Extracellular Genomic Biomarkers of Osteoarthritis

Emma Budd a, b, Giovanna Nalesso b, Ali Mobasheri a, b, c

a The D-BOARD European Consortium for Biomarker Discovery, School of Veterinary Medicine, Guildford, University of Surrey, GU2 7XH, United Kingdom

b Department of Veterinary Pre-Clinical Sciences, School of Veterinary Medicine, Faculty of Health and Medical Sciences, University of Surrey, Guildford, GU2 7XH, United Kingdom

c Arthritis Research UK Centre for Sport, Exercise and Osteoarthritis, Queen’s Medical Centre, Nottingham, NG7 2UH, United Kingdom

* Address correspondence and reprint requests to: Dr Emma Budd, School of Veterinary Medicine, University of Surrey, Guildford, GU2 7AL, United Kingdom

E-mail addresses:

e.budd@surrey.ac.uk (Emma Budd)

g.nalesso@surrey.ac.uk (Giovanna Nalesso)

a.mobasheri@surrey.ac.uk (Ali Mobasheri)

Authors’ contributions: E.B. conceived, drafted and edited the manuscript. G.N. edited and approved the submission. A.M. edited and approved the submission. All authors contributed to data interpretation and manuscript preparation and approved the final version submitted.

Acknowledgements: The research leading to these results was funded by the European Commission Framework 7 programme (EU FP7; HEALTH.2012.2.4.5-2, project number 305815; Novel Diagnostics and Biomarkers for Early Identification of Chronic Inflammatory Joint Diseases). This project has also received partial support from the Innovative Medicines Initiative Joint Undertaking under grant agreement No. 115770, resources of which are composed of financial contribution from the European Union’s Seventh Framework programme (FP7/2007-2013) and EFPIA companies’ in-kind contribution. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.
Disclosure statement: The authors declare that they have no competing interests. This paper was written by the authors within the scope of their academic and research positions. None of the authors have any relationships that could be construed as biased or inappropriate. A. Mobasher is coordinator of the D-BOARD Consortium funded by the European Commission Framework 7 programme (EU FP7; HEALTH.2012.2.4.5-2, project number 305815; Novel Diagnostics and Biomarkers for Early Identification of Chronic Inflammatory Joint Diseases) and member of the Arthritis Research UK Centre for Sport, Exercise, and Osteoarthritis, funded by Arthritis Research UK (grant reference numbers: 20194 and 21595). A.M. is a member of the Applied Public-Private Research enabling OsteoArthritis Clinical Headway (APPROACH) consortium, a 5-year project funded by the European Commission’s Innovative Medicines Initiative. APPROACH is a public-private partnership directed towards osteoarthritis biomarker development through the establishment of a heavily phenotyped and comprehensively analysed longitudinal cohort.
Abstract

Introduction
Osteoarthritis (OA), a chronic, debilitating and degenerative disease of the joints, is the most common form of arthritis. The seriousness of this prevalent and chronic disease is often overlooked. Disease modifying OA drug (DMOAD) development is hindered by the lack of soluble biomarkers to detect OA early. The objective of OA biomarker research is to identify early OA prior to the appearance of radiographic signs and the development of pain.

Areas covered
This review has focused on extracellular genomic material that could serve as biomarkers of OA. Recent studies have examined the expression of extracellular genomic material such as miRNA, IncRNA, snoRNA, mRNA and cell-free DNA (cfDNA), which are aberrantly expressed in the body fluids of OA patients. Changes in genomic content of peripheral blood mononuclear cells (PBMCs) in OA could also function as biomarkers of OA.

Expert commentary
There is an unmet need for soluble biomarkers for detecting and then monitoring OA disease progression. Extracellular genomic material research may also reveal more about the underlying pathophysiology of OA. Minimally-invasive liquid biopsies such as synovial fluid and blood sampling of genomic material may be more sensitive over radiography in the detection, diagnosis and monitoring of OA in the future.

Key words
Osteoarthritis; Extracellular; Genomic biomarkers; Body fluids; Liquid biopsy; ncRNA; miRNA; cfDNA; mRNA; extracellular nucleic acid carriers
1. The current perspective on osteoarthritis

Osteoarthritis (OA) is a chronic and debilitating disease of the joints which is more prevalent than rheumatoid arthritis (RA) and any other arthritic disease [1]. Osteoarthritis (OA) is characterised by progressive deterioration of hyaline cartilage concomitant with changes in surrounding tissues including ligaments, synovium and subchondral bone, resulting in defective integrity of affected articular joints or the spine [2]. OA is classified as a low-grade yet chronic inflammatory disease, with inflammation having a role in many of the associated pathologic changes [3]. Articular cartilage is alymphatic, aneural and avascular, with non-proliferative and non-migratory resident chondrocytes, factors which are likely to contribute the development of OA in part due to the limited capacity that articular cartilage has for endogenous regeneration [4]. OA is a heterogeneous disease which can be driven by different underlying factors. Phenotypic subgroups include a traumatic injury driven phenotype, an ageing driven phenotype, a synovitis-inflammatory driven phenotype, a subchondral bone phenotype, a cartilage driven phenotype and a metabolic phenotype [1]. The most common risk factors for OA include prior joint injury, age, sex, genetic predisposition and mechanical factors such as malalignment [5].

Not only is OA the most common form of arthritis [6], it is the most common disease of the joints [7]. It has been estimated that approximately 33% of the UK’s population aged over 45 have sought treatment for OA [8]. OA is part of a wider group of diseases which cumulatively are known as musculoskeletal conditions. In 2015 in the UK, musculoskeletal disorders remained to be the leading cause of years lived with disability and an estimated 30.8 million working days were lost in 2016 due to a musculoskeletal condition [8]. In 2013/14 musculoskeletal conditions were accountable for the 3rd largest area of national health care (NHS) programme spending, costing £4.7 billion [8]. With an ageing population and increasing life expectancy the burden of OA will only increase, which poses an unmet need for the development of new therapies, preventative measures and the identification of sensitive biomarkers for OA diagnosis.

Despite OA being the most common form of arthritis, OA still remains to gain recognition as a serious disease. It has been reported that 78% of people living with arthritis agreed that society doesn’t understand their condition because it doesn’t appear as if their condition is serious [8]. In December 2016 the Pre-competitive Consortium for Osteoarthritis (PCCOA) of the Osteoarthritis Research Society International (OARSI) submitted a white paper to the Food and Drug Administration (FDA) in order to attract the attention of the FDA and policy makers to understand the need for effective therapies for OA [9]. The submission of the white paper was as a result of OA experts to portray the seriousness of this chronic and prevalent disease, highlighting the economic burden of OA upon society and the detrimental effects of OA on quality of life, which may deny OA sufferers the right to healthy ageing and basic human rights which are outlined in the United Nations Charter for Rights of Persons with Disability [9]. The white paper was submitted to encourage the development of global and national...
health policies to address the increasing burden of OA and encourage drug development towards identifying disease modifying osteoarthritis drugs (DMOADs). Drug and biomarker development are interdependent, therefore DMOAD development is dependent on sensitive biomarkers which can detect and predict disease progression [10].

2. Limitations in diagnosing osteoarthritis

In an era where the use of biomarker analysis to confirm or assess disease state is rapidly advancing, OA lacks the use of any soluble biomarker for diagnosis in the clinical setting. To identify individuals with early-onset OA a screening method to detect biomarkers from body fluid could help in the early diagnosis, and in the future, utilised in companion with a therapy in the treatment of the disease. Radiographs and joint imaging techniques, which are currently routinely used for a diagnosis of OA, do not measure dynamic changes in the joint and are usually implemented when OA is symptomatic at which point OA has progressed to advanced stages of the disease [11]. Radiography is the most common imaging modality to assess OA, with the use of radiographic grading or scoring classification systems to classify stages of OA [12]. One of the most commonly used radiographic definitions of OA is the Kellgren and Lawrence (KL) classification which grades the whole joint with a score from 0 to 4, with grade 0 as no OA and grade 4 as severe OA. The KL classification is based on radiographic observation of joint space narrowing, marginal osteophytes, subchondral bone sclerosis and bone shape alteration [12]. While imaging techniques such as radiography can detect and monitor structural changes to the joint, giving an indication of disease stage, the rate of OA development and disease progression cannot be evaluated. Imaging techniques can classify OA by structural change only, a KL score of 0 may indicate no OA, but molecular changes at the cellular level may have occurred, for which there is no current method of detection. Joint imaging techniques can only assess structural change at which point OA development is irreversible. A requirement to define and detect the changes which occur during OA disease initiation is needed for which there is a requirement for a sensitive biomarker that can detect and diagnose early-onset OA and monitor disease progression.

A compounding limitation in OA diagnosis is the range in symptomology which accompanies the disease in different individuals. Asymptomatic OA is a common phenomenon in which disease progression continues with a lack of symptoms in some patients [13]. Knee pain, for example, can be a symptom of OA, but is an imprecise indicator for OA which may not be necessarily correlated with radiographic findings, likewise radiographic knee OA is an imprecise indicator that knee pain will be present. OA has been found to be absent in individuals following radiography despite knee pain suggestive of OA and OA has been found to be present through compounding radiographic evidence in individuals that are asymptomatic [14]. There is no defining risk profile which can predict if an individual has significant damage to cartilage, if an individual has asymptomatic OA, if an individual
will develop OA or what the progression of destruction to articular cartilage and the joint will be or the rate of OA progression [15].

OA is often not diagnosed until late stages, after significant, destructive and irreparable damage to the joint has already occurred. It has been suggested that molecular indicators exist long before radiographic evidence indicates the presence of OA and it has been estimated that the OA process begins 20 years before OA is detected by radiographic joint space narrowing (JSN) [16]. The issue lies with the diagnosis of OA, clinically detectable OA is currently confirmed by magnetic resonance imaging (MRI), ultrasound, nuclear medicine bone scans and/or radiography, but currently there is no method to detect biomarkers from extracellular body fluids. Early-stage OA has been defined and involves structural changes to the cartilage [17]. Ryd et al have attempted to define a stage in the OA disease process prior to early-stage OA known as pre-OA. Pre-OA is defined as the preclinical stage in OA whereby OA has been established for a length of time at the cellular level only, prior to early-stage OA which is distinguished by changes in cartilage structure [18]. Ryd et al suggest that the definition of pre-OA should be: ‘A knee exhibiting one or many risk factors without pain, normal standing radiographs, no structural changes on arthroscopy or standard MRI, this is, before early OA can be diagnosed’ [18]. Preventative measures should be applied at the pre-OA stage to forestall the onset of OA, but first, this stage needs to be identified in individuals.

When discussing the development of biomarkers, it is important to highlight that the identified biomarker/s will be anticipated to be utilised in screening for OA, therefore the method of how potential biomarkers are isolated from the body must be minimally invasive. Other than in arthroscopic debridement, articular cartilage biopsy is not advocated in the clinic to investigate disease status because articular cartilage does not regenerate. The evaluation of extracellular body fluids including synovial fluid, blood, plasma, serum and urine are of interest in identifying a soluble OA biomarker due their relative ease in being isolated from the body [19].

In recent years much of OA biomarker research has focused on proteomic techniques [2]. The view exists that proteomic approaches are advantageous over nucleic acid expression profiling because there is not the potential disconnection that can occur between gene expression and protein expression levels [20] as a result of post-translational modification. The secretome is comprised of secreted proteins and reflects the health of cells in real time [21]. Recent attention has turned to utilising proteomic techniques to identify proteins in the secretome of articular cartilage models of OA, as a potential indicator of disease biomarkers which may be present in the synovial fluid or blood in OA [22]. Proteomic profiling has focused on body fluids from OA patients in the search for a biomarker which is inherent to OA [23-32]. The perception exists that extracellular genomic material is unstable as a result of harsh extracellular environments which may have contributed to a lack of focus on the extracellular environment for OA genomic biomarker discovery.
The International Conference on Harmonisation (ICH) Topic 15 defined a genomic biomarker as ‘A measurable DNA and/or RNA characteristic that is an indicator of normal biologic processes, pathogenic processes, and/or response to therapeutic or other interventions’ [33]. A genomic biomarker includes measurements of gene expression, gene function and regulation of a gene and can include DNA/RNA characteristics such as SNPs, haplotypes, variability of short sequence repeats, DNA modifications, copy number variations, deletions or insertions of (a) single nucleotide(s), cytogenetic rearrangements, RNA sequences, RNA processing, RNA expression levels and miRNA levels [33]. Pre-OA which is asymptomatic, established at the cellular level and not accompanied by structural changes and therefore undetectable by radiography, requires a method of detection before structural changes ensue [34]. Extracellular genomic material originates from the intracellular environment and may be more informative than extracellular matrix proteins with regards to the intracellular changes which occur prior to the onset of structural changes in OA. Not only could genomics identify genomic biomarkers of OA but also reveal more about the underlying cellular pathophysiology.

3. Emerging extracellular genomic biomarkers of osteoarthritis

An extensive array of RNAs have been suggested to exist within extracellular spaces and analysis of sequencing data from plasma-derived RNA identified 1,192 human extracellular RNA including miRNAs, piRNAs and snRNAs [35]. Extracellular vesicles (EVs) have been found to contain DNA, mRNA and subsets of ncRNA [36]. Ubiquitous ribonucleases present in body fluids degrade extracellular RNA [37], but exosomal-derived RNA has been shown to remain stable following RNase A treatment [38]. Extracellular miRNA and potentially other classes of RNA may remain stable in extracellular nuclease-rich environments as a result of extracellular carriers such as exosomes, microvesicles, apoptotic bodies, lipoproteins or protein complexes [39]. In addition to extracellular carrier-associated nucleic acids, cfDNA not associated with any carrier exists in the extracellular environment [40]. Table 1 lists and Figure 1 illustrates both extracellular carrier-associated and extracellular carrier-free genomic material which have the potential to serve as genomic biomarkers in OA.

3.1 Functional non-coding RNA (ncRNA) in osteoarthritis

The genome contains protein-coding DNA sequences that are transcribed to produce messenger RNA (mRNA) which in turn are translated to polypeptide sequences, and non-coding DNA, sequences some of which are transcribed to functional RNA sequences but are not translated into polypeptides (Figure 2). 1-2% of the human genome encodes proteins [41], with the other 98-99% of the human genome comprising non-coding DNA. Important to note is the distinction between functional non-coding DNA, which includes DNA sequences for example cis-regulatory elements, scaffold/matrix (S/MAR) attachment regions, introns, telomeres and DNA which encodes genes for non-coding RNA, and, non-functional non-coding DNA, which has no known function [42]. Regardless of the exact proportion of
functional non-coding genomic DNA, there exists a proportion which is transcribed into functional RNA (ncRNA) which includes small nuclear RNA (snRNA), transfer RNA (tRNA), ribosomal RNA (rRNA), small nucleolar RNA (snoRNA), microRNA (miRNA), piwi-interacting RNA (piRNA), circular RNA (circRNA) and long non-coding RNA (lncRNA) [43]. Definitions of functional ncRNA are described in Table 2.

NcRNAs can be classified into two groups. NcRNAs that are constitutively expressed and have infrastructural or housekeeping functions, involved in translation, RNA splicing and DNA replication, include tRNA, rRNA and snRNA. While ncRNAs such as miRNA, piRNA and lncRNA which may have roles in gene expression regulation, classified as regulatory ncRNAs, can themselves be differentially expressed [44]. Diseases have been attributed to perturbation in protein coding gene expression, focus has now turned to looking at the role that ncRNAs may have in disease. The expression of ncRNAs, like mRNA and protein, can be quantified which offers the potential to measure subsets of ncRNAs as indicators of disease state and therefore ncRNAs may serve as genomic biomarkers. Dysregulation of ncRNAs have been observed in tumorigenesis and in cardiovascular, autoimmune, monogenic, neurological, developmental and imprinting disorders [45]. Subsets of ncRNAs have been detected and remain stable in body fluids [46], thus circulating ncRNAs may qualify as potential biomarkers for OA. Discussed below are some of the ncRNAs that have been identified in extracellular fluids in OA.

3.1.1 MicroRNA

MicroRNA (miRNA) are single stranded ncRNA of approximately 22 nucleotides in length. MiRNA function by binding to locations within mRNA, typically within the 3’UTR, following RNA transcription [47]. The binding of miRNA to mRNA, in the cytoplasm, prevents mRNA translation in the synthesis of protein. MiRNA is therefore capable of down-regulating protein expression through effective mRNA targeting [48]. MiRNA research has focused on identification of intracellular miRNA, given that miRNA exert their function of down-regulating protein expression in the cell cytoplasm. Numerous studies have shown the differential expression of different miRNAs from OA cartilage; the expression of miR-140 [49], miR-125b [50], miR-210 [51], miR-222 [52], miR-199a* [53], miR-588 [54], miR-411 [55], miR-105 [56], miR-502-5p [57] have been shown to be down-regulated in human OA chondrocytes and/or cartilage and the expression of miR-146b [4], miR-181b [58], miR-30a [59], miR-455 [60], miR-30b [61], miR-16-5p [62], miR-146a [4,63] have been shown to be up-regulated in human OA chondrocytes and/or cartilage. However, differential miRNA expression in cartilage-centric OA studies does not prove useful in identifying a non-invasive extracellular miRNA biomarker for diagnostic purposes.

Weber at al identified miRNA in 12 body fluids, including breast milk, amniotic fluid, saliva, plasma, tears, urine cerebrospinal fluid, colostrum, seminal fluid, peritoneal fluid and pleural fluid [64]. Release
of miRNAs into cell culture media has been correlated with cell death and damage in specific tissues has been shown to result in an increase in the level of tissue specific miRNAs in circulation [65]. But, miRNA may be more than by-products of cell apoptosis or necrosis [66] and may also have a potential role in extracellular cell-cell communication [65]. Expression of stable extracellular miRNA could prove useful as a biomarker in OA.

**Synovial Fluid**

The concentration of miR-16, miR-146a, miR-155 and miR-223 were found to be significantly higher in synovial fluid from RA patients compared to the synovial fluid from knee OA patients [67]. An assessment of miR-16, miR-146a, miR-155 and miR-223 concentration could be diagnostic in discriminating knee OA from RA. MicroRNA profiling of non-OA and knee OA synovial fluid-derived EVs from males and females identified miR-504-3p as the only miRNA to be expressed in both genders, which indicates the potential need to examine miRNAs as biomarkers of OA by gender segregation [68]. Validation with qPCR of randomly selected miRNAs observed a significant down-regulation of miR-26a-5p, miR-146a-5p and miR-6821-3p in female knee OA synovial fluid-derived EVs, consistent with the miRNA microarray expression results [68], indicating the use of these microRNAs as diagnostic biomarkers in female knee OA.

MicroRNA profiling and follow-up qPCR validation of synovial fluid from patients with early stage radiographic knee OA compared with the synovial fluid of patients with late stage radiographic knee OA revealed significant elevation of miR-23a-3p, miR-24-3p, miR-27a-3p, miR-27b-3p, miR29c-3p, miR-34a-5p, miR-186-5p and miR-378a-5p in late-stage OA [69]. An assessment of the expression of these miRNAs may therefore give an indication of OA progression. Treatment of synovial explants derived from patients with knee OA, with IL-1β, resulted in the significant up-regulation of miR-23a-3p and miR-27b-3p, indicating that miR-23a-3p and miR-27b-3p could serve as biomarkers of inflammatory induced knee OA [69]. Similarly, miR-140 expression analysis was conducted on synovial fluid from patients that had radiographic mild, moderate and severe OA and compared to the synovial fluid of patients with no OA. Expression of miR-140 was found to be down-regulated with increasing radiographic severity of OA [70]. Expression of synovial fluid-derived extracellular miR-140 could serve as a biomarker of OA progression.

**Components of blood**

The serum from individuals with OA-related knee and/or hip arthroplasty was compared to the serum from individuals without arthroplasty for miRNA expression. Let-7e was found to be significantly down-regulated in patients with OA-related knee and/or hip arthroplasty and was also found to be continuously decreased with increasing number of hip and/or joint replacement surgeries [71]. Expression of serum-derived Let-7e could serve as a biomarker for severe hip or knee OA.
Human TaqMan miRNA array and follow-up qPCR validation of total RNA from the plasma of patients with primary knee OA and patients without OA identified significant up-regulation of 10 miRNAs in OA plasma, including miR-16, miR-146a, miR-29c, miR-93, miR-126, miR-184, miR-186, miR-195, miR-345 and miR-885-5p [72], indicating their use as biomarkers of knee OA. Total RNA isolated from the plasma of patients with knee OA was compared to healthy controls. The expression of miR-19b-3p, miR-122-5p, miR-486-5p and miR-92a-3p were found to be significantly up-regulated in the plasma of OA patients. Univariate and multivariate logistic analysis identified miR-19-3p, miR-122-5p and miR-486-5p to be independent factors for risk of knee OA. ROC curve analysis revealed that a combination of miR-19-3p, miR122-5p and miR-486-5p resulted in a high diagnostic value for knee OA, indicating that the combined measurement of all 3 extracellular miRNAs from plasma could serve as soluble genomic biomarkers for knee OA [73].

Another source of miRNA is from cells which reside within the peripheral blood. The expression of miR-146a, miR-155, miR-181a and miR-223 were found to be significantly up-regulated in peripheral blood mononuclear cells (PBMCs) from both, patients with knee OA and in patients with RA, compared to PBMCs from healthy controls. miR-146a, miR-155 and miR-223 were also observed to be significantly up-regulated in PBMCs from patients with RA compared to PBMCs from patients with knee OA [74]. An assessment of miR-146a, miR-155 and miR-223 expression from PBMCs could be diagnostic in discriminating knee OA from RA. Additional examination found that the expression of miR-155 was found to be significantly up-regulated in late-stage radiographic OA when compared to earlier radiographic stages of OA, whereas, expression of miR-146a and miR-223 were found to be significantly up-regulated in early-stage radiographic OA compared to late-stage radiographic OA [74]. Combined assessment of miR-146a, miR-155 and miR-223 expression levels from PBMCs could serve as diagnostic indicators of disease progression.

3.1.2 Long non-coding RNA (lncRNA)

Due to their diverse structures and functions the definition of IncRNAs is limited to ncRNAs of more than 200 nucleotides in length which lack protein coding capability [75]. Attempts have been made to classify IncRNAs by means of their relationship to coding genes, genomic origin and level of processing [76]. For example, IncRNAs have been classified according to molecular mechanism and can be subtyped as having signalling, decoy, guide, scaffold and enhancer roles [77]. LncRNAs can also be subtyped according to their position and direction of transcription relative to other genes and can be classified into subtypes including intergenic, intronic, antisense, overlapping, processed and bi-directional IncRNAs [78].

Zhang et al observed the expression of the IncRNA HOTAIR to be significantly up-regulated in the synovial fluid of the temporomandibular joint (TMJ) of OA patients and also in the synovial fluid of mechanical-stress-induced OA TMJ in rabbits. The authors demonstrated that treatment of rabbit TMJ
condylar chondrocytes with IL-1β induced the up-regulation of HOTAIR and knockdown of HOTAIR resulted in down-regulation of both mRNA and protein levels of Mmp1, Mmp3 and Mmp9 and a decrease in IL-1β induced apoptosis, indicating that the lncRNA HOTAIR may play a role in chondrocyte apoptosis in OA [79]. Synovial fluid HOTAIR from TMJs could potentially function as a biomarker of OA of the TMJ joint.

The knowledge surrounding extracellular lncRNA expression in OA is limited, but studies of lncRNA expression in cells of the OA joint have been documented. Expression of lncRNAs in OA cartilage tissue and/or chondrocytes and synoviocytes, such as PCGEM1, GAS5, HOTTIP, LncRNA-CIR, H19, CILinc01 and CILinc02 have been observed, with roles in apoptosis, ECM degradation and inflammation [80]. Various studies have demonstrated expression of circulating lncRNAs in body fluids of cancer patients including in whole blood, urine, plasma, saliva and gastric juice [46] and circulating lncRNAs have been identified in endometriosis [81], acute myocardial infarction [82] and rheumatoid arthritis [83]. The field of lncRNA research is relatively young but holds promise for novel potential prognostic biomarkers for diseases including OA. It is likely that other extracellular lncRNAs involved in OA are yet to be identified.

3.1.3 Small Nucleolar RNA (snoRNA)

SnoRNAs, a sub-class of snRNA located in the nucleus, are involved with post-transcriptional modification by means of methylation or pseudouridylation of ncRNA including snRNA, rRNA and tRNA. SnoRNAs are divided into two classes, SNORDs and SNORAs, dependent on their structural features and sequences [84]. SnoRNAs are involved in rRNA processing, which is fundamental for ribosome biogenesis and ultimately translation and protein synthesis. Dysregulation of snoRNA is likely to have an impact on protein synthesis. Dysregulation of snoRNAs has been observed in certain cancers, neurodegenerative disease and viral induced diseases and changes in snoRNA levels have been observed in cells subjected to stress and drugs [85]. Differential expression of specific snoRNAs has also been observed in patients with cartilage damage associated with post-anterior cruciate ligament (ACL) damage [86] and in OA [87].

Zhang et al investigated the presence of ncRNA in the serum of patients with cartilage damage one year after patients had undergone arthroscopic ACL reconstruction surgery following ACL injury. The study looked to identify potential ncRNAs involved in early degenerative changes which may be associated with post-ACL injury-induced OA. At one year post-ACL injury, the concentration of U48 and U38 were significantly elevated in the serum of patients developing cartilage damage. The concentration of U38 was also found to be significantly elevated in cartilage with greater damage, assessed by whole organ magnetic resonance knee score, compared to cartilage with minor damage. Not only could both U48 and U38 from serum serve as biomarkers of cartilage damage which may lead to OA, but assessment of U38 concentration could be utilised to discriminate between stages of cartilage damage
If U38 concentration is correlated with increasing cartilage damage it may be a good indicator for early-onset injury-induced OA, further examination of U38 in early radiographic OA would be insightful.

In an in vivo experimental induced OA murine model, the expression of SNORD116, SNORA64 and U3 were found to be significantly up-regulated and the expression of SNORD46 was found to be significantly down-regulated in mice serum. The up-regulated expression of SNORD116 was also confirmed in another species, with a significant up-regulation in SNORD116 observed in OA equine serum [87]. SNORA64, U3, SNORD46 and SNORD116 could serve as biomarkers of OA and future examination of the expression of these snoRNAs in human OA serum is required.

3.2 Cell-Free, Carrier-Free Nucleic Acids

Cell-free nucleic acids encompass DNA and RNA which are not associated with cells, freely circulating in the blood and other body fluids [88]. Apoptosis or necrosis has been speculated to be the main source of cfDNA and RNA in plasma and serum [89,90]. CfDNA is not carrier-associated, while cell-free RNA is likely to be associated with a carrier [40]. Extracellular cfDNA is double stranded and short in length due to fragmentation, with most DNA molecules measuring approximately 170 bp in length, reminiscent of the smallest nucleosomal DNA fragment resultant of apoptosis [91]. Elevated cfDNA has been observed in various human diseases including myocardial infarction [92], acute pancreatitis [93], rheumatoid arthritis, systemic lupus erythematosus [94] and various types of cancer [95].

Leon et al examined the levels of extracellular DNA from patients with RA, OA, gout, pseudogout and post-traumatic arthritis (TRA) and identified that the concentration of DNA from both synovial fluid and serum was significantly higher in patients with RA compared to OA [96]. Hashimoto et al observed a significant increase in the concentration of cfDNA in the synovial fluid of RA patients compared to the concentration in OA patients [97]. Extracellular cfDNA may therefore serve as a biomarker to distinguish OA from RA. While DNA levels have been compared in RA and OA, no study has examined the levels of cfDNA in progressive radiographic OA. In late-stage OA chondrocyte apoptosis is evidenced by hypocellularity and lacunae emptying [98]. Given the association of increased cfDNA with elevated apoptosis or necrosis [99], an elevated level of cfDNA may be detected in late-stage radiographic OA compared to an earlier stage of radiographic OA.

3.3 Peripheral blood mononuclear cells (PBMCs) as a source of genomic biomarkers

Circulating blood contains PBMCs which encompass lymphocytes, natural killer cells, dendritic cells, monocytes [100] and a small percentage of haematopoietic stem and progenitor cells [101]. PBMCs are resident in an easily attainable fraction of blood, comprised of cells which actively undergo transcription and translation. The circulating mononuclear cells undergo immunophenotypic changes in response to disease and injury and therefore provide a peripheral source [102] for assessing gene and ncRNA expression relevant to different diseases. Chronic low-grade inflammation is a major driver of continual
degeneration of joint tissues in OA [103]. Infiltration of subsets of PBMCs including T and B lymphocytes, natural killer cells and macrophages into the synovium [103,104] and alterations in blood immune cell composition has been observed in OA [105]. Gene expression profiling of PBMCs offers a minimally invasive technique for potentially identifying novel biomarkers for OA.

Several studies have identified specific gene expression profiles of PBMCs in OA, which could serve as novel genomic biomarkers of the disease. As discussed previously, the significant up-regulation of miR-146a, miR-155, miR-181a and miR-223 was observed in PBMCs isolated from patients with knee OA [74]. RNA expression profiling of RNA isolated from the blood of patients with arthroscopically diagnosed mild knee OA revealed 3,543 genes to be differentially expressed. Significant down-regulation of HSPCA, IKBKAP, IL13RA1, LAMC1, TNFAIP6 and PF4 mRNA following qPCR validation was observed in patients with mild knee OA compared to controls [106]. Marshall et al suggest that utilising a combination of specific mRNA expression analysis of blood-derived RNA provides a more sensitive and non-invasive method to diagnose mild knee OA [106]. Ramos et al observed the differential expression of 679 protein coding genes and 15 ncRNAs in PBMCs from a cohort of patients with symptomatic OA in multiple joints in the hand or at least two joints from hand, spine, knee or hip. Differential expression of 22 genes identified out of the 694 were observed to be differentially expressed following qPCR validation and also in an independent cohort of OA patients [107].

Attur et al identified 173 differentially expressed genes in PBMCs from patients with OA and also showed that symptomatic knee OA patients could be categorised into two subclasses dependent on the expression of cytokines in PBMCs. Individuals that exhibited up-regulated expression of cytokines in PBMCs, sub-classed as cytokine overexpressors, were found to have higher pain scores and a higher risk for symptomatic radiographic knee OA progression. Cytokine gene expression profiling of PBMCs may therefore identify a subset of patients with an inflammatory phenotype who may be at risk of rapid OA progression [108].

An in vivo murine model of monosodium iodoacetate-induced OA enabled examination of whole blood gene expression profiles with the development of OA, identifying time-dependent differential gene expression and therefore potential stage-specific genomic biomarkers of OA progression. Additionally, utilising the gene expression dataset from the study conducted by Ramos et al [107], common differential gene expression changes in both OA patients and the murine model dataset were examined. The expression of TNK2, WDR37 and KCTD2 mRNA were found to be differentially expressed in both studies [109].

In addition to measurements of RNA as a source from PBMCs as potential OA genomic biomarkers, measurements of DNA or DNA characteristics from PBMCs may also function as genomic biomarkers of OA. PBMC DNA methylation examination identified differential methylation of 44 CpG sites
between PBMC DNA obtained from OA patients with rapid-progressive radiographic knee OA and PBMC DNA from patients with non-progressive radiographic knee OA. Differential DNA methylation may therefore serve as a genomic biomarker for predicting rate of OA radiographic progression in symptomatic knee OA patients [110].

Despite the relatively low number of studies conducted to date which examine extracellular genomic material in OA, some of the studies identify the same genomic marker in individuals with knee OA. The expression of miR-146a was found to be dysregulated in studies of the synovial fluid [67,68], as well as plasma [72] and PBMCs [74], indicating that miR-146a may be a favourable extracellular genomic biomarker for identifying individuals with knee OA. The observed expression of miR-146a in differing body fluids of individuals with knee OA indicates the usage of several body fluids for examining the levels of miR-146a in the potential diagnosis of knee OA. The studies outlined demonstrate the recent efforts to identify soluble extracellular genomic biomarkers of OA. Larger cohorts are required to examine and validate these potential extracellular genomic biomarkers of OA.

### 3.4 Liquid biopsies - the potential for genomic biomarkers of OA in diagnostic testing

Liquid biopsy is a non-invasive tool for analysing different components of biological fluids to make a diagnosis without the need of invasive techniques [111]. Components of body fluids such as cfDNA, RNA subtypes including mRNA, miRNA and lncRNA, proteins, exosomes, cfDNA characteristics including methylation, integrity, mutation and copy number can be measured following liquid biopsy [112]. Liquid biopsy provides a non-invasive approach which is associated with decreased morbidity and can be performed successively offering temporal disease monitoring [112]. In 2015 liquid biopsy was listed as a top ten technology breakthrough by the MIT Technology Review [113]. Liquid biopsies are already used in the clinic and some of the diagnostic tests which currently employ liquid biopsies for measuring extracellular genomic material are outlined below.

Exosome Diagnostics produced the first blood-based liquid biopsy cancer diagnostic, ExoDx Lung (ALK), which measures exosomal-derived EML4-ALK fusion transcripts from the plasma of patients with non-small cell lung cancer [114]. Exosome Diagnostics also produced the first urine-based liquid biopsy diagnostic, ExoDx Prostate (IntelliScore), which detects exosomal-derived RNA of ERG, PCA3 and SPDEF in the urine of men with prostate-specific antigen, to discriminate high grade from low-grade prostate cancer and benign disease [115]. The AlloMap test, marketed by CareDx, is a non-invasive blood test to analyse RNA from PBMCs for the expression of 20 genes, to assess the risk of moderate/severe acute cellular rejection in cardiac transplant recipients [116]. Recently CareDx has brought to market another non-invasive blood test, AlloSure, which measures donor-derived cfDNA to determine active rejection of kidney allografts in renal transplant recipients [117]. All of the genomic liquid biopsy based diagnostic assays outlined above offer non-invasive screening over previously available invasive procedures. Measurement of extracellular genomic material and measurements of
characteristics of genomic material from body fluids including synovial fluid and blood provides a promising method for the identification of genomic biomarkers of early-onset OA and in monitoring OA disease progression. Ultimately liquid biopsies such as synovial fluid and blood sampling could be used in the diagnosis and monitoring of OA disease progression in the future.

4. Extracellular carriers of genomic material as potential indicators of OA disease progression

Stability of extracellular RNA is likely to be attributable to association of RNA with extracellular carriers such as exosomes, microvesicles, apoptotic bodies, lipoproteins or protein complexes [39]. EVs can be categorised into exosomes, microvesicles and apoptotic bodies and are generated through distinct cellular mechanisms. Exosomes refer to EVs that are of endosomal origin. Endosomes constitute part of the network which guides cell contents to either lysosomes for lysosomal degradation or the cell surface membrane for secretion. Fusion of intracellular multivesicular bodies, which constitute the endosomal pathway, with the plasma membrane results in secretion of exosomes into the extracellular space [118]. Microvesicles are of plasma membrane origin and are formed through the budding and fission of the plasma membrane [118]. Apoptotic bodies are only formed during apoptosis. Cells undergoing programmed cell death follow a process of nuclear chromatin condensation, followed by membrane blebbing and the dissolution of the cell contents into membrane enclosed vesicles, known as apoptotic bodies [118].

In addition to EVs as carriers of extracellular RNA are lipoproteins and protein complexes. Lipoproteins are an assembly of phospholipids, apolipoproteins, cholesterol and a core containing triglycerides and cholesterol esters [119]. Human high density lipoproteins (HDLs) have been shown to contain small RNAs and have been suggested to function in intercellular communication through transportation of miRNA [120]. Extracellular miRNA has also been shown to be associated with argonaute proteins [66,121] and miRNA association with argonaute proteins increases miRNA stability [122]. During intracellular miRNA mediated mRNA targeting, mature miRNAs complex with argonaute proteins to form a ribonucleoprotein complex, the RNA-induced silencing complex (RISC) [123], therefore miRNA association with argonaute proteins in the extracellular environment is not unexpected.

With regards to using extracellular genomic material as potential soluble genomic biomarkers of OA, the associated nucleic acid carrier and therefore the pathway of secretion of genomic material from any of the affected cells of the joint into extracellular environments may be extremely relevant to identifying specific stage-related OA genomic biomarkers. The extent to which intracellular nucleic acids are released from cells non-selectively and the extent to which intracellular nucleic acids are predetermined and sorted for extracellular secretion in a paracrine manner is unknown [124]. Two hypotheses have been proposed for the existence of extracellular miRNA: the existence of selective miRNA export systems have been proposed and alternatively the concept that extracellular miRNA are
merely by-products of cell activity and cell death [125] and therefore released non-selectively. But it is likely that intracellular RNA can be released both selectively and non-selectively.

Following cell death, tissue-specific RNAs may be released non-selectively, whereby the elevated expression of extracellular RNA may reflect the high abundance of RNA that were located in the cells of origin. While RNA release may be as a consequence of leakage from damaged cells undergoing apoptosis and/or necrosis, RNA release may also have a biological function, in which case RNA may be selectively released utilising specific intracellular membrane trafficking pathways. Approximately 90% of circulating miRNAs have been suggested to be associated with a ribonucleoprotein complex, while approximately 10% have been suggested to be vesicle-associated [121]. Turchinovich et al proposed that the majority of circulating miRNA that are bound to the ribonucleoprotein Ago2, are likely to be by-products of dead cells [66]. In late-stage OA, chondrocyte apoptosis is evidenced by hypocellularity and lacunae emptying [98]. If the majority of circulating miRNA that are bound to Ago2 are likely to be by-products of dead cells [66], measurements of only non-vesicle Ago2-associated miRNAs in the extracellular fluids of OA patients could serve as potential biomarkers for apoptosis of specific joint cells and therefore give a potential indication of disease progression.

5. Future challenges for genomic biomarkers of osteoarthritis

A biomarker can be defined as anything which enables an observation of medical state to be made from outside of the patient, which can be measured both accurately and reproducibly [126], a defined characteristic which when measured gives an indication of normal biological processes, pathogenic processes and responses to an intervention or exposure. Recent efforts have identified that extracellular nucleic acids, both carrier-associated and non-carrier associated, including miRNA, lncRNA, snoRNA and DNA have been identified to exhibit differential expression in extracellular environments in OA and could therefore serve as appropriate soluble genomic biomarkers of OA. Future studies to examine these identified potential genomic biomarkers for reproducibility in larger validation cohorts are required.

While efforts are required to advance the identification of biomarkers in OA, efforts will also be required to identify individuals at-risk of developing OA through the implementation of screening programmes as routine procedure in clinical practice. A limitation of soluble biomarker discovery is the asymptomatic nature of pre-OA which limits the point in time at which an individual can be directed for examination. Chu et al suggest that longitudinal studies of human joint-injured cohorts provide the opportunity to characterise and define pre-OA because these populations have joint pathologies whereby OA is likely to develop [127]. With regards to identifying a genomic biomarker for pre-OA, one idea could be the implementation of a longitudinal study with use of extracellular nucleic acid profiling in individuals who have been subject to a joint injury. For example, as discussed previously, Zhang et al identified extracellular ncRNAs one year after ACL injury and examined the correlation of
serum-derived ncRNA concentration with extent of cartilage damage [86]. Studies need to focus on identification and stratification of at-risk asymptomatic individuals without radiographic OA for longitudinal studies with cohorts which are likely to have a predisposition of developing a specific phenotype of OA.

An injury driven phenotype is just one subtype of OA. Given the complex nature of OA, it is likely that no single biomarker will be appropriate for all phenotypes of OA. In addition to the differing phenotypes of OA is also the anatomical location of OA. At present, OA can be categorised into primary (idiopathic) OA and secondary OA. Secondary OA is associated with causes such as congenital or developmental deformities, preceded joint trauma or surgery, metabolic or endocrine disorders. Primary OA is most common and occurs in joints with no prior damage and can be further categorised into generalised and localised OA. Generalised OA (polyarticular) occurs in multiple joints, whereas localised OA is confined to 1 or 2 joints [128]. Most extracellular genomic material studies of OA to date, discussed in this review, examine the expression of extracellular genomic material in patients with only knee OA. While knee OA genomic biomarker discovery will be more accelerated, genomic biomarker discovery of OA of the hip, OA of the joints of the hand and generalised OA may lag behind. Future priorities need to categorise OA into disease subtypes according to potential underlying phenotype in addition to anatomical location/s, with the view to identify soluble biomarkers that are inherent to a specific type of OA.

Of the extracellular genomic material studies of OA to date, discussed in this review, only one study isolated RNA from the exosome fraction of synovial fluid [68], all of the other studies isolated total RNA from either synovial fluid, serum or plasma. The method of RNA isolation from extracellular fluid is likely to be relevant to identifying genomic biomarkers, with potential differences in extracellular RNA expression dependent on the isolation procedure utilised. However, at present, limitations exist in the lack of discrimination between extracellular nucleic acid carriers. Isolation of a pure EV fraction from extracellular fluids is challenging because non-vesicular entities such as lipoproteins and proteins complexes are not separated from EVs by isolation protocols such as centrifugation [129] and Ago-2 has been identified inside exosomes [130]. Furthermore there is much overlap in the sizes of specific EVs, with exosomes measuring between 40-200 nm in diameter, microvesicles measuring between 50-1000 nm in diameter and apoptotic bodies measuring between 500-2000 nm in diameter [131] which makes separation and isolation of subtypes of EVs difficult. Lack of defined nucleic acid carrier isolation and characterisation techniques may limit efforts to examine selective and non-selective release of nucleic acid from chondrocytes and synoviocytes. What is critical in extracellular genomic biomarker research is the necessity to appropriately document extracellular sources of genomic material with the implementation of strict guidelines for categorising nucleic acid carriers.

Personalised medicine is a rapidly growing field which is informed by the underlying molecular mechanisms of a disease in order of optimising preventative health care strategies, directing medical
care towards assessing disease risk and prevention [132]. A companion diagnostic is any in vitro diagnostic device or imaging tool that provides essential information for the safe and effective use of a corresponding therapeutic product [133]. The importance of identifying a soluble biomarker/s of OA, is the potential that biomarker may have in being utilised as a liquid biopsy-based companion diagnostic for screening participants in the detection of OA for inclusion in clinical trials, in the development of much needed DMOADSs [134].

**Expert commentary**

The definitive objective of OA biomarker research is to identify OA prior to radiographic evidence of OA. For the future, an annual biological test is necessary to detect pre- or early-onset OA while OA is either still curable or at a less advanced stage of the disease progress at which point potential DMOADSs will forestall the progression of the disease. All cells release nucleic acids into the extracellular environment and by association with RNA-binding proteins or lipoproteins or carriage within EVs, nucleic acids are thought to be protected from degradation. Biological fluids are more accessible than tissues and provide an appealing and non-invasive source for the measurement of extracellular genomic material. Future OA research prerequisites include identifying stable and reproducible extracellular genomic biomarkers of pre- and early-onset OA and extracellular genomic biomarkers which can discriminate between stages of OA progression, which may lead to the creation of OA specific non-invasive liquid biopsy diagnostics.

Documented to date, genomic material that has been observed in extracellular environments in OA, such as synovial fluid and plasma and/or serum includes cfDNA, miRNA, IncRNA and snoRNA. A diverse composition of RNA subsets have been identified in EVs including tRNA fragments, rRNA, vault RNA, circRNA, piRNA, Y RNA (Table 2) [36] and cell-free telomeric repeats-containing RNA (cfTERRA) [135], all of which could be present and exhibit differential expression in the extracellular fluids of OA patients. Future studies to investigate the presence of RNA subsets in OA may reveal additional RNA candidates that could serve as genomic biomarkers of OA. Other epigenomic alterations such as histone modifications [136], enabling measurements of histone methylation, acetylation, phosphorylation and ubiquitination in PBMCs of OA patients, may also function as potential genomic biomarkers. Extracellular fluid analysis holds much promise for the identification of genomic biomarkers of OA, but a combination of techniques inclusive of genomics, proteomics, and metabolomics used together, is likely to identify a more powerful biomarker panel capable of detecting OA early.

Recently liquid biopsy has gained much attention in the cancer sciences field, but this novel technique set to revolutionise diagnostics has far reaching implications not just in the field of oncology but in other diseases too. As advances are made in OA research, in the future not only could liquid biopsy be a useful method for OA disease progression monitoring, but also in the initial screening of individuals
for OA detection and also in the stratification of individuals according to OA phenotype, made possible by anticipated discovery of reliable informative single nucleotide polymorphisms.

Beyond the scope of this review is the potential link that mechanisms of selective and non-selective release of genomic material from cells of the OA joint may have in OA disease pathogenesis. In the search for a suitable extracellular biomarker of OA opens up questions about the potential role that conventional and unconventional secretory pathways may have in secreting genomic material to the extracellular environment. In addition to exploring intracellular trafficking pathways including exocytosis, microvesical shedding and exosome secretion, exploring exosome secretion as a result of plasma membrane fusion with secretory lysosomes and amphisomes in addition to multivesicular bodies may reveal links to autophagy and an indication of underlying cellular health in OA. The extent of which specific nucleic acid carrier subtypes are released from cells of the joint may have indications for cellular homeostatic status. The type of extracellular carrier associated with extracellular genomic material and the mode of secretion may indicate how cells of the joint function in health and disease, revealing roles for genomic material in cell-cell communication and the potential to reveal more about the underlying pathophysiology of OA. A more profound understanding of the mechanisms of release of genomic material into the extracellular environment in the cells of the joint is required.

**Five-year view**

At present, the field of extracellular genomic biomarker research is primitive and in the next few years, extracellular genomic biomarker research is anticipated to grow, with the identification of novel subsets of extracellular RNA isolated from the extracellular fluids of OA patients. *In vitro* cell and explant culture systems will likely be routinely used to validate and further explore identified extracellular genomic biomarkers. It is expected that there will be a focus on EV-derived genomic material and it is anticipated that technologies utilised to extract extracellular genomic material will continue to grow and move towards isolation of pure fractions of extracellular nucleic acid carrier subtypes, such as subsets of EVs. As a greater understanding regarding extracellular genomic material develops, pathways of intracellular cell trafficking of genomic material to the plasma membrane will also develop and may reveal more about the underlying dysregulated mechanisms which may be involved in the pathophysiology of OA. In addition to expected identification of extracellular metabolomic and proteomic biomarkers of OA, with the identification of differential measurements of extracellular genomic material in OA which could function as genomic biomarkers of OA, the number of human biomarker studies to assess panels of a combination of biomarkers from different body fluids is expected to increase. It is anticipated that in the event of successful identification of soluble biomarkers for diagnosing OA, soluble biomarkers will be in the first instance adjunct to and then definitively in place of radiographic measurements as the regulatory endpoint in DMOAD development.

**Key Issues**
- Osteoarthritis (OA) is a chronic and debilitating disease which affects approximately a third of the UK’s population that are aged over 45 years and a diagnosis of OA is currently based on radiographic findings at which point structural changes are observed in the affected joint.
- While recent efforts have looked to address the notion that OA is indeed a serious disease, much more needs to be done to drive forwards biomarker discovery and the development of DMOADs.
- The OA disease process is estimated to begin 20 years prior to detection by radiographic joint space narrowing with the likelihood that molecular indicators exist at the cellular level only, a stage that can be defined as pre-OA.
- No soluble biomarkers exist which can detect OA or determine the progression of the disease. Asymptomatic OA is a common phenomenon. Identification of suitable soluble biomarkers would potentially help identify asymptomatic individuals and also pre-OA in routine screening and to assess disease progression.
- Examination of extracellular body fluids for biomarkers offers a non-invasive method for assessing disease state. Body fluids including synovial fluid and blood which can be analysed for measurements of extracellular genomic material have the potential to serve as liquid biopsies in the diagnosis of OA.
- Several recent studies observe the differential expression of subsets of ncRNA in extracellular fluids including synovial fluid, plasma and serum of OA patients, indicating the potential use of ncRNAs such as miRNAs, lncRNAs and snoRNAs as diagnostic and prognostic biomarkers for OA.
- Studies of PBMCs in OA has observed the differential expression of miRNA, mRNA and differential measurements of DNA methylation, suggesting that isolation of PBMCs could provide a source of genomic material as diagnostic and prognostic biomarkers for assessing OA disease.
- RNA exists in extracellular fluids by association with a carrier, which protects the RNA from degradation. Carriers of genomic material may include lipoproteins, protein carriers and EVs such as exosomes, microvesicles and apoptotic bodies.
- The type of carrier that genomic material may be associated with may give indications about how intracellular genomic material is secreted or released to the extracellular environment. How genetic material is secreted or released from cells of the joint including chondrocytes and synoviocytes, with regards to specific intracellular trafficking pathways, may give an indication to the underlying health of the cell of origin and reveal more about OA pathophysiology.
- Drug and biomarker development are interdependent. The development of DMOADs requires a biomarker which can detect OA and predict OA disease progression. In the future,
identification of OA before radiographic evidence of OA, using extracellular-derived soluble biomarkers, could see the use of effective DMOADs forestalling the onset of OA.
Extracellular carrier-associated and extracellular carrier-free genomic material from body fluids with the potential to serve as genomic biomarkers in OA. To date extracellular genomic material has been identified in synovial fluid, plasma and serum. Peripheral blood mononuclear cells (PBMCs) have been isolated from blood and identified to exhibit differential expression of mRNA and miRNA in OA patients. PBMCs include monocytes, lymphocytes, dendritic cells and natural killer cells. Extracellular genomic material found in synovial fluid, plasma and/or serum may be contained inside extracellular vesicles including apoptotic bodies, microvesicles and exosomes or associated with a protein complex such as Argonaute2 or high density lipoproteins (HDLs) or not associated with a carrier, such as cell-free DNA (cfDNA). TMJ – Temporomandibular joint. ↑ – Up-regulated expression or concentration. ↓ – Down-regulated expression or concentration. Documented studies have observed the differential expression/concentration/level of specific extracellular genomic material between rheumatoid arthritis patients (RA) and patients with knee OA, between non-OA patients and patients with mild and severe knee OA, between non-OA patients and mild OA patients, between non-OA patients and patients with TMJ OA.
DNA

Non-coding DNA (98-99%)

Protein coding DNA (1-2%)

mRNA

Protein

Non-functional non-coding DNA

Functional non-coding DNA

Cis-regulatory elements

Scaffold/matrix (S/MAR) attachment regions

Introns

Telomeres

Functional non-coding RNA (ncRNA)

miRNAs

IncRNAs

snoRNAs

snRNAs

piRNAs

Y RNA

circRNA

vtRNA

tRFs

rRNA

tRNA
Figure 2. DNA encodes both protein coding genes and non-coding DNA. Only a small proportion of DNA is protein coding DNA, in which the DNA is transcribed to mRNA which in turn is translated to protein. Non-coding DNA includes non-functional DNA which has no known function and functional DNA which includes DNA sequences which have known functions and include for example: cis-regulatory elements, scaffold/matrix (S/MAR) attachment regions, introns, telomeres and DNA which encodes genes for non-coding RNA. Functional non-coding RNA encompasses RNAs which are not translated into protein but have functional roles and include miRNAs, IncRNAs, snoRNAs, snRNAs, piRNAs, vtRNAs, Y RNA, circRNAs, tRFs, rRNAs and tRNAs. Many ncRNAs have housekeeping and regulatory roles. SnRNAs for example have roles in processing pre-mRNA during splicing and tRNA is fundamental to protein biosynthesis in ribosomes. MiRNAs have roles in regulating gene expression, via post-transcriptional regulation of mRNA, generally via binding to the 3'UTR of target mRNA. One function of circRNA is the regulation of gene expression through the targeting of miRNA, enabling derepression of miRNA targeted mRNA.
<table>
<thead>
<tr>
<th>External Extra</th>
<th>Anatomical location of OA</th>
<th>Genomic material source</th>
<th>OA Study</th>
<th>Extracellular genomic material expression profiling results</th>
<th>Extracellular genomic material expression/level/concentration following validation</th>
<th>Potential indication for genomic material</th>
<th>Ref.</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>MicroRNA</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Human synovial Fluid</strong></td>
<td>Knee</td>
<td>Total RNA from synovial fluid.</td>
<td>Differential expression of synovial fluid miRNAs between radiographic early-stage and late-stage OA (assessed by KL grading).</td>
<td>14 miRNAs differentially expressed in miRNA PCR array.</td>
<td>miR-23a-3p ↑ miR-24-3p ↑ miR-27a-3p ↑ miR-27b-3p ↑ miR-29e-3p ↑ miR-34a-5p ↑ miR-186-5p ↑ miR-378a-5p ↑ (Significantly upregulated in late-stage OA compared to early-stage OA).</td>
<td>Prognostic biomarkers of knee OA disease progression.</td>
<td>[69]</td>
</tr>
<tr>
<td><strong>Conditioned media from human synovial explant culture</strong></td>
<td>Synovial explants from the knee joint</td>
<td>Total RNA from cell culture supernatant.</td>
<td>Differential expression of cell culture supernatant miRNAs in response to IL-1β stimulation.</td>
<td>-</td>
<td>miR-23a-3p ↑ miR-27b-3p ↑ (Significantly upregulated with IL-1β stimulation).</td>
<td>Diagnostic biomarkers of inflammatory knee OA.</td>
<td>[69]</td>
</tr>
<tr>
<td><strong>Human synovial Fluid</strong></td>
<td>Knee</td>
<td>Total RNA from synovial fluid.</td>
<td>Differential expression of synovial fluid miRNAs between patients with OA and RA (assessed by 1986 classification of knee OA in diagnostic criteria of the American Rheumatism Association [137]).</td>
<td>-</td>
<td>miR-16 ↓ miR-146a ↓ miR-155 ↓ miR-223 ↓ (Lower concentration in OA compared to RA).</td>
<td>Diagnostic biomarkers of knee OA and RA. Assessment of concentration required to differentiate knee OA from RA - lower concentration in OA than RA.</td>
<td>[67]</td>
</tr>
<tr>
<td><strong>Human plasma</strong></td>
<td>Knee</td>
<td>Total RNA from plasma</td>
<td>Differential expression of plasma miRNAs between patients with OA, RA and healthy controls (HC) (assessed by 1986 classification of knee OA in diagnostic criteria of the American Rheumatism Association [137]).</td>
<td>-</td>
<td>miR-16 ↓ miR-132 ↓ (Significantly lower concentration in OA compared to HC).</td>
<td>miR-132 - diagnostic biomarker of knee OA and RA, but not inherent to knee OA as no differences observed when OA compared to RA. miR-16 – diagnostic biomarker of knee OA.</td>
<td>[67]</td>
</tr>
<tr>
<td><strong>Human synovial Fluid</strong></td>
<td>Knee</td>
<td>Total miRNA from synovial fluid-derived extracellular vesicle-derived miRNAs between males</td>
<td>Differential expression of synovial fluid-derived extracellular vesicle-derived miRNAs between</td>
<td>114 miRNAs differentially expressed in males</td>
<td>miR-26a-5p ↓ miR-146a-5p ↓ miR-6821-3p ↓</td>
<td>Diagnostic biomarkers of knee OA</td>
<td>[68]</td>
</tr>
<tr>
<td><strong>Human synovial fluid</strong></td>
<td><strong>Knee</strong></td>
<td>extracellular vesicles</td>
<td>patients with OA and without OA and between male and female.</td>
<td>144 miRNAs differentially expressed in females in miRNA array <strong>miR-504-3p ↑</strong> (Only miRNA to be up-regulated in both male and female).</td>
<td>(Significantly down-regulated expression in female OA compared to female non-OA).</td>
<td>in females only.</td>
<td></td>
</tr>
<tr>
<td>--------------------------</td>
<td>----------</td>
<td>-----------------------</td>
<td>-------------------------------------------------</td>
<td>----------------------------------------------------------------------------------</td>
<td>------------------------------------------------------------------------------------------</td>
<td>-------------------------</td>
<td></td>
</tr>
<tr>
<td>Human serum</td>
<td>Knee and/or hip</td>
<td>Total RNA from serum.</td>
<td>Differential expression of serum miRNAs between patients with OA-related knee or hip arthroplasty compared to individuals without arthroplasty.</td>
<td>12 miRNAs differentially expressed in miRNA array.</td>
<td><strong>let-7e ↓</strong> (with increasing number of hip/knee joint replacement surgeries).</td>
<td>Diagnostic biomarker of knee and/or hip OA and likelihood for need of knee and/or hip arthroplasty.</td>
<td></td>
</tr>
</tbody>
</table>
| Human plasma             | Knee     | Total RNA from plasma. | Differential expression of plasma miRNAs between patients with radiographic knee OA and patients without clinical diagnosis of OA (assessed by KL grading). | 12 miRNAs differentially expressed in miRNA array. | **miR-29c ↑**  
**miR-93 ↑**  
**miR-126 ↑**  
**miR-184 ↑**  
**miR-186 ↑**  
**miR-195 ↑**  
**miR-345 ↑**  
**miR-885-5p ↑**  
**miR-16 ↑**  
**miR-146a ↑** | Diagnostic biomarkers of knee OA. |
| Human plasma             | Knee     | Total RNA from plasma. | Differential expression of plasma miRNAs between patients with knee OA and healthy controls (HC) (assessed by 1986 classification of knee OA in diagnostic criteria of the American Rheumatism Association [137]). | 70 miRNAs differentially expressed in miRNA array. | **miR-19b-3p ↑**  
**miR-92a-3p ↑**  
**miR-122-5p ↑**  
**miR-486-5p ↑** | Diagnostic biomarkers of knee OA. |
| Human blood - PBMCs      | Knee     | Total RNA from PBMCs. | Differential expression of PBMC-derived miRNA between patients with RA, knee OA and healthy controls (HC) (assessed by 1986 classification of knee OA in diagnostic criteria of the American Rheumatism Association [137]). | - | **miR-146a ↑**  
**miR-155 ↑**  
**miR-181a ↑**  
**miR-223 ↑**  
(Significantly up-regulated expression in both RA and OA compared to HC). | Diagnostic biomarkers of knee OA and RA. Assessment of expression required to differentiate OA from RA – lower |
<table>
<thead>
<tr>
<th><strong>miRNA</strong></th>
<th><strong>IncRNA</strong></th>
<th><strong>snoRNA</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td>Differential expression of PBMC-derived miRNA between radiographic stages of OA (assessed by KL grading).</td>
<td>Differential expression of synovial fluid IncRNAs between patients with clinical signs and symptoms of OA and patients without (assessed by research diagnostic criteria for TMJ disorders).</td>
<td>Differential concentration of serum snoRNAs between patients one year post ACL injury with cartilage damage (assessed by MRI analysis using Whole Organ Magnetic Resonance Knee Scoring (WORMS)) and in healthy donors. &amp; Between patients one year post ACL injury developing cartilage damage and in patients one year post ACL injury without evidence of cartilage damage (assessed by MRI analysis using WORMS).</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Differential expression of serum snoRNAs between old sham mice and mice with experimentally induced destabilisation of the medial meniscus (DMM).</td>
</tr>
<tr>
<td><strong>miR-223 ↓</strong> (Significantly up-regulated expression in RA when compared to OA).</td>
<td><strong>HOTAIR ↑</strong> (Significantly up-regulated in OA compared to normal controls).</td>
<td><strong>SNORD116 ↑</strong></td>
</tr>
<tr>
<td><strong>miR-146a ↑ in early OA</strong></td>
<td></td>
<td><strong>U48 ↑</strong> (Significantly higher levels in patients with ACL injury with cartilage damage compared to healthy donors).</td>
</tr>
<tr>
<td><strong>miR-223 ↑ in early OA</strong></td>
<td></td>
<td><strong>U38 ↑</strong> (Significantly higher levels in patients with ACL injury with cartilage damage compared to healthy donors) &amp; Significant higher levels in patients with greater cartilage damage compared to patients with minor cartilage damage).</td>
</tr>
<tr>
<td><strong>miR-155 ↓ in early OA</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>miR-146a and miR-223 biomarkers of early OA miR-155 biomarker of late-stage OA.</td>
<td>Diagnostic biomarker of TMJ OA. [79]</td>
<td>Diagnostic biomarkers of cartilage damage. U38 can differentiate between minor and greater cartilage damage induced by ACL injury – U38 biomarker of early ACL induced damage. [86]</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Horse serum</td>
<td>Not stated in study</td>
<td>Serum</td>
</tr>
<tr>
<td>Human serum</td>
<td>Not stated in study</td>
<td>DNA from serum</td>
</tr>
<tr>
<td>Human synovial fluid</td>
<td>Not stated in study</td>
<td>DNA from synovial fluid</td>
</tr>
<tr>
<td>Human plasma</td>
<td>Not stated in study</td>
<td>cfDNA from plasma</td>
</tr>
</tbody>
</table>

### DNA

| Human blood | Knee | Total RNA from cellular fraction of blood. | Differential gene expression of total cellular blood between patients with mild OA and control subjects (assessed by arthroscopic scoring method [138]). | 3,543 differentially expressed genes in cDNA microarray. | HSPCA mRNA ↓ IKBKAP mRNA ↓ IL13RA1 mRNA ↓ LAMC1 mRNA ↓ TNFAIP6 mRNA ↓ PF4 mRNA ↓ (Significantly down-regulated in mild knee OA compared to controls). | Diagnostic biomarkers of mild knee OA. | [106] |
| Human blood | Knee | Total RNA from PBMCs. | Correlation of OA patients which identified as having up-regulated cytokine expression (from blood-derived PBMCs RNA) compared to OA patients identified as having low cytokine expression, and OA radiographic progression (assessed by) | - | IL-1β mRNA ↑ TNFα mRNA ↑ IL-6 mRNA ↑ IL-8 mRNA ↑ COX-2 mRNA ↑ (Significant up-regulation in inflammatory OA subclass). | Biomarkers of inflammatory knee OA with increased risk of rapid radiographic progression. | [108] |
| Human blood | Primary OA – at multiple joint sites in the hand, or at least two joints from hand, spine, knee or hip. | Total RNA from PBMCs. | Differential gene expression of blood-derived PBMCs between OA patients and non-OA control subjects (assessed by definition of OA diagnosis by the American College of Rheumatology [137]). | 679 genes and 15 ncRNAs differentially expressed in microarray. | GPR18 mRNA ↑ ZNF564 mRNA ↑ XNF20 mRNA ↑ RIPK1 mRNA ↑ ADRB2 mRNA ↑ PELO mRNA ↑ SIAH1 mRNA ↑ SLA mRNA ↑ PLEKHF1 mRNA ↑ CASP3 mRNA ↑ RABEP1 mRNA ↑ FADD mRNA ↑ CLK1 mRNA ↑ ESF2 mRNA ↑ (Significant up-regulation in OA compared to non-OA). | Diagnostic biomarkers of OA. | [107] |
### Table 1

Extracellular carrier-associated and extracellular carrier-free genomic material with the potential to serve as genomic biomarkers in OA. † - Up-regulated expression/concentration/level in OA. ↓ - Down-regulated expression/concentration/level in OA.

<table>
<thead>
<tr>
<th>Blood Type</th>
<th>Joint</th>
<th>Method</th>
<th>Genes/DNA</th>
<th>Biomarkers</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rat blood</td>
<td>Monosodium iodoacetate (MIA)-induced OA of the knee joint.</td>
<td>Time course gene expression of MIA induced OA and comparison of microarray dataset to microarray dataset obtained from the human GARP study [107].</td>
<td>OSM mRNA ↓ GSTM2 mRNA ↓ ERG1 mRNA ↓ IL8 mRNA ↓ (Significant down-regulation in OA compared to non-OA).</td>
<td>Diagnostic biomarkers of OA. [109]</td>
</tr>
<tr>
<td>Human blood</td>
<td>Knee</td>
<td>Differential methylation of CpG sites of blood-derived PBMC DNA between patients with rapid progressive OA and patients with non-progressive OA (assessed by radiographic joint space loss at follow-up and need for arthroplasty).</td>
<td>Tnk2 mRNA Wdr37 mRNA Kcid2 mRNA (Differentially expressed in both rat model and OA patients).</td>
<td>Methylation status as a biomarker for predicting future radiographic progression of knee OA. [110]</td>
</tr>
</tbody>
</table>

<p>| Human blood | Knee | DNA from PBMCs. | 44 CpG sites (Differentially methylated between rapid and non-progressive OA). | Methylation status as a biomarker for predicting future radiographic progression of knee OA. [110] |</p>
<table>
<thead>
<tr>
<th>Functional non-coding RNA (ncRNA)</th>
<th>Definition and function</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>MicroRNA (miRNA)</strong></td>
<td>MiRNA are single stranded ncRNA of approximately 22 nucleotides in length. MiRNA function by binding to locations within mRNA, typically within the 3'UTR, following RNA transcription [47]. The seed sequence of the miRNA is responsible for miRNA-mRNA binding and is a sequence of at least 6 nucleotides positioned at the 5’ end of the miRNA and base pairs with perfect complementarity to a target sequence within the mRNA 3'UTR [139]. The binding of miRNA to mRNA, in the cytoplasm, prevents mRNA translation in the synthesis of protein. MiRNA is therefore capable of down-regulating protein expression through effective mRNA targeting [48].</td>
</tr>
<tr>
<td><strong>Long non-coding RNA (lncRNA)</strong></td>
<td>The definition of lncRNAs is limited to ncRNAs of more than 200 nucleotides in length which lack protein coding capability [75]. LncRNAs have been classified according to molecular mechanism and can be subtyped as having signalling, decoy, guide, scaffold and enhancer roles [77]. LncRNAs can also be subtyped according to their position and direction of transcription relative to other genes and can be classified into subtypes including intergenic, intronic, antisense, overlapping, processed and bidirectional lncRNAs [78]. For example lncRNAs include some cis-natural antisense transcripts (NATs), antisense ncRNA of sequence complementarity to coding RNA from the same genomic locus, which function to regulate expression of sense transcripts. LncRNAs include some trans-NATs, antisense ncRNA of sequence complementarity to coding RNA from a distal genomic locus, which functions to regulate sense transcripts. LncRNAs include long intergenic ncRNA (lincRNA) which are encoded in intergenic regions of DNA, sense overlapping lncRNAs which are transcribed from the same strand of DNA as another transcript, sense intronic lncRNAs which are encoded in introns of coding genes and processed lncRNAs which can be spliced and/or polyadenylated [76]. Different lncRNAs are involved in different biological functions in both the nucleus and cytoplasm. LncRNAs have been observed to have roles in chromatin modification through recruitment of chromatin modifying complexes. LncRNAs have been observed to have roles in transcriptional regulation, for example by functioning as decoys for transcription factors, functioning as transcription factor co-regulators and by inhibiting formation of transcription machinery. LncRNAs have also been observed to have roles in post-transcriptional regulation, for example by interacting with splicing factors and hindering spliceosome formation, by increasing or decreasing mRNA stability, by masking miRNA binding sites in mRNA 3'UTR and by acting as miRNA sponges to prevent miRNA binding to target mRNA [140].</td>
</tr>
<tr>
<td><strong>Small nuclear RNA (snRNA)</strong></td>
<td>SnRNA, located in the nucleus, function in the process of splicing, aiding in the processing of pre-mRNA to mature mRNA in the removal of introns. SnRNAs associate with small nuclear ribonucleoproteins to form the spliceosome. The snRNAs function as ribozymes, mediating base pairing with RNA [141]. Spliceosomal snRNAs include U1, U2, U4, U5, and U6 have roles in substrate recognition at splice sites, in lariat formation and splicing catalysis [142].</td>
</tr>
<tr>
<td><strong>Small nucleolar RNA (snoRNA)</strong></td>
<td>SnoRNAs, a sub-class of snRNA located in the nucleus, are involved with post-transcriptional modification by means of methylation or pseudouridylation of non-coding RNA including snRNA, rRNA and tRNA. SnoRNAs are divided into two classes, SNORDs and SNORAs, dependent on their structural features and sequence [84]. SNORDs are box C/D snoRNAs, are approximately 70-120 nucleotides in length and contain two conserved elements: box C (PuUGAUGA) located at the 5’- termini and box D (CUGA) located at the 3’- termini. The sequence elements form</td>
</tr>
</tbody>
</table>
stem-bulge-stem folding domains which serve as scaffolds for small nucleolar ribonucleoprotein (snoRNP) assembly. The Box C/D snoRNA: snoRNP complex catalyses site specific 2’-O-methylation in target RNA [85]. SNORAs are box H/ACA snoRNAs are hairpin-hinge-hairpin-tail structures which contain two conserved elements: box H (ANANNA) (N is any nucleotide) and box ACA which is a trinucleotide located 3 nucleotides upstream of the 3’-terminus. Box H/ACA snoRNAs associate with snoRNPs which catalyse pseudouridylation in target RNA [85]. Cajal-body-specific RNAs (scaRNAs) are snoRNAs which commonly contain features of both SNORDs and SNORAs and also contain an element called the Cajal body box (CAB box). ScaRNAs are located within Cajal bodies and are involved with methylation and pseudouridylation of spliceosomal snRNAs [143]. In addition to their role in the modification of other ncRNA, snoRNAs can be further processed to generate smaller fragments known as sno-derived RNAs (sdRNAs) which possess miRNA-like features [144], [145].

<table>
<thead>
<tr>
<th>Transfer RNA (tRNA)</th>
<th>tRNAs, approximately 76 nucleotides in length, are fundamental to the process of protein biosynthesis and participate in translating codons of mRNA into corresponding amino acids and for each canonical amino acid a subset of tRNA species exist [146]. tRNAs are aminoacylated by aminoacyl-tRNA synthetase and the resultant aminoacyl-tRNAs are directed to the ribosome where the tRNA locates to the tRNA binding site. At the tRNA binding site the anticodon sequence in the tRNA base-pairs with the codon sequence in the mRNA, which specifies the aminoaylated attached amino acid [147]. tRNAs exhibit a cloverleaf-shaped secondary structure which ultimately folds into an L-shaped tertiary structure which is fundamental to its role [146].</th>
</tr>
</thead>
<tbody>
<tr>
<td>Transfer RNA-derived RNA fragments (tRFs)</td>
<td>tRFs are processed from pre-tRNA and mature t-RNA, generated from all parts of the tRNA molecule. tRFs were assumed to be inert but have recently been observed to be produced to function in processes including stress response, translation inhibition and proliferation [148].</td>
</tr>
<tr>
<td>Ribosomal RNA (rRNA)</td>
<td>rRNA ranges in size from 120 to 5000 nucleotides and are a key component of ribosomes and therefore play a fundamental role in protein biosynthesis. Ribosomes are composed of proteins and several different rRNA molecules which are organised into a small subunit and a large subunit [147]. The ribosome comprises of rRNA which possesses ribozymal activity. The small subunit of the ribosome mediates the interactions between the tRNA anticodons and mRNA codons and the large subunit catalyses the formation of the peptide bond, with rRNA 28S being responsible for peptidyl transferase activity of the ribosome [149].</td>
</tr>
<tr>
<td>Circular RNA (circRNA)</td>
<td>circRNA is non-linear, with the 3’ and 5’ ends joined together resulting in a circular structure, generated from the backsplicing of introns and/or exons. circRNA has been proposed to be stable as a result of its structure which prevents degradation by RNA exonucleases and therefore is abundant in the cytoplasm. CircRNAs have been suggested to function as miRNA sponges and therefore may play a role in gene expression regulation by regulating miRNA expression [150]. Additional potential roles for circRNA include interacting with RNA binding proteins, modulating mRNA stability and regulating gene transcription [151].</td>
</tr>
<tr>
<td>Piwi-interacting RNA (piRNA)</td>
<td>piRNAs are 24-31 nucleotides in length and interact with PIWI proteins. PIWI proteins and piRNA constitute piRNA-induced silencing complexes which function to repress transposons [152]. piRNAs are enriched in germline tissues, and through the silencing of mobile elements, are thought to prevent accumulation of genome changes [153].</td>
</tr>
</tbody>
</table>
| Vault RNA  
| (vtRNA) | vtRNAs are approximately 80 to 150 nucleotides in length and are associated with vault particles. Vault particles are large cytoplasmic ribonucleoproteins which may have roles in intracellular nucleocytoplasmic transport [154], intracellular detoxification, apoptosis resistance, signalling, DNA damage repair, innate immune response and formation of the nuclear pore complex [155]. |
| Y RNA | Y RNAs vary in length between approximately 83 and 112 nucleotides and are folded into conserved stem-loop-structures. Functions for Y RNA have been proposed and include cell stress response and proliferation regulation, DNA replication and small RNA quality control. Y RNA are associated with different core proteins, for example the association of Y RNA with Ro60 [156] |

**Table 2.** Definitions of functional ncRNAs.
References


**The Osteoarthritis Research Society International white paper submitted to the FDA to highlight the seriousness of OA and the need for effective therapies.**


   • This study introduces the definition of pre-OA, OA which is established at the cellular level only, prior to changes in cartilage structure.


• This study demonstrates the existance of RNA within the extracellular space, including miRNA, piRNA ad snRNA.


**This study demonstrates the examination of extracellular vesicles from synovial fluid for aberrant miRNA expression. A down-regulated expression of miRNAs in extracellular vesicles from the synovial fluid of female knee OA patients was observed.**


**In 2015 the MIT Technology Review listed liquid biopsy as a top ten technology breakthrough. In the future, with the identification of a soluble biomarker for OA, liquid biopsies could be utilised in the screening and monitoring of OA.**


