symptomatic).
References: Osteoarthritis Cartilage 2009; 17: 1137 (ROAD). Arthritis Rheum 1987; 30: 914 (Framingham). Ann Rheum Dis 1989;48:271 (Zoetermeer). J Rheumatol 2007; 34:172 (Johnston County). Arthritis Rheum 001;44:2065 (Beijing). Rheumatol Int 2005;25:585 (Shanghai). J Rheumatol 2006;33:2271 (NHANES III)

compare a case group containing a substantial number of subjects with ambiguous definition versus a control group containing a substantial number of affected subjects, plus without adjustment for confounding environmental/clinical factors. Contrary to the genetic studies, studies of clinical trial and observational epidemiology are performed under a sound scientific rigidity in compliance with very strict rules to examine the accurate effect sizes of interventions and environmental/clinical factors, respectively. Introduction of strict regulation in the genetic field, just like CONSORT guidelines in the clinical trial field [16], might improve the scientific rigidity.

REFERENCES

1. Spector TD, MacGregor AJ. Risk factors for osteoarthritis: genetics. Osteoarthritis Cartilage. 2004; 12 Suppl A: S39-44.
2. Kerkhof HJ, Meulenbelt I, Akune T, Arden NK, Aromaa A, Bierma-Zeinstra SM, et al. Recommendations for standardization and phenotype definitions in genetic studies of osteoarthritis: the TREAT-OA consortium. Osteoarthritis Cartilage. 2011; 19: 254-264.
3. Miyamoto Y, Mabuchi A, Shi D, Kubo T, Takatori Y, Saito S . A functional polymorphism in the 5' UTR of GDF5 is associated with susceptibility to osteoarthritis. Nat Genet. 2007; 39: 529-533.
4. Valdes AM, Loughlin J, Timms KM, van Meurs JJ, Southam L, Wilson SG. Genome-wide association scan identifies a prostaglandin-endoperoxide synthase 2 variant involved in risk of knee osteoarthritis. Am J Hum Genet. 2008; 82: 1231-1240.
5. Miyamoto Y, Shi D, Nakajima M, Ozaki K, Sudo A, Kotani A, et al. Common variants in DVWA on chromosome 3p24.3 are associated with susceptibility to knee osteoarthritis. Nat Genet. 2008; 40: 994-998.
6. Kerkhof HJ, Lories RJ, Meulenbelt I, Jonsdottir I, Valdes AM, Arp P, et al. A genome-wide association study identifies an osteoarthritis susceptibility locus on chromosome 7q22. Arthritis Rheum. 2010; 62: 499-510.
7. Nakajima M, Takahashi A, Kou I, Rodriguez-Fontenla C, Gomez-Reino JJ, Furuichi T, et al. New sequence variants in HLA class II/III region associated with susceptibility to knee osteoarthritis identified by genome-wide association study. PLoS One. 2010; 18: e9723.
8. Oka H, Muraki S, Akune T, Mabuchi A, Suzuki T, Yoshida H. Fully automatic quantification of knee osteoarthritis severity on plain radiographs. Osteoarthritis Cartilage. 2008; 16: 1300-1306.
9. Yoshimura N, Muraki S, Oka H, Kawaguchi H, Nakamura K, Akune T. Cohort Profile: Research on Osteoarthritis/Osteoporosis against Disability study. Int J Epidemiol. 2010; 39: 988-995.
10. Saito T, Fukai A, Mabuchi A, Ikeda T, Yano F, Ohba S, et al. Transcriptional regulation of endochondral ossification by HIF-2a during skeletal growth and osteoarthritis development. Nat Med. 2010; 16: 678-686.
11. Nakajima M, Shi D, Dai J, Tsezou A, Zheng M, Norman PE. Replication studies in various ethnic populations do not support the association of the HIF-2Î± SNP rs17039192 with knee osteoarthritis. Nat Med. 2011; 17: 26-27.
12. Eckstein F, Guermazi A, Roemer FW. Quantitative MR imaging of cartilage and trabecular bone in osteoarthritis. Radiol Clin North Am. 2009; 47: 655-673.
13. Muraki S, Oka H, Akune T, Mabuchi A, En-yo Y, Yoshida M. Prevalence of radiographic knee osteoarthritis and its association with knee pain in the elderly of Japanese population-based cohorts: the ROAD study. Osteoarthritis Cartilage. 2009; 17: 1137-1143.
14. Rodriguez-Fontenla C, López-Golán Y, Calaza M, Pombo-Suarez M, Gómez-Reino JJ, González A. Genetic risk load and age at symptom onset of knee osteoarthritis. J Orthop Res. 2012; 30: 905-909.
15. Takahashi H, Nakajima M, Ozaki K, Tanaka T, Kamatani N, Ikegawa S. Prediction model for knee osteoarthritis based on genetic and clinical information. Arthritis Res Ther. 2010; 12: R187.
16. Schulz KF, Altman DG, Moher D; CONSORT Group. CONSORT 2010 Statement: updated guidelines for reporting parallel group randomised trials. BMC Med. 2010; 8: 18-26.

Cite this article

Kawaguchi H (2016) Can Genetic Studies Identify the Target Molecules of Osteoarthritis? JSM Arthritis 1(1): 1002.