Biomarkers for osteoarthritis: investigation, identification, and prognosis

Guangju Zhai1,2
Erfan Aref Eshghi1
1Faculty of Medicine, Memorial University of Newfoundland, St John’s, NL, Canada; 2Department of Twin Research and Genetic Epidemiology, King’s College London, London, UK

Abstract: Osteoarthritis (OA) is the most common form of arthritis and results in substantial morbidity and disability in the elderly, imposing a great economic burden on society. While there are drugs available on the market that mitigate pain and improve function, there are no disease-modifying osteoarthritis drugs, partly because there is no reliable method that can be used to identify early OA changes. There is a pressing need to develop reliable biomarkers that can inform on the process of joint destruction in OA. Such biomarkers could aid in drug development by identifying fast progressors and detecting early response to therapy, thus reducing patient numbers and time required for clinical trials. Over the last several years, dramatic advances in our understanding of the biochemistry of cartilage have led to a cascade of studies testing proteins as biomarkers of OA. Investigation of single-nucleotide polymorphisms as genetic biomarkers and the application of technologies such as metabolomics to OA are generating potentially additional biomarkers that could help detect early OA changes. This review summarizes the data on the investigation of biochemical and genetic markers in OA and highlights the new biomarkers that are recently reported and their application and limitation in the management of OA. However, despite the dramatic growth of knowledge concerning the discovery of a number of useful biomarkers, the real breakthrough in this area is still not achieved.

Keywords: osteoarthritis, biochemical markers, metabolomics, genetics, epigenetics

Introduction

Osteoarthritis (OA) is characterized pathologically by focal areas of damage to the articular cartilage, centered on load-bearing areas, associated with new bone formation at the joint margins (osteophytosis), changes in the subchondral bone, variable degrees of mild synovitis, and thickening of the joint capsule.1 It primarily involves the joints of the knee, hip, spine, hand, and foot, is strongly associated with increasing age, and affects approximately twice as many females as males. It is the most common form of arthritis, results in substantial morbidity and disability in the elderly,2 and imposes a great economic burden on society.3 This societal burden (both in terms of personal suffering and use of health resources) is expected to increase worldwide with the increasing prevalence of obesity and the aging of the community.4

Despite high prevalence and societal impact, OA is far behind other skeletal diseases like osteoporosis in the availability of disease-modifying treatments. This is mainly because little is known about the underlying molecular mechanisms that could be exploited by therapeutic targets.5 Further, this is compounded by the inability to detect early OA changes by current evaluation methods.
Plain radiography has been by far the standard diagnostic method in OA; however, it is insufficient to determine the progression and outcome of new treatments in short timescales due to their semiquantitative grading scale and low sensitivity. Furthermore, OA changes that can be observed on X-rays are usually in more advanced stage of the disease, in which joint tissue damage is considered irreversible. Magnetic resonance imaging has the ability to visualize simultaneously all joint tissues and has been increasingly used in OA research studies, but the parameters that can be used for early diagnosis and clinical trials are still unclear.6

The field of OA study is desperately in need of biomarkers, as highlighted by the OA Biomarkers Global Initiative. A sound biomarker is OA-specific, reflects actual disease development and progression, facilitates earlier diagnosis, is sensitive to changes due to therapeutic intervention, and can predict disease outcome. Biomarkers are generally considered to be biological substances, although some researchers view imaging and even traditional disease risk factors as biomarkers. In this review, we shall consider only biological substances, which include genetic (DNA and RNA) and biochemical (carbohydrate, proteins, proteins fragments, peptides, metabolites) molecules. We shall summarize the data on the investigation of biochemicals and genetics in OA, and highlight new biomarkers that have been reported recently and their application and limitations in the management of OA.

Biochemical markers

Progressive loss of articular cartilage is a central feature of OA. Articular cartilage, a nonvascular tissue, is composed of chondrocytes embedded in an extracellular matrix (ECM), which provides the biomechanical and physiologic characteristics that are essential for articular movement.7 Type II collagen (CII) provides the major portion of the organic components (15%–22%) in the ECM, followed by aggrecan (4%–7%), and other noncollagenous proteins (0.5%–1%), including cartilage oligomeric matrix protein.8 Because an imbalance in cartilage synthesis and degradation is central to cartilage loss in OA, biomarkers reflecting these metabolic processes in the ECM have been under extensive investigation.

So far, there are eight biochemical markers concerning CII metabolism. Six of them are for CII degradation (CTX-II, Helix-II, C2C, Coll2-1, Coll2-1 NO2, TIINE) and two for CII synthesis (PIANP, PIICP).9 Among them, CTX-II was investigated most extensively despite its known limitations in that its tissue origins remain ambiguous10 and the immunoreactive epitope is not well characterized.11 Owing to its small size, it is freely filtered by the renal system and is concentrated in the urine.12 Urinary (u) CTX-II correlates well with the total body burden of osteophytes and satisfies four categories of the burden of disease, investigative, prognostic, efficacy of intervention, diagnosis of the disease, and safety of intervention (BIPED) classification scheme.13 Coll2-1 is a good disease-specific marker that is sensitive to the structural changes occurring in a single joint.14 Coll2-1 and Coll2-1 NO2 are useful for studying oxidative-related CII network degradation in OA.15 Of note, Helix-II might originate from type III collagen, rather than from CII, as thought previously.16

Aggrecan is a major proteoglycan found in the ECM and largely responsible for the high resistance to compression of this load-bearing tissue.7 Aggrecan depletion in OA cartilage can be ascribed to increased proteolytic cleavage of the core protein and is mediated by various matrix proteinases.7 Detection of aggrecanase-cleaved fragments of aggrecan in human serum and urine has been developed,17,18 but hard data are still very scant.

Cartilage oligomeric matrix protein (COMP) is a tissue-specific matrix thrombospondin-family protein that is synthesized by chondrocytes.19 It is abundant in OA cartilage, and can also be measured in serum and synovial fluids. Its concentration is ten times higher in synovial fluids than in serum, indicating preferential release from the affected joints. However, precisely what role it plays in OA pathogenesis remains unclear.20 It is the second most studied OA biomarker and satisfies four categories of BIPED. It is worthy to mention that analysis of the data, including a series of potential biomarkers measured on 137 patients with symptomatic knee OA in the Boston Osteoarthritis of the Knee Study, showed that among the six biomarker studies (C1-2C, C2C, C-propeptide, Col2CTx, Aggrecan 846 epitope, and COMP), COMP was the only marker that was a statistically significant predictor of cartilage loss even after adjustment for age, gender, and BMI.21

Further, OA is widely accepted as a disease of the whole joint affecting not only the cartilage but also the subchondral bone and the synovial membrane. Changes in these structures were found to be related to the disease and an interesting target in OA biomarker research. However, biomarkers derived from these tissues have received far less attention in the context of OA than the biomarkers from cartilage.

Hyaluronic acid, also known as hyaluronan (HA) is a glycosaminoglycan formed by alternating units of glucosamine and glucuronic acid.22 It is a constituent of synovium and
cartilage, and is thought to contribute to the lubricating mechanisms of synovial fluid. HA concentration can be measured in synovial fluid and serum. Serum HA levels correlate positively with total osteophyte burden after adjustment for age, weight, and height,\(^1\) is the third most studied OA biomarker, and satisfies four categories of BIPED.

Degradation of type I collagen during subchondral bone resorption can be reflected by elevated levels of N- and C-terminal cross-linked telopeptides (NTX-I and CTX-I). Berry et al\(^2\) found that both CTX-I and NTX-I were significantly associated with reduced cartilage loss. Bettica et al\(^3\) found that these two bone resorption markers were significantly associated with progressive OA patients, but there was no difference between healthy and nonprogressive OA subjects. However, these conclusions are probably not unequivocal. A review suggested that biochemical markers of bone metabolism performed less adequately in comparison to cartilage and synovium factors.\(^9\)

Other less investigated OA biochemical markers include type I collagen synthesis (PICP, PINP), collagen type III synthesis (PIIINP), collagen type I and II degradation (C1, 2C), osteocalcin, bone sialoprotein, keratin sulphate, chondroitin sulphate 846 (CS846), human glycoprotein of cartilage 39 (YKL-40), collagen cross-links (pyridinoline [Pyr], deoxypyridinoline [D-Pyr], glucosyl-galactosyl pyridinoline [Glc-Gal-Pyr]), and pentosidine. Information on the molecular basis of the aforementioned biochemical markers is described in the comprehensive review by Garnero et al in 2000,\(^2\) and a detailed review on the investigation of these biochemical markers has been provided by Livshits et al.\(^2\)

The performance of a biomarker is characterized by sensitivity and specificity, or positive and negative predictive values. A biomarker for clinical use needs to have a good sensitivity of 0.9 or higher, and a good specificity of 0.9 or higher.\(^2\) However, most of these biochemical OA markers do not have data on their predictive capability yet. The data on the two most investigated markers – CTX-II and COMP – show that they do not have such a predictive capability for clinical use.\(^2\) Animal model data and results from a pilot human study using risedronate, a bisphophonate, showed that urine levels of CTX-II were dramatically decreased by risedronate treatment, together with expected decreases on bone turnover marker levels.\(^2\) These results were confirmed in a large randomized clinical trial of knee OA.\(^2\) However, neither knee joint structure as monitored by standardized radiographs nor symptoms were affected by risedronate treatment over 2 years.\(^2\) These results suggested that CTX-II failed as a biomarker to reflect efficacy of response in clinical trials. On the other hand, the majority of these biomarkers were investigated using radiographic OA as an end-point measure, making it difficult to judge their usefulness in early diagnosis of the disease. A recent study\(^2\) examined 14 biomarkers (uCTX-II, uCTX-I, sCOMP, sPIANP, sCS846, sC1, 2C, sOC, sPINP, sHA, sPIIINP, pLeptin, pAdiponectin, pResistin) in a 10-year prospective cohort of 1002 individuals with early symptomatic knee and/or hip OA. Using principal component analysis, the authors identified five clusters of interrelated biomarkers within the biomarker spectrum, consecutively designated as “bone-CTX-II,” “inflammation,” “synovium,” “C1, 2C-adipokines,” and “cartilage synthesis” clusters. The identified clusters extended knowledge on individual biomarkers, suggesting the potential of combined biomarkers in early diagnosis of OA.

**Novel biomarkers by a metabolomics approach**

There is mounting evidence that OA may also be a “metabolic disorder,” as lipid, metabolic, and humoral factors appear to contribute to the initiation and progression of OA.\(^2\)

Metabolomics is a state-of-the-art technique that allows a large number of small-molecule metabolites from body fluids or tissues to be detected quantitatively in a single step, and promises immense potential for early diagnosis, therapy monitoring, and understanding the origin and development of many diseases. The method has proved very useful in the rapid assessment of several disease states, such as diabetes,\(^2\) coronary heart disease,\(^2\) and blood pressure.\(^2\)

Application of the metabolomics approach in OA biomarker discovery is emerging. Williamson et al\(^2\) studied the levels of a range of components measured by \(^1\)H NMR in samples of synovial fluid taken from three groups of patients comprising ten with OA, 18 with rheumatoid arthritis (RA), and eleven with traumatic effusions. They found that patients with traumatic effusions had high levels of saturated triglycerides, while those with OA had low levels. The chain length of the triglycerides found in OA synovial fluid appears to be shorter than that for the other groups. Lamers et al\(^2\) studied urine samples of 47 non-OA controls and 45 individuals with radiographic OA of the knees or hips. They showed that urine NMR spectra can discriminate OA cases and controls in both males and females, and the metabolic profiles largely resembled the profile identified in the guinea pig model. They also demonstrated a high correlation between Kellgren–Lawrence score and the metabolite profile with \(R^2 = 0.82–0.93.\)
Recently, using the targeted metabolomic approach, we studied 163 serum metabolites in a discovery sample of 123 knee OA cases and 299 controls from the TwinsUK cohort and a replication sample of 76 knee OA cases and 100 controls from the Chingford Study. We identified 14 metabolite ratios that were significantly associated with knee OA in the TwinsUK cohort. Two of these 14 metabolite ratios were confirmed in the Chingford Study as correlating with knee OA: the ratio of branched-chain amino acids to histidine. The findings are supported by other animal and human data. An animal model of osteoarthritis showed an enhancement of the resonance at 0.85 ppm of the $^1$H high-resolution magic-angle spinning NMR spectra of the osteoarthritis-affected cartilage sample, which could be attributable to the increase in leucine and isoleucine. A recent study obtained metabolic profiles of synovial tissue cultures from patients with end-stage OA or from control individuals. The researchers identified 105 distinct compounds, and concentrations of eleven of these biochemicals were markedly different between the samples from patients with OA and those from controls. Although metabolic profiles from the cultured tissues may not be expected to reproduce in vitro profiles, the study confirms that potential biomarkers of OA could be involved in cellular metabolic and energetic processes such as branched-chain amino acid catabolism and support our findings.

These results are still preliminary in terms of clinical use, and further characterization is needed, with their predictive capability to be assessed. However, these studies highlight the potential of metabolic analyses to provide an alternative perspective of the altered biochemical processes responsible for OA onset and/or progression.

Genetic markers

Evidence suggests that genetic factors play a major role in OA, although they may be site- and sex-specific. From twin studies, this genetic influence has been estimated to be between 40% and 65%, and first-degree relatives have a two- to threefold increased risk of disease. The nature of the genetic influence in OA is still unclear, but is likely to involve a combination of effects on structure (ie, collagen), alterations in cartilage, or bone metabolism or inflammation. It is believed that identification of specific genetic factors for OA can help our understanding in the pathogenesis of OA and identify people and families with high risk for OA earlier.

Numerous efforts have been made at great expense on human genetic studies on OA worldwide. Several linkage scans have been performed and identified large chromosome regions associated with OA, but these are of limited value for detecting any specific susceptible genes. A number of candidates have been reported with OA, although many early studies were based on small numbers and were not replicated. One of the most consistent has been the vitamin D receptor gene (a candidate gene for osteoporosis), although size and direction of results differed when subjects were defined by osteophytes or joint space or site of OA. Another potential candidate related to bone is transforming growth factor–β, and an association with disk degeneration. A single-nucleotide polymorphism (SNP) (rs12901499) mapping to intron 1 of SMAD3, a key intracellular messenger in the transforming growth factor–β signaling pathway, was associated with both knee and hip OA. Lumbar disk degeneration has also been associated with mutations in the gene for collagen type IX in the Finnish population. Other candidate genes include estrogen receptor genes, aggrecan, CRTL1, TNF, and the interleukin (IL)-1 gene cluster, which has been validated in several populations of knee OA. Another is the gene FRZB, implicated in linkage studies and associated in females with hip OA. FRZB was also found to be associated with knee OA in women but not men. DIO2 is a regulator of thyroid hormone metabolism in the growth plate and may confer susceptibility for OA at multiple joint sites, as suggested in a linkage study. Valdes et al showed an association with specific candidates with progression of knee OA, which given that OA is nearly universal in the elderly, may be a more revealing phenotype. The genes implicated were ADAM12, CILP, OPG, and TN4. These genes have recently been replicated in an independent sample with clinical knee OA. A convincing association study has been performed in Japanese patients, where the gene that encodes for asporin – ASPN – an extracellular matrix protein, was associated in two populations with knee OA. A meta-analysis suggested that an ASPN allele is protective against the risk of knee OA in Caucasians. Spector et al reported that the gene leucine-rich repeats and calponin homology containing 1 (LRCH1) was consistently associated with knee OA in three samples of pooled DNA from two populations with northern European ancestry but this has not been replicated independently. Using 2170 patients with OA and 2849 controls, Valdes et al found three SNPs in the ANP32A gene were significantly associated in hip OA, but not knee OA. The most convincing and robust association was a single SNP (rs143383, T/C) located in the 5′-UTR of the GDF5 gene, which was reported in Japanese and Chinese case-control cohorts. The major T allele of the SNP was
common in Asian populations, with frequencies >70% in controls, and was at an elevated frequency in OA cases, with odds ratios ranging from 1.30 to 1.79 for knee and hip cases. The same T allele was found to be increased in hip and knee OA cases from Spain and the UK relative to controls with a very modest odds ratio of 1.10.72,73

However, all these previous candidate gene studies have to be interpreted with extreme caution, as most are likely to be false positives.74 Large meta-analyses of published candidates in diabetes75 and osteoporosis76 have suggested that less than 10% of reported published associations are real. However, a large meta-analysis from the TreatOA consortium did confirm the association of the gene GDF5 but not FRZB, despite many publications on the latter.72

Genome-wide association study (GWAS) is a powerful approach for unlocking the genetic basis of complex diseases such as OA. The method has uncovered >800 SNP associations for more than 150 disease and other traits.77 Notable advantages include its comprehensiveness and the potential for finding susceptibility genes with previously unknown loci and relationship to the diseases.

The first wave of GWAS in OA has been published. While these studies clearly show that there is no definitive and common highly penetrant allele that causes OA, some interesting candidate genes emerged from these studies.

To date, five large case-control association scans have been reported. Mototani and coworkers78 tested 72,000 markers for association with hip OA, and identified a variant in the CALM1 gene to be strongly associated in the Japanese population. However, studies in United Kingdom samples failed to show an association of this variant with hip79 or knee OA.63 A pooled, large-scale (500,000 markers) GWAS on knee OA has been published.80 The variants identified by this scan, although not achieving genome-wide significance (P < 5 × 10−8), have been subsequently replicated in independent cohorts, and fell in the 5′ region of the gene encoding the COX-2 and the cytosolic phospholipase enzymes (both involved in prostaglandin synthesis), in the 2q33 linkage region, and near a gene involved in transcriptional repression of thyroid hormone receptors. Using a two-stage approach genome-wide association, we identified the SNP rs716508 located in the A2BP1 gene was associated with hand radiographic OA. The same allele of the SNP was also associated with reduced bone density at both hip and lumbar spine, suggesting the potential mechanism of the gene in hand OA might be via effects on subchondral bone.81 Kerkhof et al82 performed a GWAS testing >500,000 SNPs in 1341 OA cases and 3496 Caucasian controls from The Netherlands. SNPs associated with at least two OA phenotypes were analysed in 14,938 OA cases and approximately 39,000 controls. The minor allele of rs3815148 on chromosome 7q22 was associated with a 1.14-fold increased risk for knee and/or hand OA (P = 8 × 10−6), and also with increased risk for knee OA progression. The region encompasses six genes: PRKAR2B, HPB1, COG5, GPR22, DUSAL, and BCAP29. None of these are obvious candidates for OA susceptibility, and functional studies have so far failed to prioritize compellingly one over the others.83 arcOGEN84 is a UK study that recently reported stage 1 of its GWAS, which involved a discovery sample of 3177 OA cases and 4894 controls. The strongest signals found were rs2277831 in MICAL3, rs11280 in Cboorf130, and rs2615977 in COL11A1, but none of these hits reached a genome-wide significance.

However, to date, no single large genetic effect has been found. Rather, the increased risks for carrying a predisposing gene variant are likely to be modest (odds ratio less than 1.3),85,86 thus limiting their individual predictive capability for clinical use. However, combinations of large numbers of risk alleles may have diagnostic value. Valdes et al87 examined 36 SNPs in 17 candidate genes previously associated with OA in 603 knee OA cases who met American College of Rheumatology criteria and 596 age- and ethnicity-matched controls. The odds ratio for individuals in the top quartile of the genetic risk variable compared to those in the bottom quartile was found to be 8.68 (95% confidence interval 5.20–14.49, P < 2 × 10−6) for women and 5.06 (95% confidence interval 3.10–8.27, P < 1 × 10−6) for men, suggesting that the additive information from a number of genetic variants can predict a substantial proportion of the risk of knee OA.

Epigenetic markers

Epigenetics encompasses changes to marks on the genome that are copied from one cell generation to the next, which may alter gene expression but do not involve changes in the primary DNA sequence. These marks include DNA methylation, histone modifications, and noncoding RNAs. Epigenetic patterns undergo dynamic changes during development, cell differentiation, and in response to environmental stimuli, leading to changes in temporal and spatial gene expression. Alteration in epigenetic state has been correlated with cancers,88–93 and several other complex diseases.94–97 A recent genome-wide DNA methylation study98 on bladder cancer not only found that methylation of TBX2, TBX3, GATA2, and ZIC4 was associated with progression to muscle-invasive disease in pTa tumors but also demonstrated that methylation
of TBX2 alone has a sensitivity of 100%, a specificity of 80%, a positive predictive value of 78%, and a negative predictive value of 100%, with an area under the curve of 0.96 for predicting progression, suggesting DNA methylation markers have great potential to serve as useful biomarkers for disease prognosis.

Data on epigenetics in OA are still sparse. Poschl et al investigated whether the loss of aggrecan expression in OA was linked to methylation changes in the promoter by studying eleven normal (age 60–90 years) and six OA cases (age 62–79 years), but were unable to find an association. Roach et al studied cartilage tissues obtained from the femoral heads of 16 patients with OA and ten control patients with femoral neck fracture to investigate whether the abnormal expression of matrix metalloproteinases MMP-3, -9, and -13 and ADAMTS-4 is associated with epigenetic unsilencing. They found that the overall percentage of nonmethylated sites was increased in OA patients (48.6%) compared with controls (20.1%): 20% versus 4% for MMP-13, 81% versus 47% for MMP-9, 57% versus 30% for MMP-3, and 48% versus 0% for ADAMTS-4. Cheung et al further demonstrated that the abnormal expression of ADAMTS-4 in OA chondrocytes corresponds to a heritable loss of DNA methylation at some CpG sites in the proximal promoter region of ADAMTS-4. Nieu et al and Iliopoulos et al found that leptin expression level was different between OA and controls and demethylation of leptin promoter region might upregulate leptin gene expression level and contribute to OA. Hashimoto et al found that demethylation at the specific CpG sites in the IL-1β promoter in response to inflammatory cytokines in human articular chondrocytes results in long-term induction of IL-1β. Scott et al showed that a reduction of superoxide dismutase 2 is associated with early stage of OA, and the superoxide dismutase 2 promoter had significant DNA methylation alteration in OA cartilage. Expression of the gene GDF5, which is consistently reported to be associated with OA across populations, is found to be modulated epigenetically by DNA methylation.

Further studies are needed to confirm these results and examine their potential value as diagnostic and prognostic biomarkers in OA. Also, large-scale studies should be initiated to examine the DNA methylation patterns in OA, systematically using recent advanced genomic technologies.

Small noncoding RNAs, known as miRNAs – short (∼21 nucleotides) single-stranded RNA molecules – play an important role in post-transcriptional regulation of gene expression. There are approximately 1000 miRNAs in the human genome, some of which are tissue-specific. Since extracellular miRNAs are detectable in most body fluids and excretions and they are resistant against factors like enzymes and freezing, they might be considered promising diagnostic markers for diseases.

Several miRNAs are found to be differentially expressed between normal and osteoarthritic cartilage. Increased expression levels of miR-9 and miR-98 and decreased levels of miR-146, miR-27b, and miR-140 have been reported in OA cartilage. miR-140 was found to be expressed only in cartilaginous tissues in developing zebrafish and in murine skeletal development. In humans, the expression of miR-140 increases during chondrogenesis and is more abundant in articular cartilage, but reduced in OA. However, most of these studies are on animal models or in vitro, and population-based data are still sparse. Okuhara et al examined the expression patterns of miRNAs in the peripheral blood mononuclear cells of OA patients, and found that the relative expression levels of miR-146a, 155, 181a, and 223 in the OA patients were significantly higher than those found in healthy controls, suggesting these miRNAs might be related to the pathogenesis of OA. Murata et al found that synovial fluid concentrations of miR-16, 132, 146a, and 223 were significantly lower than their plasma concentrations in RA and OA patients, and there was no correlation between plasma and synovial fluid miRNAs. Plasma miR-132 differentiated healthy controls well from patients with RA and OA, while synovial fluid miRNAs differentiated RA and OA. Plasma miR-32 has a sensitivity of 84% and a specificity of 81.2% for OA diagnosis. These results are promising, and indicate the great potential value of miRNAs as biomarkers for OA.

Conclusion
A significant body of work on biomarkers of OA exists, and there is no doubt that it allows better understanding of the OA disease process. However, to date, none of these proposed biomarkers could be used in daily practice for diagnosing, monitoring, prognosticating, and clinical trials for OA, partly because of the lack of information about sensitivity, specificity, normal range, and clinically important differences. More research is needed to further characterize the previously identified biomarkers. Application of the latest state-of-the-art genomic and metabolomic technologies in novel OA biomarker discovery will help us not only in better understanding the pathophysiology of OA but also in the generation of a new wave of biomarkers.
Disclosure
Financial support was received from the Research and Development Corporation of Newfoundland and Labrador Ignite Program, and the Cox Award, Faculty of Medicine, Memorial University of Newfoundland.

References


