**Loughlin J. Genetic contribution to osteoarthritis development: current state of evidence. *Current Opinion in Rheumatology* 2015, 27(3), 284-288.**



**DOI link to article:**

http://dx.doi.org/10.1097/BOR.0000000000000171

**Date deposited:**

06/05/2015

REVIEW

OPEN

# Genetic contribution to osteoarthritis development: current state of evidence

*John Loughlin*

**Purpose of review**
Powerful association studies have identified a number of genetic signals that can be confidently judged as associated with osteoarthritis. Efforts have continued to discover new loci, whilst functional studies are being applied to assess which genes are the likely targets of the risk-conferring alleles. The study of epigenetics has highlighted an interaction between osteoarthritis genetics and DNA methylation. This review will summarize some of the recent key studies in osteoarthritis genetics, including functional and epigenetic analyses.

**Recent findings**
Several novel osteoarthritis susceptibility loci have been reported recently, including the regulatory genes *NCOA3* and *ALDH1A2*. Functional analyses of these genes and of others reported previously support earlier suggestions that osteoarthritis susceptibility is principally mediated by modulations to gene expression. DNA methylation analyses provide additional insights into the osteoarthritis disease process, at both a genome-wide level and when investigating direct interactions with risk-conferring alleles.

**Summary**
Osteoarthritis genetic risk predominantly acts by modulating gene expression, an effect typically mediated via transcriptional regulation. Effects on various pathways have been detected, including cell differentiation and cartilage homeostasis. The continued identification of risk loci, their functional study, and the unification of genetic and epigenetic analyses will be key themes in the future.

**Keywords**
*ALDH1A2*, DNA methylation, genetic association, *NCOA3*, single-nucleotide polymorphism

## INTRODUCTION

Epidemiological investigations of the past, including twin-pair analyses and family-based segregation studies, provided clear evidence of a heritable component to osteoarthritis susceptibility. These studies stimulated the initial search for osteoarthritis risk alleles. Osteoarthritis susceptibility is coded for by DNA polymorphism [principally single-nucleotide polymorphisms (SNPs)] and as such these studies involved a comparison of the frequency of polymorphisms in individuals with the disease (cases) versus those without (disease-free controls), or those in whom osteoarthritis status had not been determined (population controls). Such case-control studies have tremendous power, so long as certain criteria are met. These include: an analysis of an adequate number of cases and controls, to account for the relatively modest effect individual risk alleles have on disease cause; the analysis of an appropriate number of polymorphisms, to account for the extensive genetic variation within the genome; the use of cases and controls drawn from the same ethnic group, to avoid stratification (and false-positives) resulting from the natural variation in gene frequencies between different population groups; and the use of an unambiguous disease phenotype. This latter requirement is one that particularly taxes osteoarthritis researchers, with numerous choices available, ranging from symptom-free radiographic disease through to severe, painful end-stage osteoarthritis necessitating joint replacement. The

Institute of Cellular Medicine, Newcastle University, Newcastle upon Tyne, UK

Correspondence to Professor John Loughlin, Newcastle University, Institute of Cellular Medicine, Musculoskeletal Research Group, 4th Floor Catherine Cookson Building, The Medical School, Framlington Place, Newcastle upon Tyne NE2 4HH, UK. Tel: +44 191 208 7178; e-mail: john.loughlin@ncl.ac.uk

**Curr Opin Rheumatol** 2015, 27:284–288

DOI:10.1097/BOR.0000000000000171

## KEY POINTS

- Osteoarthritis genetic risk loci have small individual effects, but they can be identified so long as cohort sample size is adequate and the phenotype is robust.
- Effects on gene expression appear to be the common mechanism by which osteoarthritis risk-conferring alleles mediate their effect.
- A range of informative functional tools are now available to osteoarthritis geneticists, including access to relevant cells and tissues, removing any excuses for not following up interesting genetic hits.
- Epigenetic analyses, in particular DNA methylation studies, are clarifying how genetic effects interact with, and are modulated by, the genome and the environment. This may be the most exciting discovery to have emerged in the past year or two.

skeletal site to be studied is also a source of debate; disease of the knee or hip has dominated in the past decade, but the hand is now making a comeback, whilst other skeletal sites remain under-investigated. The past decade has seen a number of reasonably powered osteoarthritis case-control studies, in the form of candidate gene-based analyses or genome-wide association studies (GWAS). In the past few years, the focus has started to shift to functional analyses of genes implicated from these genetic studies, and to an analysis of the role of epigenetics in osteoarthritis and its interaction with risk-conferring alleles. This review will touch on these areas, with a particular focus on novel and insightful discoveries from the past 18 months.

## OSTEOARTHRITIS GENETIC ASSOCIATION STUDIES

The most powerful osteoarthritis GWAS performed to date was published in 2012 and entailed an analysis of over 7400 osteoarthritis cases [1]. This study, termed arcOGEN, was the first GWAS to report multiple, independent association signals that replicated at a level considered significant after accounting for the multiple tests that are performed in a GWAS ($P$ value $< 5 \times 10^{-8}$). arcOGEN investigated patients with hip or knee osteoarthritis and clearly demonstrated that, at the molecular genetic level, osteoarthritis risk is not uniform between these two skeletal sites: a polymorphism can be associated with disease at the hip without showing any evidence of association at the knee. Epidemiological studies are also revealing that osteoarthritis pathophysiology is not uniform across the skeleton [2]. As such, genetic studies are substantiating these finds. A bioinformatics analysis of the association signals that emerged from arcOGEN, and from less powerful GWAS and candidate gene-based studies, highlighted a number of biological pathways through which osteoarthritis genetic susceptibility is operating [3]. These include skeletogenesis, encompassing development and differentiation of osteoblasts and chondrocytes, with transcriptional regulation and cell signaling key. Since the arcOGEN publication, there have been four additional osteoarthritis genetic association papers published in which the investigators have utilized large cohorts and which have added to our understanding of the molecular basis of this disease.

## FOUR RECENT EXAMPLES

A meta-analysis of over 78 000 Europeans has identified a genome-wide significant signal associated with hip osteoarthritis [4▪▪]. The signal, at chromosome 20q13, is marked by SNP rs6094710, with a genome-wide significant $P$ value of $7.9 \times 10^{-9}$ and an odds ratio (OR) of 1.28. This SNP is located upstream of two very plausible candidate genes for osteoarthritis: *NCOA3,* which codes for nuclear receptor coactivator 3, and *SULF2*, which codes for an extracellular heparan sulfate 6-O endosulfatase 2. Both genes are expressed in cartilage, with expression of *NCOA3* reduced in osteoarthritis-affected cartilage compared with preserved cartilage from the same joint [4▪▪], and with *SULF2* expression increased in osteoarthritis cartilage as compared with normal cartilage [5]. rs6094710 is in perfect linkage disequilibrium ($r^2 = 1$, $D' = 1$) with several other SNPs within the 20q13 region; these SNPs co-segregate without recombination and as such it is not possible, genetically, to determine which SNP is responsible for the association signal. Instead, direct functional investigations of the SNPs will be required. One of the SNPs, rs6094752, codes for an amino acid change within the NCOA3 protein. However, this change is benign when examined using protein prediction tools. No other SNPs alter amino acid sequence of the NCOA3 or SULF2 proteins, and as such it appears likely that the association signal acts by modulating the expression of *NCOA3* or *SULF2*, or possibly both genes.

The first powerful GWAS of hand osteoarthritis has identified a signal at chromosome 15q22, marked by SNP rs3204689, with a genome-wide significant $P$ value of $1.1 \times 10^{-11}$ and an OR of 1.46 [6▪▪]. This SNP is located in the 3′ untranslated region (3′UTR) of the gene *ALDH1A2*, which codes for an enzyme involved in retinoic acid synthesis. As the authors note in their report, retinoic acid is a signaling molecule involved in a variety of

developmental pathways, including skeletogenesis. Intriguingly, rs3204689 was not associated with osteoarthritis at other skeletal sites despite an expression analysis of *ALDH1A2*, revealing that the hand osteoarthritis risk-conferring C-allele of the SNP correlates with reduced expression of the gene in knee and hip cartilage from osteoarthritis patients. Assuming that this functional effect is not coincidental, it would therefore represent an example of a functional effect (altered expression of *ALDH1A2* in knee and hip cartilage) not translating into a genetic risk, once again highlighting the differing osteoarthritis pathophysiology between different skeletal sites. To be confident that reduced expression of *ALDH1A2* is in fact causal for hand osteoarthritis, an expression analysis of the gene in relevant tissues from the finger joints of osteoarthritis patients will be required.

A second recent hip osteoarthritis GWAS reported a signal at chromosome 7p12.3, marked by SNP rs788748, with a genome-wide significant *P* value of $2.0 \times 10^{-8}$ and an OR of 0.71 [7]. Unlike the above two reports, replication analysis performed by the investigators led to the signal becoming weaker rather than stronger, with a meta-analysis *P* value of $1.0 \times 10^{-6}$. A less significant *P* value following meta-analysis is very often the sign of a false-positive. If that is the case, it may have derived from the small sample size used in the GWAS (only 654 cases), which will have limited its power to detect a genuine signal. The investigators nevertheless performed a number of functional analyses on the gene located closest to rs788748, *IGFBP3*, which codes for insulin-like growth factor-binding protein 3. They discovered that knockdown and overexpression of IGFBP3 protein had measurable effects compared to control treatments in several chondrogenesis models, and reported that genotype at rs788748 correlated with levels of IGFBP3 in the blood. The question, however, is whether these functional effects are enough to assuage concerns about the reduced significance of the association signal after replication. IGFBP3 clearly has a functional role, but that does not automatically mean that its gene should be the repository of osteoarthritis genetic risk. An assessment of the association of rs788748 in additional osteoarthritis cohorts, including the large arcOGEN study, would therefore be wise.

The final example is a meta-analysis of European osteoarthritis candidate gene studies [8▪]. Before GWAS became an affordable option, investigators would choose their favorite gene and subject it to a genetic association analysis. These studies would often be performed on case-control cohorts of modest size (typically only a few hundred individuals) and with a small number of polymorphisms tested. Osteoarthritis researchers succumbed to this temptation as much as investigators in other disease areas. Despite the fact that the majority of such studies found no compelling evidence for an association, there always seemed to be enough marginally positive data to ensure some investigators kept the faith. In fact, there has only been one osteoarthritis candidate that has convincingly stood the test of repeated analysis: *GDF5*, whose story has been told before [9]. In the recent candidate analysis, the investigates studied published reports from another 199 osteoarthritis candidates and confirmed that none was overly compelling, with only the type XI collagen gene *COL11A1* and the vascular endothelial growth factor gene *VEGF* demonstrating any semblance of association. However, even their *P* values were quite modest compared with the signals that have emerged from GWAS studies, with values greater than $1 \times 10^{-6}$. This candidate study has therefore reinforced the need to maintain a hypothesis-free approach when trying to identify osteoarthritis risk loci. It also reminds us that our comprehension of the disease is still somewhat nascent.

In summary, these four recent reports have added to the genes that can be confidently considered as repositories of osteoarthritis genetic risk (*NCOA3*/*SULF2* and *ALDH1A2*); they have reconfirmed the observation that osteoarthritis risk is often acting by modulating gene regulation, that functional analyses can be highly insightful, but that vigilance and objectivity need to be maintained and applied when assessing the likelihood of an association signal being genuine.

## EPIGENETIC EFFECTS

Epigenetics refers to heritable changes in gene expression or phenotype that occur without changes in the underlying DNA sequence. There are three mechanisms of epigenetic gene regulation: DNA methylation of CpG dinucleotides, post-translational modifications of histone proteins, and non-coding RNAs such as microRNAs. They regulate gene expression either by affecting gene transcription (DNA methylation, histone modifications) or by acting post-transcriptionally, leading to changes in the levels of the encoded protein (e.g. microRNAs). Epigenetic patterns are both plastic, especially during development and cell differentiation when they undergo dynamic changes, and stable, allowing cellular identify to be maintained during mitotic cell divisions. A recent review provides an update on the status of epigenetics in osteoarthritis [10].

The advent of high-throughput arrays for the analysis of DNA methylation has significantly

boosted the amount of epigenetic data that has emerged in the past few years. Osteoarthritis has taken part in this enhanced activity and is in fact quite well placed to do so, for two principal reasons: the central tissue involved in the disease process, cartilage, is readily accessible via joint replacement surgery; and the cartilage contains only a single cell type, the chondrocyte, thus limiting the scope for confounding heterogeneity that can be encountered when studying the epigenetics of a multicellular tissue in which cell-specific effects are likely.

Five methylation array analyses of cartilage DNA have been published in the past 18 months, focused on knee and/or hip osteoarthritis [11▪,12–15]. A number of interesting observations have emerged. For example, knee and hip cartilages are strikingly different, with DNA methylation differences at a number of genes, including homeobox genes [12,15], which are key regulators of skeletogenesis. This is reminiscent of the joint-specific genetic effects touched on earlier, and suggests that cartilages from different parts of the skeleton are not only genetically but also epigenetically distinct. The DNA methylation studies [11▪,12–14] have also revealed that genes that have previously been implicated in osteoarthritis typically harbor CpG sites that are differentially methylated between osteoarthritis and non-osteoarthritis cartilage. This implies that the regulation of gene expression via DNA methylation is a major driver of the osteoarthritis disease process.

An alternative avenue of investigation has been to assess whether there is a direct, functional relationship between epigenetics and genetics at particular risk alleles. For example, could DNA methylation changes impact the penetrance of a nearby risk polymorphism? If the risk polymorphism modulates gene expression, then the methylation status of the surrounding DNA could attenuate or amplify the risk effect by regulating expression of the gene. This is a particularly interesting question in genetics, since epigenetic changes can occur in response to environmental factors, including diet, and during aging, leading to the suggestion that a possible mechanism for the late onset of common human diseases such as osteoarthritis is the age-related loss of normal epigenetic control. Two such studies directly investigating interactions between DNA methylation and osteoarthritis risk-conferring alleles have been reported, for the aforementioned *GDF5* and for *DIO2*.

The *GDF5* functional SNP, rs143383, is a C/T transition located in the 5'UTR of the gene. In its C-allele form, it forms a CpG site. The SNP mediates differential allelic expression (DAE) of *GDF5*, with the osteoarthritis risk-conferring T allele demonstrating reduced expression due to the more avid binding of repressor proteins, including DEAF1 [16]. rs143383 is a particular risk factor for knee osteoarthritis [17]. Functional analyses using human transformed cell lines, chondrocytes, and cartilages have highlighted that differential DNA methylation of rs143383, and of CpG sites flanking the SNP, modulates the binding of the repressor proteins and therefore alters the expression differences between the C and T alleles [18,19▪▪]. Furthermore, a CpG site located 4 base pairs upstream of rs143383 shows highly significant demethylation in osteoarthritis knee cartilage compared with osteoarthritis hip cartilage, which correlates with reduced expression of the gene [19▪▪]. This may therefore account for why rs143383 is a particular risk factor for the knee: the combination of an epigenetic effect leading to demethylation and reduced expression of *GDF5*, combined with the presence of the low-expressing T allele, may be enough to make this a more penetrant osteoarthritis allele in knee cases.

*DIO2* codes for iodothyronine-deiodinase enzyme type 2 (D2), a selenoprotein that converts intracellular inactive thyroid hormone to its active form. A common *DIO2* haplotype composed of the C-allele of SNP rs225014, and the C-allele of SNP rs12885300 is associated with osteoarthritis [20]. An analysis of *DIO2* expression has revealed that, as for *GDF5*, the gene is subject to DAE, with the C-allele of rs225014 correlating with increased expression of the gene in cartilage [21]. In a recent functional analysis of *DIO2*, including DNA methylation studies, it was reported that differential methylation of CpGs located upstream of the gene correlated with *DIO2* expression changes, and that these effects were particularly striking for individuals harboring the risk-conferring allele of rs225014 [22▪]. As for *GDF5*, DIO2 is subject to epigenetic regulation related to genotype.

## CONCLUSION

The genetic study of osteoarthritis susceptibility has now evolved from the mapping of risk loci through to their comprehensive functional analysis, combined with detailed epigenetic investigations. We have though only scratched the surface, with just a handful of risk alleles so far identified. Osteoarthritis is particularly polygenic [23] and as such large cohorts are required to map osteoarthritis susceptibility loci with confidence. The heterogeneity between skeletal sites emphasizes that each needs to be treated almost as a different genetic disease, although, as with the emerging *ALDH1A2* story, it

may be possible to do functional analyses on cells or tissues from more accessible skeletal sites that are not themselves impacted on by the risk-conferring allele. More mapping, additional functional studies, and further integration with other genome-wide approaches will continue to be the priorities for this common and particularly complex human disease.

## Acknowledgements

*The author acknowledges research support from Arthritis Research UK, the UK Medical Research Council, the Dr William Harker Foundation, the NIHR Newcastle Biomedical Research Centre, and the European Union Seventh Framework Programme (FP7/2007-2013) under grant agreement n° 305815 (D-BOARD). He is a member of the MRC–Arthritis Research UK Centre for Integrated Research into Musculoskeletal Ageing (CIMA).*

## Financial support and sponsorship

*No funding was received for this work.*

## Conflicts of interest

*There are no conflicts of interest.*

## REFERENCES AND RECOMMENDED READING

Papers of particular interest, published within the annual period of review, have been highlighted as:

- ■ of special interest
- ■■ of outstanding interest

1. arcOGEN Consortium, arcOGEN Collaborators, Zeggini E, *et al.* Identification of new susceptibility loci for osteoarthritis (arcOGEN): a genome-wide association study. Lancet 2012; 380:815–823.
2. Karlsson MK, Karlsson C, Magnusson H, *et al.* Individuals with primary osteoarthritis have different phenotypes depending on the affected joint: a case control study from southern Sweden including 514 participants. Open Orthop J 2014; 8:450–456.
3. Reynard LN, Loughlin J. Insights from human genetic studies into the pathways involved in osteoarthritis. Nat Rev Rheumatol 2013; 9:573–583.
4. ■■ Evangelou E, Kerkhof HJ, Styrkarsdottir U, *et al.* A meta-analysis of genome-wide association studies identifies novel variants associated with osteoarthritis of the hip. Ann Rheum Dis 2014; 73:2130–2136.

A powerful GWAS uncovering a novel hip osteoarthritis signal harboring two intriguing genes.

5. Otsuki S, Taniguchi N, Grogan SP, *et al.* Expression of novel extracellular sulfatases Sulf-1 and Sulf-2 in normal and osteoarthritic articular cartilage. Arthritis Res Ther 2008; 10:R61.
6. ■■ Styrkarsdottir U, Thorleifsson G, Helgadottir HT, *et al.* Severe osteoarthritis of the hand associates with common variants within the ALDH1A2 gene and with rare variants at 1p31. Nat Genet 2014; 46:498–502.

The first hand osteoarthritis GWAS that has identified a highly significant signal encompassing a gene never predicted to be involved in the osteoarthritis disease process, demonstrating the power of such a hypothesis-free approach.

7. Evans DS, Cailotto F, Parimi N, *et al.* Genome-wide association and functional studies identify a role for IGFBP3 in hip osteoarthritis. Ann Rheum Dis 2015 (in press).
8. ■ Rodriguez-Fontenla C, Calaza M, Evangelou E, *et al.* Assessment of osteoarthritis candidate genes in a meta-analysis of 9 genome-wide association scans. Arthritis Rheumatol 2014; 66:940–949.

A salient lesson regarding our limited ability to pick winners when it comes to selecting osteoarthritis candidates.

9. Reynard LN, Loughlin J. The genetics and functional analysis of primary osteoarthritis susceptibility. Expert Rev Mol Med 2013; 18:e2.
10. Barter MJ, Bui C, Young DA. Epigenetic mechanisms in cartilage and osteoarthritis: DNA methylation, histone modifications and microRNAs. Osteoarthritis Cartilage 2012; 20:339–349.
11. ■ Fernández-Tajes J, Soto-Hermida A, Vázquez-Mosquera ME, *et al.* Genome-wide DNA methylation analysis of articular chondrocytes reveals a cluster of osteoarthritic patients. Ann Rheum Dis 2014; 73:668–677.

The first array-based DNA methylation analysis of osteoarthritis cartilage, demonstrating the power of the technique for this disease.

12. Rushton MD, Reynard LN, Barter MJ, *et al.* Characterization of the cartilage DNA methylome in knee and hip osteoarthritis. Arthritis Rheumatol 2014; 66:2450–2460.
13. Jeffries MA, Donica M, Baker LW, *et al.* Genome-wide DNA methylation study identifies significant epigenomic changes in osteoarthritic cartilage. Arthritis Rheumatol 2014; 66:2804–2815.
14. Moazedi-Fuerst FC, Hofner M, Gruber G, *et al.* Epigenetic differences in human cartilage between mild and severe OA. J Orthop Res 2014; 32:1636–1645.
15. den Hollander W, Ramos YFM, Bos SD, *et al.* Knee and hip articular cartilage have distinct epigenomic landscapes: implications for future cartilage regeneration approaches. Ann Rheum Dis 2014; 73:2208–2212.
16. Syddall CM, Reynard LN, Young DA, Loughlin J. The identification of trans-acting factors that regulate the expression of GDF5 via the osteoarthritis susceptibility SNP rs143383. PLoS Genet 2013; 9:e1003557.
17. 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 Rheumatol 2009; 60:1710–1721.
18. Reynard LN, Bui C, Canty-Laird EG, *et al.* Expression of the osteoarthritis-associated gene GDF5 is modulated epigenetically by DNA methylation. Hum Mol Genet 2011; 20:3450–3460.
19. ■■ Reynard LN, Bui C, Syddall CM, Loughlin J. CpG methylation regulates allelic expression of GDF5 by modulating binding of SP1 and SP3 repressor proteins to the osteoarthritis susceptibility SNP rs143383. Hum Genet 2014; 133:1059–1073.

Linking genetics with epigenetics at the allelic level and highlighting how epigenetics can impact on the penetrance of a risk-conferring osteoarthritis allele.

20. Meulenbelt I, Min JL, Bos S, *et al.* Identification of DIO2 as a new susceptibility locus for symptomatic osteoarthritis. Hum Mol Genet 2008; 17:1867–1875.
21. Bos SD, Bovee JV, Duijnisveld BJ, *et al.* Increased type II deiodinase protein in OA-affected cartilage and allelic imbalance of OA risk polymorphism rs225014 at DIO2 in human OA joint tissues. Ann Rheum Dis 2012; 71:1254–1258.
22. ■ Bomer N, den Hollander W, Ramos YF, *et al.* Underlying molecular mechanisms of DIO2 susceptibility in symptomatic osteoarthritis. Ann Rheum Dis 2015 (in press).

A good example of the functional analysis of an OA risk-conferring locus, including epigenetic analyses.

23. 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; 70:864–867.