REVIEW

**cDNA arrays in degenerative arthritis research**

Thomas Aigner†, Pia M Gebhard, Robert Kueffner, Hongwei Zhang & K Wayne Marshall

†Author for correspondence
University of Leipzig, Osteoarticular and Arthritis Research, Department of Pathology, FRG, Liebkgrt. 26, 04103 Leipzig, Germany
Tel.: +49 341 971 5000;
Fax: +49 341 971 5009;
Thomas.Aigner@medizin.uni-leipzig.de

Keywords: anabolism, arthritis, blood, cartilage, catabolism, diagnosis, functional genomics, gene expression, osteoarthritis, proteases

Gene expression & osteoarthritis

In recent years, osteoarthritis (OA) research has shifted from investigating primarily the biochemical aspects of articular cartilage matrix destruction to focusing on the molecular and genetic aspects of the disease. Articular chondrocytes determine cartilage matrix homeostasis, and chondrocyte biology plays a key role in the initiation, progression and outcome of OA. Functional genomics is a challenging new way to address complex diseases such as OA and rheumatoid arthritis at the molecular level. This review will explore the potential of genomic technologies in the analysis of degenerative joint disease, including diagnosis and therapy.

Although OA is primarily a biochemical problem (i.e., the degradation of the extracellular cartilage matrix), a deeper understanding of molecular events within the tissue cells – the articular chondrocytes – will help to clarify pathogenetic mechanisms. Understanding the molecular biology of chondrocytes will also aid in the identification of new diagnostic markers and cellular targets for therapeutic intervention. Gene expression profiling represents an innovative and challenging approach complementing, although not replacing, classical research. A basic consideration for designing a gene profiling study is to decide which material should form the basis of analysis: the articular cartilage, the synovial membrane or the peripheral blood cells and, potentially in the future, also subchondral bone. As will become clear during this review, different starting points offer different potentials in terms of insights to be gained and complexities and problems to be faced. Overall, as in many other studies of functional genomics, a better definition of the patient population examined is essential in order to bring a comprehensive understanding of the disease process.

General issues

Theoretically, functional genomics promises to investigate the whole transcriptome in parallel. However, a number of technical, statistical, bioinformatic and disease-related issues limit the extent to which functional genomics has lived up to its promise. Firstly, issues such as sensitivity and specificity are still unresolved. Secondly, none of the techniques can account for (post)translational control of molecules and molecular activities (i.e., activation, degradation and sequestering) nor for any important molecules that are not proteins at all (e.g., eicosanoids), and most technologies do not account for the splicing variants that occur in many genes in vivo. Thirdly, optimal bio-statistical evaluation methodologies have yet to be determined, from simple and basic procedures such as normalization of given data sets [1] to more complex issues such as bioinformatic network formation and data evaluation, which is the path from pure data to information on questions and systems of interest [2]. In addition, many tissue sample-related issues need to be resolved. For example, normal (or early diseased) cells/tissue (for comparison) is difficult to acquire. In addition, there is no clear consensus on the classification and staging of OA; defined criteria are needed in order to group and cluster results to the stage of disease progression more appropriately.

---

**cDNA microarray technology is a powerful tool that allows the expression profiling of thousands of mRNA transcripts simultaneously. Despite technical and analytical challenges, the application of gene expression profiling in degenerative arthritis research will provide a better understanding of the disease at the molecular level and lead to new diagnostic markers and therapeutic targets. Profiling the gene expression of articular cartilage will lead to the identification of genes involved in cartilage matrix homeostasis and in disease initiation, progression and outcome. Analysis of gene expression patterns of synovium in rheumatoid arthritis and osteoarthritis (OA) may help identify targets for future disease management. Recently, gene profiling strategies have been applied to peripheral blood from subjects with OA, a novel development in OA diagnosis.
Gene expression profiling of cartilage degeneration

Adult human articular cartilage offers the advantage over other tissues that it contains only one cell type, the chondrocyte. Although there is clearly some heterogeneity regarding the zones and areas the cells occupy, the gene expression levels detected in cartilage can at least be confidently attributed to this cell type, and there is no mix-up with stromal, inflammatory, vascular or other types of cells as long as one carefully avoids, in particular, sample areas with pannus-like tissue or showing secondary cartilage formation. To date, however, few larger (more than 100 genes) gene expression profiling studies of OA and chondrocytes (i.e., looking at more than 100 genes in parallel) have been carried out.

In a study published by Aigner and colleagues in 2000 [3], more than 1000 genes were examined in parallel using the Clontech™ Cancer Arrays. In this study the authors were able to characterize chondrocyte gene expression patterns and were able to relate these patterns to the main function of this cell type, that is, the preservation and turnover of the cartilage matrix. This study confirmed that cartilage collagen expression was largely absent in normal adult cartilage, and mRNA levels of several cartilage collagen expression was largely absent in the cartilage matrix. This study confirmed that this cell type, that is, the preservation and turnover able to relate these patterns to the main function of this enzyme is centrally involved in physiological chondrocyte gene expression patterns and were differentially expressed between normal and osteoarthritic cartilage. This will require more detailed insights into the gene profiles of these phenotypes as they occur in vivo (i.e., during development) or in vitro (e.g., after dedifferentiation in monolayer culture).

Another good example of the use of cDNA array technology to identify novel or interesting molecules is a study by Marshall and colleagues[9]. This group identified β2-microglobulin (B2M) as being differentially upregulated in a large gene screen that compared osteoarthritic with normal cells (results validated by quantitative PCR). The group then evaluated differences in gene expression in chondrocytes after stimulation with B2M. However, no real insights into either the pathophysiology of OA or the role and activity of this molecule in the context of the chondrocyte could be established.

Biomarkers for OA are urgently required for patient management [10]. Biomarkers are also required for the diagnosis of early OA as well as for monitoring disease progress. Conventional diagnosis by x-ray imaging is an insensitive methodology; it does not allow an estimate of short-term disease developments. YKL-40 was one of the candidate biochemical markers suggested to be of value for the diagnosis, prognosis and monitoring of the osteoarthritic and rheumatoid disease process [11,12]. A recent study (Aigner and colleagues) showed that YKL-40 was expressed at high levels in normal cartilage, but was not upregulated in osteoarthritic cartilage. However, YKL-39, a closely related protein, was significantly upregulated in osteoarthritic chondrocytes and might prove to be a more accurate marker for chondrocyte activation in the disease process. YKL-39 was also upregulated in early degenerative cartilage specimens. However, this molecule has yet to be established as a suitable marker for chondrocyte activation when measured in synovial fluid and serum. Unfortunately, the function of both molecules remains unclear at present.

Gene expression profiling of osteoarthritic synoviolpathy

The synovial membrane represents another important tissue in the pathogenesis and progression of arthritis and, particularly, for rheumatoid arthritis [13–17]. But changes in the synovial membrane are also a feature of OA, even for early disease states [18]. Synovial pathology might be an
important factor in the progression of OA and analysis of gene expression patterns of osteoarthritic synovium has great potential. The synovial membrane is pharmacologically easier to tackle than the remote avascular articular cartilage.

A major issue in gene screening studies, however, is the heterogeneity of cell subpopulations in inflamed tissues. Thus, osteoarthritic synovia may show significant infiltrates by T and B cells [18] and a very different gene expression repertoire to the normal synoviocytes, fibroblasts, stromal and vascular cells. This heterogeneity of cellular subpopulations blurs any gene expression profiles obtained. Regardless of this, it will be highly interesting when the first studies are published.

Gene expression profiling of blood cells Unlike articular cartilage and synovial tissues, blood is readily available and can be obtained with little patient discomfort. Blood is considered an ideal surrogate for disease diagnosis, it is a highly dynamic tissue and constantly interacts and communicates with every organ and tissue in the body. In view of the dynamic and interactive properties of circulating blood, blood cells may function as 'sentinels', which can reflect the status of the current state of health or disease throughout the body [101]. This Sentinel Principle was first demonstrated in the study by Ma and Liew [19]. This study successfully identified differentially expressed genes in the blood from healthy control subjects and from coronary artery disease patients [19]. A number of subsequent studies have also independently explored and verified the use of blood-derived RNA expression to diagnose disease in a broad range of conditions [20], most notably neuro-pathological conditions such as schizophrenia and bipolar disorder [20].

Most recently, Marshall and colleagues investigated the possible utility of using a blood gene expression profiling approach for diagnosing subjects with mild OA of the knee [21]. In this study, a total of 68 subjects (29 controls and 39 arthroscopically diagnosed mild knee OA patients) were profiled using a custom cDNA microarray to identify differentially expressed genes (Figure 1).

A subset of genes was then assessed by real-time PCR. Logistic regression was used to generate linear combinations of biomarkers. Linear combinations of nine genes, early growth response 1 (EGR1), α-glucosidase II subunit (G2AN), heat shock 90 kDa protein 1-α (HSPCA), inhibitor of κ-light polypeptide gene enhancer in B cells; kinase complex-associated protein (IKB KAP), interleukin-13 receptor α-1 (IL13RA1), laminin γ-1 (LAM C1), v-maf musculoaponeurotic fibrosarcoma oncogene homolog B (MAFB), platelet factor 4 (PF4) (also known as chemokine [C-X-C motif]), tumor necrosis factor-α-induced protein 6 (TNF–AIP6), (or TSG–6), yielded Receiver Operating Characteristic (ROC) curve areas of 0.90 or more for a reference set of 78 samples (52 samples used in microarray and 26 additional samples). Furthermore, these biomarker combinations gave a sensitivity of 72% and a specificity of 66% when tested against a blind set of an additional 67 samples (32 mild knee OA and 35 controls). This study indicates that a panel of gene (mRNA) expression of peripheral blood cells might have clinical utility for detecting mild OA of the knee [21].

Interestingly, of the nine candidate blood biomarker genes identified, three had been linked previously to OA cartilage biology: TNF–AIP6 (or TSG–6), IL13RA1 and EGR1 [22,24]. TNF–AIP6 was shown to be chondroprotective in an antigen-induced model of arthritis [25]. Moreover, in a recent study reported by Adarichev and colleagues, TNF–AIP6 was one of the genes most significantly upregulated in the arthritis gene signature of pre-inflamed joints [26]. This gene was also found to be upregulated in human OA cartilage reported by Zhang and colleagues [27]. However, TNF–AIP6 expression was downregulated in early OA blood samples [21]. The downregulation of TNF–AIP6 in OA blood probably reflects systemic rather than local changes of the OA disease. Blood biomarker IL13RA1 is one of the functional receptor components of IL-13 in blood monocytes [28] and is critical for IL-13 expression in response to inflammation. IL13 has been suspected of being protective against cartilage dysfunction [22,29]. EGR1, the third blood biomarker, is a transcriptional repressor of collagen Type II promoter activity [30]. EGR1 was reported to be downregulated in human OA cartilage [24]. In concordance, EGR1 expression was also downregulated in OA blood.

The fact that some of the blood biomarkers have known associations with cartilage biology further supports the hypothesis that blood cell gene expression could reflect the pathophysiological process in an organ such as articular cartilage. However, the identified blood biomarkers that have not been previously associated with OA cartilage biology could play a role in OA. The blood gene expression changes could be a systemic change or downstream response of the disease. It is hard to postulate at this point whether the change observed in peripheral blood is the cause or effect of the early disease process.
The authors’ study represents the first application of profiling blood gene expression to OA. It demonstrates that the panel of identified blood biomarkers is relatively sensitive (72%) and specific (66%) for early knee OA. However, the panel was not tested against other arthritic disease, such as rheumatoid arthritis. The blind test sample size is still small and does not reflect a typical population. Nonetheless, as more genes and more samples will be tested, more potential biomarkers will be identified.

Compared with articular cartilage and synovium, blood is easily available. Gene expression profiling of blood will not only complement our understanding of OA research on articular cartilage and synovium, but it also holds great potential to be the source of diagnosing, classifying and staging disease, and monitoring disease progress and therapeutic response. Further studies are needed to address the clinical utility of blood gene expression for the diagnosis of OA.

Pathogenesis, diagnostics & therapeutics

New markers are required in order to characterize cellular behavior during degenerative joint diseases and to characterize and classify the disease process itself. New biological markers will also help the development of approaches to stop, delay, or even to reverse cartilage degeneration. Increasing our understanding of the
pathogenesis of OA and gaining new angles to tackle the disease therapeutically is the focus of current studies in functional genomics. This technology also has the potential for improving diagnosis and staging of disease. Blood cells have the advantage of being easily accessible. However, synovial and cartilage specimens are also of potential use in profiling, provided the amount necessary for analysis is reduced to minor bioptical specimens in the future. This may allow the identification of subgroups of patients enabling more individualized therapeutic approaches. Patients with matrix degeneration might comprise one group showing high protease expression and normal (or even increased) matrix component synthesis, while another group may show normal protease expression but reduced levels of aggrecan resynthesis. The ability to subgroup patients would have obvious important implications for the prescription and treatment with protease inhibitors and with anabolic stimulants, respectively.

Interindividual variability might at first sight seem disadvantageous for statistical evaluation and the establishment of general rules. However, clustering may enable the identification and prediction of responders and nonresponders within defined therapeutic regimens, and may help reduce the size of clinical trials. Recently, bioinformatic clustering of

![Figure 2. Schematic representing a potential workflow of functional genomic analysis.](https://www.futuremedicine.com)

Given reliable data on gene expression levels, the very challenge of functional genomics has to be faced: how to gain biologically meaningful information and knowledge, i.e., real understanding of biological systems and diseases, from a mass of data [2]. Fundaments of such efforts can be pre-existing knowledge on relevant biological systems (genes of known relevance), ranking genes along biostatistical significance (p-value approach), the comparison of gene expression patterns to in vivo (evo–devo approach) [37] or in vitro models, bioinformatical tools (clustering and neighborhood analyses), as well as an envisioned combination of both pre-existing knowledge and computational processing of obtained data (i.e., bioinformatical pathway modelling) [38]. Final confirmation of a revealed pattern will require functional validation of the genes and molecular networks identified.
samples was able to establish diagnostic algorithms to distinguish between acute myeloid leukemia and acute lymphatic leukemia [31].

Although this is only a first step, it demonstrates the potential of such methods to identify subsets of diseases and might show potential in the field of degenerative joint disease. Diagnosis, in particular of early disease stages, has been an issue for research and for the management of patients suffering over decades from osteoarthritic cartilage degeneration. Despite numerous studies, biochemical markers detected in body fluids are far from being of reliable diagnostic value [32-34]. One drawback of all the methods utilized so far has been that for technical reasons, it has only been possible to screen for a few genes or proteins. The use of a single marker for diagnosis is an unlikely goal; in most areas of medicine a broad panel of diagnostic parameters is usually applied. To take a core needle biopsy for multigene analysis of this size of tissue might well be justifiable, if the results offer significant benefits in terms of diagnosis and treatment to the patient.

Conclusions
Despite limitations such as low sensitivity and insensitivity to alternative splicing, post-transcriptional regulation and post-translational modification, cDNA array technology is a powerful tool to obtain an overview of gene expression pattern; this is not achievable with other techniques. This technology can identify known genes as well as new genes of potential interest. Computational tools for evaluating data obtained in gene expression screens need to be improved considerably with respect to biostatistical analysis; furthermore, bioinformatical methods are needed to enable the accumulation of gene expression networks illustrating biological systems on the cell and tissue level. Another important issue will be the

Figure 3. Gene expression profiling hints to gene groups involved in the osteoarthritic disease process.

Gene expression profiling reveals a large amount of up- and down-regulated genes in a given context. One way to gain insight is to cluster differentially expressed genes on gene groups of interest: this shows significant regulatory patterns of anabolic, catabolic and phenotype-related genes – as to be expected from the literature – as well as genes related to cell integrity and degeneration [Aigner et al., Unpublished Data].
Combining mRNA expression analysis with proteomics, interactomics and all other levels of biological investigation [35].

Even more importantly, gene expression data (and also proteomics data) need to be embedded in the context of the physiology and pathobiology of cells and tissues. For OA research, this will allow the interpretation of gene regulation within the context of the disease to promote our understanding rather than providing only or mostly large amounts of ‘naked’ data. Thus, the major challenge of modern large-scale (functional) genomics in the field of OA research, as elsewhere, is how to channel abundantly revealed ‘data’ into useful ‘information’ on genes and molecular pathways and how to finally achieve knowledge in terms of pathogenetic concepts allowing the identification of therapeutic targets (a typical workflow is shown in Figure 2). It is well known that the cellular reaction pattern of chondrocytes in osteoarthritic cartilage degeneration is poorly understood, mainly because most of the involved genes are not yet identified. cDNA array technology has the power to change this.

### Executive summary

#### Gene expression & osteoarthritis
- The molecular and genetic phenotype of chondrocytes plays a key role in the initiation, progression and outcome of osteoarthritis (OA).
- Gene expression profiling represents an innovative and challenging approach to identify new disease-relevant genes.
- In addition to chondrocyte gene expression patterns, the profiling of the synovial membrane is also of high importance.

#### General issues
- Further advances in technical (e.g., sensitivity and specificity), statistical, bioinformatic and disease-related issues are required to realize the full potential of functional genomics.

#### Gene expression profiling of cartilage degeneration
- The great advantage of human articular cartilage is that it contains – both in the normal and in the diseased state – only one cell type, the chondrocyte. Thus, the detected gene expression levels can be attributed fully to chondrocytes.
- The first published gene profiling studies were able to characterize the chondrocyte gene expression pattern related to the main function of chondrocytes, the preservation and turnover of the cartilage matrix; they confirmed the largely absent expression of cartilage collagens in normal adult cartilage and very much increased mRNA levels of several collagen genes in advanced OA. Important upregulated proteases were matrix-degrading metalloproteinases-2 (gelatinase A) and -13 (collagenase 3).
- Other studies added additional regulated genes, such as β2-microglobulin.
- Also, potential new biological markers were identified, such as YKL-39.

#### Gene expression profiling of OA synovioathy
- Analysis of the gene expression pattern of OA synovium has great potential, in particular because the synovial membrane is pharmacologically easier to tackle than the remote, avascular, articular cartilage.
- The major issue of gene screen studies, however, is the heterogeneity of cell subpopulations in inflamed tissues (e.g., T and B cells), which have a very different gene expression repertoire than the usually present synoviocytes, fibroblasts, stromal and vascular cells, for example. This heterogeneity of cellular subpopulations certainly blurs the gene expression profiles obtained. Still, it will be highly interesting when the first studies are published.

#### Gene expression profiling of blood cells
- Blood, due to its availability, is an ideal surrogate for disease diagnosis. A first study suggests that gene profiling of blood cells is a feasible and powerful tool to identify and classify patient groups, even in early OA.

#### Pathogenesis, diagnostics & therapeutics
- For the understanding of the pathogenesis of degenerative joint disease, new markers are required in order to analyze further OA and to develop therapeutic approaches.
- The identification of subgroups of patients by gene expression profiling might enable more individualized therapeutic approaches and the identification and prediction of responders and nonresponders within defined therapeutic regimens, which will allow the reduction of clinical trial size.

#### Conclusions
- Despite limitations such as a low sensitivity and insensitivity to alternative splicing, post-transcriptional regulation and post-translational modification, cDNA array technology provides a powerful tool to obtain an overview of gene expression pattern that is not achievable with other techniques.
Future perspective

Future attempts to interpret gene expression data will need to focus more on understanding signaling pathways and gene networks related to defined disease hypotheses [36]. A first step towards this goal is to define hypotheses and to decide whether they are suitable for testing within the data set available. A next step is to define the genes likely to be indicative of disease processes fitting to the hypothesis (a rough approach is presented in terms of detected pattern in Figure 3). For example, the traditional disease hypotheses of OA has been that OA is a failure of chondrocyte anabolic activity, largely attributable to hypercatabolism of cartilage matrix molecules; or alternatively, that OA is caused by changes in the subchondral bone or in the synovial membrane. The first hypothesis is amenable to analyses based on gene expression data obtained from normal and osteoarthritic cartilage tissue, this might be impossible or difficult for the latter.

Identification, classification and staging of disease processes using large-scale gene expression analysis will become an important option in the future and may allow the identification of subgroups of patients enabling more individualized therapeutic approaches as discussed above. The near future will ascertain technical versatility and comparability in the whole area of gene expression profiling. Besides adding to previous studies, which provided interesting data on single or few genes, functional genomics will provide an overall picture of genes involved in the osteoarthritic disease process. Molecular portraits of these processes will stimulate the testing of new markers, which are desperately needed for the diagnosis and monitoring of the disease, and outline new targets for therapeutic intervention.

Acknowledgements

We acknowledge the German Ministry of Research for financial support (grant 01GG9824).

Bibliography

Papers of special note have been highlighted as either of interest (•) or of considerable interest (••) to readers.

15. Important best review on gene expression profiling technology, with many interesting insights regarding its application to arthritis research.


• Interesting overview of strategies, potentials and limitations of gene expression profiling in rheumatoid arthritis.


• Landmark publication in the new field of identifying markers for degenerative joint diseases in blood cells.


• Very interesting review on gene profiling approaches related to arthritis research; in particular, the cross-linking to other methodological platforms such as proteomics is made.

109